Геометрические методы в математической физике 13 мая 2012 г.

Катанаев Михаил Орионович[1]

Математический институт имени В. А. Стеклова
Российской Академии Наук

13 мая 2012 г.

[1]Любые замечания, указания на ошибки, неточности и опечатки прошу отправлять на e-mail: katanaev@mi.ras.ru

Оглавление

1 Основные обозначения и соглашения **10**

2 Введение **1**
2.1 Множества . . . 1
2.2 Поле вещественных чисел $\mathbb{R}$ и прямая . . . 4
2.3 Евклидово пространство $\mathbb{R}^n$ . . . 12
2.3.1 $\mathbb{R}^n$ как метрическое пространство . . . 12
2.3.2 $\mathbb{R}^n$ как топологическое пространство . . . 17
2.3.3 $\mathbb{R}^n$ как векторное пространство . . . 28
2.3.4 $\mathbb{R}^n$ как аффинное пространство . . . 49
2.4 Отображения . . . 54
2.4.1 Отображения топологических пространств . . . 58
2.5 Преобразования координат . . . 63
2.6 Группа двумерных вращений $\mathbb{O}(2)$ . . . 66
2.7 Группа двумерных преобразований Лоренца $\mathbb{O}(1,1)$ . . . 71
2.8 Трехмерное евклидово пространство . . . 79
2.9 Пространство Минковского . . . 88
2.9.1 Группа Пуанкаре . . . 90
2.9.2 Группа Галилея . . . 96
2.9.3 Группа конформных преобразований . . . 98
2.9.4 Трехмерное пространство Минковского . . . 100
2.9.5 Четырехмерное пространство Минковского . . . 103
2.10 Специальная теория относительности . . . 108
2.10.1 Нерелятивистские модели . . . 108
2.10.2 Релятивистские модели . . . 111
2.10.3 Замедление времени и лоренцево сокращение . . . 116
2.10.4 Сложение скоростей и эффект Доплера . . . 119
2.10.5 Равноускоренное движение . . . 123

3 Многообразия и тензорные поля **128**
3.1 Многообразие . . . 128
3.2 Разбиение единицы . . . 137
3.3 Многообразия с краем . . . 139
3.4 Расслоения . . . 141
3.5 Скалярные поля и плотности . . . 144
3.6 Векторные поля и 1-формы . . . 147
3.6.1 Локальное определение . . . 147
3.6.2 Глобальное определение векторных и ковекторных полей . . . 150
3.6.3 Кокасательные векторные поля и ростки . . . 153

3.6.4 Векторные поля и дифференцирования . . . 154
3.6.5 Векторные поля и интегральные кривые . . . 158
3.6.6 1-формы и гиперповерхности . . . 169
3.6.7 Алгебра Ли векторных полей . . . 170
3.7 Тензорные поля . . . 172
3.8 Полностью антисимметричные тензоры . . . 178
3.9 Отображения многообразий . . . 180
3.10 Подмногообразия . . . 185
3.11 Теорема Фробениуса . . . 190
3.12 Слоения . . . 196
3.13 Бесконечно малые преобразования координат . . . 201
3.14 Производная Ли . . . 205

4 Дифференциальные формы и интегрирование 209
4.1 Внешняя алгебра . . . 209
4.2 Дифференциальные формы . . . 216
4.3 Внешнее дифференцирование . . . 219
4.4 Теорема Дарбу . . . 224
4.5 Оператор Лапласа–Бельтрами . . . 229
4.6 Интегрирование дифференциальных форм . . . 234
4.6.1 Форма объема . . . 236
4.6.2 Формула Стокса . . . 239

5 Метрика 243
5.1 Определение и свойства . . . 243
5.2 Метрика на лоренцевых многообразиях . . . 247
5.3 Векторные поля и вложения . . . 253
5.4 Выбор системы координат . . . 255
5.5 Разложение Ходжа . . . 257

6 Связность на векторном расслоении и расслоении реперов 260
6.1 Векторные расслоения . . . 260
6.2 Связность на векторном расслоении . . . 264
6.3 Аффинная связность . . . 270
6.4 Связность на расслоении реперов . . . 272

7 Аффинная геометрия. Локальное рассмотрение 280
7.1 Локальное определение аффинной связности . . . 280
7.2 Кручение и неметричность . . . 283
7.3 Ковариантная производная тензорных плотностей . . . 287
7.4 Параллельный перенос . . . 289
7.5 Геометрический смысл кручения . . . 293
7.6 Свойства аффинной связности . . . 295
7.7 Локальное определение тензора кривизны . . . 298
7.8 Свойства тензора кривизны . . . 300
7.9 Неголономный базис . . . 304
7.10 Тождества Бианки . . . 308

8 Криволинейные координаты в $\mathbb{R}^3$ 310
8.1 Сферические координаты . . . 310
8.2 Цилиндрические координаты . . . 315

9 Группы Ли 318
9.1 Группы Ли и локальные группы Ли . . . 318
9.2 Действие группы слева . . . 324
9.3 Действие группы справа . . . 331
9.4 Присоединенное представление . . . 332
9.5 Группы Ли как римановы пространства . . . 335
9.6 Группы Ли как пространства Римана–Картана . . . 339
9.7 Двумерная неабелева группа Ли . . . 340
9.8 Гомоморфизмы групп Ли . . . 343
9.9 Экспоненциальное отображение для групп Ли . . . 348
9.10 Интегрирование на группах Ли . . . 350
9.11 Некоторые общие свойства групп и алгебр Ли . . . 351
9.12 Группа Ли $\mathbb{GL}(n,\mathbb{C})$ . . . 355
9.12.1 Алгебра Ли $\mathfrak{gl}(n,\mathbb{C})$ . . . 355
9.12.2 Группа Ли $\mathbb{GL}(n,\mathbb{C})$ . . . 361

10 Группы преобразований 369
10.1 Действие групп преобразований . . . 369
10.2 Инфинитезимальные преобразования . . . 375
10.3 Инвариантные структуры . . . 377
10.4 Отображения групп преобразований . . . 380

11 Гомотопии и фундаментальная группа 382
11.1 Пути . . . 382
11.2 Гомотопия непрерывных отображений . . . 387
11.3 Фундаментальная группа . . . 390
11.4 Фундаментальная группа и ориентируемость . . . 395

12 Накрытия 402
12.1 Определения и примеры . . . 402
12.2 Фундаментальная группа пространства орбит . . . 406
12.3 Группа скольжений и существование накрытий . . . 409

13 Главные и ассоциированные расслоения 412
13.1 Главные расслоения . . . 412
13.2 Ассоциированные расслоения . . . 421
13.3 Отображение расслоений . . . 429

14 Связности на главных и ассоциированных расслоениях 434
14.1 Связность на главном расслоении . . . 434
14.2 Форма кривизны и структурное уравнение . . . 445
14.3 Параллельный перенос . . . 454
14.4 Группы голономии . . . 455
14.5 Петля Вильсона . . . 458
14.6 Отображение связностей . . . 461
14.7 Связность на ассоциированном расслоении . . . 464

14.8 Свойства групп голономий . . . 467
14.9 Плоские связности . . . 469
14.10Локальные и инфинитезимальные группы голономии . . . 472
14.11Инвариантные связности . . . 476

15 Приложения в квантовой механике 481
15.1 Адиабатическая теорема . . . 482
15.1.1 Двухуровневая система . . . 490
15.2 Фаза Берри . . . 493
15.2.1 Абелев случай: невырожденное состояние . . . 493
15.2.2 Частица со спином 1/2 в магнитном поле . . . 495
15.2.3 Неабелев случай: вырожденное состояние . . . 498
15.3 Эффект Ааронова–Бома . . . 501
15.3.1 Электрический потенциал . . . 503
15.3.2 Магнитный потенциал . . . 505

16 Векторные поля Киллинга 509
16.1 Изометрии и инфинитезимальные изометрии . . . 509
16.2 Однородные и изотропные многообразия . . . 513
16.3 Свойства векторных полей Киллинга . . . 519
16.4 Лоренц-инвариантные метрики . . . 521

17 Геодезические и экстремали 526
17.1 Геодезические . . . 526
17.2 Экстремали . . . 533
17.3 Интегрирование уравнений для экстремалей и геодезических . . . 537
17.4 Вторая вариация уравнений для экстремалей . . . 538
17.5 Уравнение Гамильтона–Якоби для экстремалей . . . 542
17.6 Волновое уравнение . . . 544
17.7 Приближение эйконала . . . 547
17.8 Гармонические координаты . . . 550
17.9 Нормальные, геодезические или римановы координаты . . . 551
17.9.1 Нормальные координаты. Локальное рассмотрение. . . . 552
17.9.2 Нормальные координаты в (псевдо-)римановом пространстве . 558
17.9.3 (Псевдо-)римановы пространства постоянной кривизны . . . 559
17.9.4 Нормальные координаты и экспоненциальное отображение . . . 564
17.10Полнота римановых многообразий . . . 566
17.11Формулы Френе . . . 569

18 Симплектические и пуассоновы многообразия 572
18.1 Симплектические группы . . . 572
18.2 Симплектические многообразия . . . 577
18.3 Пуассоновы многообразия . . . 582
18.4 Структура Ли–Пуассона . . . 587
18.5 Отображения пуассоновых многообразий . . . 588

19 Принцип наименьшего действия 594
19.1 Постановка вариационных задач . . . 594
19.1.1 Задача с заданными граничными условиями . . . 595
19.1.2 Задача со свободными граничными условиями . . . 597
19.1.3 Задача с подвижной границей . . . 598
19.1.4 Задача на условную стационарную точку . . . 600
19.1.5 Другие задачи и терминология . . . 602
19.2 Первая теорема Нетер . . . 604
19.2.1 Тензор энергии-импульса . . . 606
19.2.2 Тензор момента количества движения . . . 608
19.3 Вторая теорема Нетер . . . 609
19.4 Эффективное действие . . . 612
19.5 Редуцированное действие . . . 614

20 Канонический формализм 616
20.1 Канонический формализм в механике точечных частиц . . . 616
20.1.1 Преобразование Лежандра . . . 616
20.1.2 Гамильтонова динамика точечных частиц . . . 618
20.1.3 Потенциальное движение точечной частицы . . . 624
20.1.4 Лемма Стокса . . . 627
20.1.5 Канонические уравнения Гамильтона . . . 628
20.1.6 Принцип Мопертюи . . . 630
20.1.7 Уравнение Гамильтона–Якоби . . . 634
20.1.8 Принцип Гюйгенса . . . 636
20.1.9 Переменные действие-угол . . . 639
20.1.10 Канонические преобразования . . . 642
20.1.11 Производящие функции канонических преобразований . . . 650
20.1.12 Разделение переменных в уравнении Гамильтона–Якоби . . . 658
20.2 Гамильтонова динамика частиц со связями . . . 661
20.2.1 Связи в гамильтоновом формализме . . . 661
20.2.2 Гамильтонова динамика частиц со связями II рода . . . 663
20.2.3 Гамильтонова динамика частиц со связями I рода . . . 669
20.2.4 Калибровочная модель нерелятивистской частицы . . . 675
20.2.5 Частица в псевдоримановом пространстве . . . 678
20.2.6 Граничные слагаемые в калибровочных моделях . . . 687

21 Основы общей теории относительности 695
21.1 Пространство-время, метрика и гравитация . . . 695
21.2 Действие Гильберта–Эйнштейна . . . 698
21.3 Вариация действия Гильберта–Эйнштейна . . . 703
21.4 Зависимость уравнений Эйнштейна . . . 705
21.5 Действие для полей материи в обобщенных моделях гравитации . . . 707
21.6 Скалярно-тензорные модели . . . 710
21.7 Полиномиальная форма действия Гильберта–Эйнштейна . . . 711
21.8 Точечные частицы в теории гравитации . . . 716
21.8.1 Нерелятивистский предел для точечной частицы . . . 723
21.8.2 Теория гравитации Ньютона . . . 725
21.8.3 Свойства тензора энергии-импульса точечных частиц . . . 727

21.9 Ньютонов предел . . . 731
21.10 Гравитационные волны . . . 735
21.11 Сплошная среда в общей теории относительности . . . 738
21.11.1 Акустические фононы в нерелятивистской гидродинамике . . . 744
21.12 Выбор системы координат . . . 749
21.12.1 Сопутствующая система координат . . . 749
21.12.2 Временна́я калибровка . . . 755
21.12.3 Калибровка светового конуса . . . 764
21.13 О постановке задач в теории гравитации . . . 771

22 Гамильтонова формулировка общей теории относительности 774
22.1 Лагранжиан Гильберта–Эйнштейна . . . 775
22.2 АДМ параметризация метрики и репера . . . 776
22.3 Геометрия гиперповерхностей . . . 782
22.4 Кривизна в АДМ параметризации метрики . . . 787
22.5 Гамильтониан . . . 791
22.6 Вторичные связи . . . 793
22.7 Полиномиальная гамильтонова форма . . . 796
22.8 Проблема энергии в теории гравитации . . . 802
22.8.1 Тензор энергии-импульса полей материи . . . 802
22.8.2 Псевдотензор энергии-импульса для гравитации . . . 803
22.8.3 Полная гравитационная энергия асимптотически плоского пространства-времени . . . 806

23 Скалярные и калибровочные поля 810
23.1 Действительное скалярное поле . . . 810
23.1.1 Скалярное поле в пространстве Минковского . . . 810
23.1.2 Скалярное поле в аффинной геометрии . . . 814
23.1.3 Скалярное поле в общей теории относительности . . . 816
23.2 Комплексное скалярное поле . . . 819
23.2.1 Комплексное скалярное поле в пространстве Минковского . . . 819
23.2.2 Комплексное скалярное поле в аффинной геометрии . . . 823
23.2.3 Комплексное скалярное поле в общей теории относительности . . . 824
23.3 Электромагнитное поле . . . 825
23.3.1 Лагранжева формулировка . . . 829
23.3.2 Законы сохранения . . . 834
23.3.3 Гамильтонова формулировка . . . 835
23.3.4 Скалярная электродинамика . . . 843
23.3.5 Электромагнитное поле в аффинной геометрии . . . 846
23.3.6 Электромагнитное поле в общей теории относительности . . . 849
23.4 Поле Прок´а . . . 854
23.4.1 Гамильтонова формулировка . . . 856
23.5 Поле Янга–Миллса . . . 860
23.5.1 Лагранжева формулировка . . . 866
23.5.2 Гамильтонова формулировка . . . 868
23.5.3 Поля Янга–Миллса в аффинной геометрии . . . 873
23.5.4 Поле Янга–Миллса в общей теории относительности . . . 874

24 Геометрия поверхностей **877**
24.1 Геометрия Римана–Картана на поверхности . . . 877
24.2 Неголономный базис в двух измерениях . . . 880
24.3 Выбор системы координат . . . 881
24.3.1 Диагональная калибровка . . . 881
24.3.2 Конформная калибровка для римановых поверхностей . . . 882
24.3.3 Конформная калибровка для лоренцевых поверхностей . . . 883
24.4 Координаты светового конуса . . . 886

25 Поверхности постоянной кривизны **889**
25.1 Сфера $\mathbb{S}^2$ . . . 889
25.2 Двуполостный гиперболоид $\mathbb{H}^2$ . . . 893
25.3 Однополостный гиперболоид $\mathbb{L}^2$ . . . 898
25.4 Уравнение Лиувилля . . . 905

26 Поверхности с одним вектором Киллинга **911**
26.1 Локальный вид лоренцевой метрики . . . 912
26.2 Конформные блоки . . . 916
26.3 Экстремали . . . 918
26.3.1 Форма экстремалей . . . 918
26.3.2 Асимптотика экстремалей . . . 923
26.3.3 Полнота экстремалей . . . 924
26.4 Построение глобальных решений . . . 926
26.5 Примеры . . . 929
26.5.1 Решение Шварцшильда . . . 929
26.5.2 Решение Рейснера–Нордстрема . . . 931
26.5.3 Экстремальная черная дыра . . . 932
26.5.4 Плоскость Минковского . . . 932
26.5.5 Поверхности постоянной кривизны . . . 935
26.6 Координаты Эддингтона–Финкельстейна . . . 938
26.7 Дифференцируемость метрики в седловой точке . . . 941
26.8 Локальный вид римановой метрики . . . 943
26.9 Экстремали . . . 945
26.10Построение глобальных решений . . . 949
26.11Решение Шварцшильда . . . 953

27 Сплетенные решения в общей теории относительности **955**
27.1 Сплетенное произведение . . . 955
27.2 Двумерная редукция . . . 956
27.3 Произведение поверхностей постоянной кривизны . . . 959
27.4 Пространственно симметричные решения . . . 960
27.4.1 Сферически симметричные решения $K = 1$ . . . 966
27.4.2 Планарные решения $K = 0$ . . . 972
27.4.3 Гиперболические глобальные решения $K = -1$ . . . 976
27.5 Лоренц-инвариантные решения . . . 977
27.5.1 Лоренц-инвариантные решения $K = 1$ . . . 980
27.5.2 Решения с плоскостью Минковского $K = 0$ . . . 986
27.6 Итоги главы . . . 987

28 Дополнение **989**
28.1 Матрицы 989
28.2 Матрицы Паули 997
28.3 Кватернионы 998
28.4 Полностью антисимметричные тензоры 1000

Глава 1

Основные обозначения и соглашения

Дифференцирование выполняется раньше алгебраических операций.

$\mathbb{N} = \{1, 2, \dots\}$	– множество натуральных чисел,
$\mathbb{Z} = \{\dots -1, 0, 1, \dots\}$	– группа целых чисел по сложению,
$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$	– поле рациональных, вещественных, комплексных чисел,
$\mathbb{R}_+$	– множество положительных вещественных чисел,
$\dagger$	– комплексное или эрмитово сопряжение,
$:=$	– равно по-определению,
x и $(x^1, \dots, x^n)$	– точка многообразия и ее координаты,
$\{x^\alpha\} = \{x^0, x^\mu\} = \{x^0, \boldsymbol{x}\}$	– декартовы координаты в пространстве Минковского, или координаты на псевдоримановом многообразии,
$\partial_\alpha := \frac{\partial}{\partial x^\alpha}$	– частная производная,
$\partial^2_{\alpha\beta} := \frac{\partial^2}{\partial x^\alpha \partial x^\beta}$	– частная производная второго порядка,
∇_α	– ковариантная производная,
$\triangle$	– оператор Лапласа или Лапласа–Бельтрами,
$\Box$	– оператор Даламбера, конец доказательства, примера или определения,
$(\partial f)^2 := g^{\alpha\beta}\partial_\alpha f \partial_\beta f$	– квадрат градиента функции f,
$g_{\alpha\beta}$	– компоненты метрики,
$g := \mathsf{det}\,(g_{\alpha\beta})$	– определитель метрики,
${e_\alpha}^a$	– компоненты репера,
$\sqrt{\lvert g\rvert} := \mathsf{det}\,({e_\alpha}^a)$	– элемент объема (определитель репера),
$\upsilon := dx^1 \wedge \dots \wedge dx^n \sqrt{\lvert g\rvert}$	– форма объема на многообразии $\mathbb{M}$,
e	– единица группы,
e	– основание натурального логарифма,
${\Gamma_{\alpha\beta}}^\gamma$	– компоненты аффинной связности,
${\omega_{\alpha a}}^b$	– компоненты линейной или лоренцевой связности,
$\mathsf{sgn}\,\sigma(\alpha_1, \dots, \alpha_n)$	– знак перестановки σ индексов $\alpha_1, \dots, \alpha_n$.

Некоторые многообразия и классы объектов имеют специальные обозначения:

$\mathbb{R}^n$ – n-мерное евклидово пространство,
$\mathbb{R}^n_+$ – полупространство в $\mathbb{R}^n$, определяемое условием $x^n > 0$,
$\mathbb{R}^{1,n-1}$ – n-мерное пространство Минковского,
$\mathbb{S}^n_r$ – n-мерная сфера радиуса r,
$\mathbb{B}^n_r$ – n-мерный шар радиуса r.
$\mathcal{C}^k(\mathbb{M})$ – класс функций на многообразии $\mathbb{M}$, непрерывных вместе с производными вплоть до k-го порядка,
$\mathcal{X}(\mathbb{M})$ – множество гладких векторных полей на многообразии $\mathbb{M}$,
$\mathcal{T}(\mathbb{M})$ – множество гладких тензорных полей на многообразии $\mathbb{M}$,
$\mathbb{M} \approx \mathbb{N}$ – многообразие $\mathbb{M}$ диффеоморфно (гомеоморфно) многообразию $\mathbb{N}$,
$\mathbb{G} \simeq \mathbb{H}$ – группа (алгебра, векторное пространство, ...) $\mathbb{G}$ изоморфна группе (алгебре, векторному пространству, ...) $\mathbb{H}$,
$\mathbb{M} \to \mathbb{N}$ – отображение множеств,
$\mathbb{M} \ni a \mapsto b \in \mathbb{N}$ – отображение элементов множеств,

Антисимметризация по индексам обозначается квадратными скобками:

$$A^{[ab]} := \frac{1}{2}(A^{ab} - A^{ba}),$$
$$A^{[abc]} := \frac{1}{6}(A^{abc} + A^{bca} + A^{cab} - A^{bac} - A^{acb} - A^{cba}).$$

В общем случае, когда имеется n индексов, сумма берется по всем $n!$ перестановкам и делится на $n!$. При этом четные перестановки индексов входят со знаком плюс, а нечетные – со знаком минус.

Симметризация индексов обозначается круглыми скобками:

$$A^{(ab)} := \frac{1}{2}(A^{ab} + A^{ba}),$$
$$A^{(abc)} := \frac{1}{6}(A^{abc} + A^{bca} + A^{cab} + A^{bac} + A^{acb} + A^{cba}).$$

Евклидова метрика δ_{ab} имеет два нижних индекса и равна единичной матрице. Например, в четырехмерном евклидовом пространстве $\mathbb{R}^4$

$$\delta_{ab} := \mathsf{diag}\,(++++) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.1}$$

Символ Кронекера δ^b_a является тождественным оператором, действующим в векторном пространстве, и также равен единичной матрице. В $\mathbb{R}^4$

$$\delta^b_a := \mathsf{diag}\,(++++) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.2}$$

Для краткости, произведение символов Кронекера обозначается одним символом

$$\delta^{ab\ldots d}_{ef\ldots h} := \delta^a_e \delta^b_f \ldots \delta^d_h. \tag{1.3}$$

Иногда используется обобщенный символ Кронекера, помеченный шляпкой,

$$\hat{\delta}^{ab\ldots d}_{ef\ldots h} := n!\delta^{[a}_{\ e}\delta^b_f \ldots \delta^{d]}_h = n!\delta_{[e}^{\ \ a}\delta^b_f \ldots \delta^d_{h]}, \tag{1.4}$$

который получается из произведения (1.3) антисимметризацией по верхним или нижним индексам (что эквивалентно).

Метрика Минковского в четырехмерном пространстве-времени $\mathbb{R}^{1,3}$ имеет вид

$$\eta_{ab} := \mathsf{diag}\,(+---) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.5}$$

Каноническая симплектическая форма в евклидовом пространстве $\mathbb{R}^{2n}$ имеет вид

$$\varpi := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \tag{1.6}$$

где $\mathbb{1}$ – единичная $n \times n$ матрица.

Готический шрифт

$\mathfrak{A}\ \mathfrak{a}$	$\mathfrak{B}\ \mathfrak{b}$	$\mathfrak{C}\ \mathfrak{c}$	$\mathfrak{D}\ \mathfrak{d}$	$\mathfrak{E}\ \mathfrak{e}$	$\mathfrak{F}\ \mathfrak{f}$	$\mathfrak{G}\ \mathfrak{g}$	$\mathfrak{H}\ \mathfrak{h}$	$\mathfrak{I}\ \mathfrak{i}$
$\mathfrak{J}\ \mathfrak{j}$	$\mathfrak{K}\ \mathfrak{k}$	$\mathfrak{L}\ \mathfrak{l}$	$\mathfrak{M}\ \mathfrak{m}$	$\mathfrak{N}\ \mathfrak{n}$	$\mathfrak{O}\ \mathfrak{o}$	$\mathfrak{P}\ \mathfrak{p}$	$\mathfrak{Q}\ \mathfrak{q}$	$\mathfrak{R}\ \mathfrak{r}$
$\mathfrak{S}\ \mathfrak{s}$	$\mathfrak{T}\ \mathfrak{t}$	$\mathfrak{U}\ \mathfrak{u}$	$\mathfrak{V}\ \mathfrak{v}$	$\mathfrak{W}\ \mathfrak{w}$	$\mathfrak{X}\ \mathfrak{x}$	$\mathfrak{Y}\ \mathfrak{y}$	$\mathfrak{Z}\ \mathfrak{z}$	

Используемый греческий шрифт

α	β	$\Gamma\ \gamma$	$\Delta\ \delta$	ϵ	ε	ζ	η
$\Theta\ \theta$	ϑ	ι	κ	$\Lambda\ \lambda$	μ	ν	$\Xi\ \xi$
o	$\Pi\ \pi$	ϖ	ρ	ϱ	$\Sigma\ \sigma$	ς	τ
$\Upsilon\ \upsilon$	$\Phi\ \phi$	φ	χ	$\Psi\ \psi$	$\Omega\ \omega$		

Глава 2

Введение

После напоминания основных понятий теории множеств и введения обозначений, будет достаточно подробно рассмотрено евклидово пространство с различных точек зрения. Это сделано по двум причинам. Во-первых, чтобы подчеркнуть, что в евклидовом пространстве можно задавать различные структуры, которые в дальнейшем будут обобщаться в аффинной геометрии. Во-вторых, чтобы при изложении аффинной геометрии не напоминать относительно сложные понятия геометрии Евклида. В заключительных главах Введения рассматриваются евклидовы пространства и пространства Минковского низших размерностей, которые играют очень важную роль в приложениях. При написании введения, которое содержит хорошо известный материал, использованы монографии [1–20].

2.1 Множества

В математике некоторые исходные понятия не имеют определения. Эти понятия основаны на интуиции и служат для определения других, более сложных, конструкций. Такими интуитивными понятиями являются *множество* и *элемент множества*. Под множеством понимают произвольную совокупность объектов, которые называются элементами множества. Говорят, что множество состоит из своих элементов. Один из способов задания множества состоит просто в перечислении его элементов.

Пример 2.1.1. Натуральные числа являются множеством, которое будем обозначать следующим образом:

$$\mathbb{N} = \{1, 2, \dots\}.$$

Множество натуральных чисел не определяется и лежит в основе всей математики. □

Пример 2.1.2. Множество целых чисел, включающее также нуль и все отрицательные числа, будем обозначать

$$\mathbb{Z} = \{\dots -2, -1, 0, 1, 2, \dots\}$$

Как правило, мы рассматриваем множество целых чисел $\mathbb{Z}$ как абелеву группу по отношению к сложению. □

Пример 2.1.3. Пусть m, n – произвольные целые числа, причем $n \neq 0$. Тогда множество чисел вида m/n называется множеством рациональных чисел и обозначается $\mathbb{Q}$. □

Пример 2.1.4. Вещественные и комплексные числа также являются множествами. Для них приняты обозначения $\mathbb{R}$ и $\mathbb{C}$. □

Множества натуральных, целых и рациональных чисел являются бесконечными *счетными множествами*, т.е. их элементы можно пронумеровать натуральными числами. Множества вещественных $\mathbb{R}$ и комплексных чисел $\mathbb{C}$ представляют собой примеры *несчетных множеств*, поскольку их элементы невозможно пронумеровать натуральными числами. Множество называется *конечным*, если оно состоит из конечного числа элементов.

Пример 2.1.5. Множество поворотов евклидовой плоскости $\mathbb{R}^2$ на угол $\pi/2$ является конечным и состоит из четырех элементов. Это циклическая группа четвертого порядка. □

В большинстве случаев для обозначения множеств, таких как многообразия, группы, и др. мы будем употреблять латинские буквы, напечатанные ажурным шрифтом, $\mathbb{A}, \mathbb{B}, \mathbb{C}, \ldots$. Иногда будут использоваться также заглавные буквы, напечатанные курсивом $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$.

Синонимами понятий множества и элемента множества являются *пространство* и *точка пространства*. Разница в употреблении терминов множество и пространство сложилась исторически. О множестве векторов говорят, как о векторном пространстве. Множество, снабженное топологией, называют топологическим пространством. Часто, говоря о пространстве, подразумевают, что оно, в отличие от множества, снабжено какой-либо дополнительной структурой, будь то топология или структура векторного пространства. Множество элементов называется также *семейством*. Как правило, термин семейство употребляется тогда, когда его элементами являются некоторые множества.

Если элемент x принадлежит множеству $\mathbb{U}$, то мы пишем $x \in \mathbb{U}$. В противном случае применяется обозначение $x \notin \mathbb{U}$.

Определение. Если каждый элемент множества $\mathbb{U}$ принадлежит также множеству $\mathbb{V}$, то $\mathbb{U}$ является *подмножеством* $\mathbb{V}$. Это обозначается $\mathbb{U} \subset \mathbb{V}$ или $\mathbb{V} \supset \mathbb{U}$. При этом совпадение или равенство множеств $\mathbb{U} = \mathbb{V}$ значит, что $\mathbb{U} \subset \mathbb{V}$ и $\mathbb{V} \subset \mathbb{U}$. Это записывается в виде

$$\mathbb{U} = \mathbb{V} \quad \Leftrightarrow \quad \mathbb{U} \subset \mathbb{V},\ \mathbb{V} \subset \mathbb{U},$$

где стрелка $\Leftrightarrow$ обозначает утверждение “тогда и только тогда”. Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом $\emptyset$. Оно, по-определению, является подмножеством любого множества, $\emptyset \subset \mathbb{U}$, $\forall \mathbb{U}$. Подмножество $\mathbb{U} \subset \mathbb{V}$ называется *собственным*, если оно не пусто и не совпадает со всем $\mathbb{V}$: $\mathbb{U} \neq \emptyset$ и $\mathbb{U} \neq \mathbb{V}$. □

Пример 2.1.6. Множество целых чисел $\mathbb{Z}$ имеет много подмножеств. Среди них есть подмножества, которые делятся без остатка на натуральное число $n \in \mathbb{N}$, отличное от нуля. Это подмножество обозначается $n\mathbb{Z} \subset \mathbb{Z}$ и получается умножением каждого целого числа на n. □

Подмножество элементов $x \in \mathbb{U}$, удовлетворяющих некоторому свойству P обозначается следующим образом

$$\{x \in \mathbb{U} :\ P\},$$

где свойство P, как правило, задается некоторым уравнением или неравенством для элементов x.

Пример 2.1.7. Множество положительных и отрицательных вещественных чисел определяется равенствами:

$$\mathbb{R}_+ := \{x \in \mathbb{R} : \ x > 0\},$$
$$\mathbb{R}_- := \{x \in \mathbb{R} : \ x < 0\}.$$ □

Определение. *Объединение, пересечение* и *разность* двух множеств $\mathbb{U}$ и $\mathbb{V}$ обозначается соответственно через

$$\mathbb{U} \cup \mathbb{V} := \{x : \ x \in \mathbb{U} \ \text{ или } \ x \in \mathbb{V}\},$$
$$\mathbb{U} \cap \mathbb{V} := \{x : \ x \in \mathbb{U} \quad \text{и} \quad x \in \mathbb{V}\},$$
$$\mathbb{U} \setminus \mathbb{V} := \{x \in \mathbb{U} : \ x \notin \mathbb{V}\}.$$

Если $\mathbb{U} \cap \mathbb{V} = \emptyset$, т.е. множества $\mathbb{U}$ и $\mathbb{V}$ не содержат общих элементов, то они называются *непересекающимися.* Если множество $\mathbb{V}$ является подмножеством в $\mathbb{U}$, $\mathbb{V} \subset \mathbb{U}$, то разность $\mathbb{U} \setminus \mathbb{V}$ называется *дополнением* к $\mathbb{V}$ в $\mathbb{U}$. □

Из определения, очевидно, следует

Предложение 2.1.1. *Пусть* $\mathbb{W} = \mathbb{U} \setminus \mathbb{V}$*, тогда* $\mathbb{W} \cup \mathbb{V} = \mathbb{U}$ *и* $\mathbb{W} \cap \mathbb{V} = \emptyset$*.*

Замечание. Операции объединения, пересечения и дополнения соответствуют логическим связкам “или”, “и”, “нет”. □

Предложение 2.1.2. *Операции объединения, пересечения и вычитания множеств удовлетворяют тождествам:*

$$\begin{aligned} \mathbb{M} \setminus (\mathbb{U} \cup \mathbb{V}) &= (\mathbb{M} \setminus \mathbb{U}) \cap (\mathbb{M} \setminus \mathbb{V}), \\ \mathbb{M} \setminus (\mathbb{U} \cap \mathbb{V}) &= (\mathbb{M} \setminus \mathbb{U}) \cup (\mathbb{M} \setminus \mathbb{V}) \end{aligned} \tag{2.1}$$

и

$$\mathbb{M} \cap (\mathbb{U} \cup \mathbb{V}) = (\mathbb{M} \cap \mathbb{U}) \cup (\mathbb{M} \cap \mathbb{V}),$$
$$\mathbb{M} \cup (\mathbb{U} \cap \mathbb{V}) = (\mathbb{M} \cup \mathbb{U}) \cap (\mathbb{M} \cup \mathbb{V}).$$ □

Доказательство. Выполняется простой проверкой. Рис.2.1 иллюстрирует формулы (2.1). □

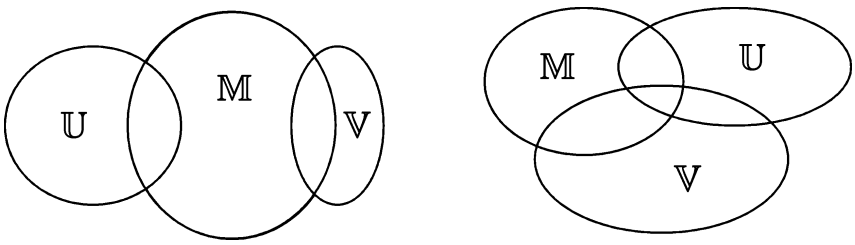


Рис. 2.1: Объединение, пересечение и разность множеств. Слева и справа затемнены те области, которые соответствуют равенствам (2.1).

Определение. Пусть $\mathbb{M}_1$ и $\mathbb{M}_2$ – два непустых множества, тогда их *прямым* или *декартовым* произведением $\mathbb{M}_1 \times \mathbb{M}_2$ называется множество упорядоченных пар (x_1, x_2), где $x_1 \in \mathbb{M}_1$ и $x_2 \in \mathbb{M}_2$:

$$\mathbb{M}_1 \times \mathbb{M}_2 := \{(x_1, x_2) : \quad x_1 \in \mathbb{M}_1, x_2 \in \mathbb{M}_2\}.$$

Если одно из множеств $\mathbb{M}_1$ или $\mathbb{M}_2$ пусто, то их прямое произведение пусто. □

Для того, чтобы понятие множества было более содержательным, на нем вводятся различные структуры. Важным примером является упорядочение.

Определение. Множество $\mathbb{M}$ называется *частично упорядоченным*, если на нем задано бинарное отношение $\leq$, удовлетворяющее следующим условиям:

$$\begin{aligned} &x \leq x, && \text{– рефлексивность,} \\ &x \leq y \ \text{ и } \ y \leq z \ \Rightarrow \ x \leq z, && \text{– транзитивность,} \\ &x \leq y \ \text{ и } \ y \leq x \ \Rightarrow \ x = y, && \text{– антисимметричность.} \end{aligned}$$

В общем случае может оказаться, что для некоторой пары элементов соотношение $\leq$ не определено. Если для любой пары элементов $x, y \in \mathbb{M}$ либо $x \leq y$, либо $y \leq x$, то множество $\mathbb{M}$ называется *линейно* упорядоченным или *цепью*. □

Пример 2.1.8. Множество всех подмножеств $\mathbb{P}(\mathbb{M})$ множества $\mathbb{M}$ является частично упорядоченным по отношению включения $\subseteq$. Если множество состоит из двух элементов $\mathbb{M} := \{x, y\}$, то его подмножествами являются $\emptyset$, $\{x\}$, $\{y\}$ и $\{x, y\}$. Оно не является линейно упорядоченным, т.к. для подмножеств $\{x\}$ и $\{y\}$ отношение $\subseteq$ не определено. □

Пример 2.1.9. Множество вещественных чисел $\mathbb{R}$, где символ $\leq$ означает “меньше или равно”, является линейно упорядоченным. □

2.2 Поле вещественных чисел $\mathbb{R}$ и прямая

В геометрии к неопределяемым понятиям относится *прямая линия* или, короче, прямая, которую будем обозначать буквой $\mathbb{R}$. При этом ее представляют как отрезок, начерченный по линейке и мысленно продолженный до бесконечности в обе стороны. Прямая линия находится во взаимно однозначном соответствии с полем вещественных (действительных) чисел. Другими словами, прямая является наглядным изображением поля вещественных чисел. На ней выбирается произвольная точка, *начало отсчета*, которой ставится в соответствие число нуль. Затем каждому положительному числу $x \in \mathbb{R}_+$ ставится в соответствие точка, лежащая справа от нуля на расстоянии, равном этому числу. Каждому отрицательному числу $x \in \mathbb{R}_-$ ставится в соответствие точка, лежащая слева от начала отсчета на расстоянии, равном модулю этого числа. При этом расстояние измеряется с помощью линейки, а число x называется координатой точки. Таким образом между точками прямой и вещественными числами устанавливается взаимно однозначное соответствие. Суммируя, можно сказать, что прямая это не более, чем наглядный образ вещественных чисел.

Расстояние l между двумя точками a и b, по-определению, равно модулю разности двух вещественных чисел $l(a, b) = |b - a|$. Его можно записать в виде интеграла

$$l(a, b) := \left| \int_a^b dx g \right| = |b - a|, \qquad g = 1. \tag{2.2}$$

Здесь введена функция g, равная единице, которая называется *метрикой*. Метрика в (2.2) всюду равна единице и называется *евклидовой*. В дифференциальной геометрии понятие расстояния обобщается за счет расширения класса рассматриваемых метрик. Понятие же евклидовой метрики чрезвычайно важно, т.к. служит той точкой отсчета, с которой сравниваются все остальные метрики.

Координаты точек складываются и умножаются так же, как и обычные вещественные числа. Для вещественных чисел определены две операции: сложение и умножение. По отношению к сложению вещественные числа образуют абелеву группу. Напомним общие определения и основные сведения из теории групп.

Определение. Непустое множество $\mathbb{G}$ называется *группой*, если выполнены четыре условия:

1) *Закон композиции.* Каждой паре элементов $a, b \in \mathbb{G}$ сопоставляется третий элемент этого же множества, называемый *произведением* элементов и обозначаемый $a \cdot b$ или ab. Закон композиции называется также *бинарной операцией*.
2) *Закон ассоциативности.* Для любых трех элементов $a, b, c \in \mathbb{G}$ имеет место равенство
$$(ab)c = a(bc).$$
3) В $\mathbb{G}$ существует левая *единица* e:
$$ea = a, \qquad \forall a \in \mathbb{G}.$$
4) Для каждого элемента $a \in \mathbb{G}$ существует по крайней мере один левый *обратный элемент* $a^{-1} \in \mathbb{G}$:
$$a^{-1}a = e.$$

Множество элементов с одной бинарной операцией, которая удовлетворяет только условию ассоциативности, называется *полугруппой*. Полугруппа с единичным элементом называется *моноидом*. Если для любых двух элементов $ab = ba$, то группа (или полугруппа) называется *коммутативной* или *абелевой*. В противном случае группа (или полугруппа) называется *неабелевой*. □

Пример 2.2.1. Пусть $n\mathbb{Z}$ – множество целых чисел, делящихся на n, где n – произвольное натуральное число. Это множество содержит число 0 при всех n, и в нем определены операции сложения $(+)$ и умножения $(\cdot)$. Пара $(n\mathbb{Z}, +)$ является коммутативной группой, где роль единицы выполняет число 0. Обозначим через $n\mathbb{Z}_+$ все неотрицательные числа из $n\mathbb{Z}$. Тогда пара $(n\mathbb{Z}_+, +)$ будет коммутативным моноидом. Если $n \geq 2$, то пара $(n\mathbb{Z}, \cdot)$ является коммутативной полугруппой без единицы. □

Пример 2.2.2. Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами x, y, z. Пусть $\mathbb{G}$ – множество всех отображений $\mathbb{R}^3$ в себя. Под композицией двух отображений мы понимаем их последовательное выполнение. В предложении 2.4.1 будет доказано, что композиция отображений произвольного множества является ассоциативной операцией. Множество отображений $\mathbb{G}$ содержит единицу, которой является тождественное отображение, оставляющее точки евклидова пространства неподвижными. Ясно, что множество всех отображений представляет собой некоммутативный моноид. В общем случае множество отображений $\mathbb{G}$ группу не образует, т.к. содержит, например, проекцию на плоскость x, y, которая не имеет обратного отображения. □

В дифференциальной геометрии изучаются чаще всего такие преобразования многообразий, которые образуют группу. Поэтому опишем некоторые свойства групп более подробно.

Предложение 2.2.1. *Каждая левая единица $e \in \mathbb{G}$ является одновременно и правой единицей, $ae = a$ для всех $a \in \mathbb{G}$. Единица в группе единственна. Каждый левый обратный элемент $a^{-1} \in \mathbb{G}$ одновременно является и правым обратным элементом, $aa^{-1} = e$. Обратный элемент a^{-1} для всех $a \in \mathbb{G}$ единственен. Справедливо правило $(ab)^{-1} = b^{-1}a^{-1}$.*

Доказательство. Приведено в большинстве учебников по теории групп. □

Определение. Группа, состоящая из конечного числа элементов, называется *конечной*. При этом количество элементов в группе называется *порядком* группы. Группа, состоящая из степеней одного элемента a называется *циклической*. Если циклическая группа имеет порядок n, то она обозначается $\mathbb{C}_n$, и $a^n = e$. Циклические группы обязательно абелевы. *Кручением* конечно порожденной абелевой группы $\mathbb{G}$ называется ее подгруппа, состоящая их всех элементов конечного порядка. Говорят, что конечно порожденная абелева группа не имеет кручения, если в ней нет элементов конечного порядка, т.е. не существует элемента конечная положительная степень которого равна единице. □

Замечание. В дифференциальной геометрии термин кручение имеет совершенно другой смысл (см. раздел 6.3). □

Пример 2.2.3. Множество невырожденных (с отличным от нуля определителем) квадратных $n \times n$ матриц над полем вещественных или комплексных чисел представляют собой группы по отношению к умножению матриц, которые обозначаются $\mathbb{GL}(n, \mathbb{R})$ и $\mathbb{GL}(n, \mathbb{C})$. Эти группы при $n > 1$ некоммутативны (неабелевы). □

Пример 2.2.4. Группа вращений трехмерного евклидова пространства $\mathbb{O}(3)$ и группа Лоренца $\mathbb{O}(1, 3)$ являются неабелевыми. □

Пример 2.2.5. Два целых числа 1 и -1 с обычным умножением образуют циклическую группу $\mathbb{Z}_2$. Ее также отождествляют с группой одномерных вращений, $\mathbb{Z}_2 = \mathbb{O}(1)$. □

Пример 2.2.6. Множество всех целых чисел $\mathbb{Z}$ образует бесконечномерную циклическую группу по отношению к сложению. Она порождена одним элементом – числом 1 – и не имеет кручения. □

Пример 2.2.7. Дробно-линейные преобразования расширенной комплексной плоскости (см. раздел **??**)

$$\overline{\mathbb{C}} \ni z \ \mapsto \ \frac{az + b}{cz + d} \in \overline{\mathbb{C}}, \qquad a, b, c, d \in \mathbb{C}, \quad ad - bc \neq 0 \tag{2.3}$$

образуют группу Ли, которая называется *группой Мёбиуса*. Эта группа шести параметрическая, т.к. числитель и знаменатель можно разделить на произвольное отличное от нуля комплексное число, и изоморфна группе невырожденных комплексных матриц $\mathbb{SL}(2, \mathbb{C})$. Подробно эта группа рассмотрена в разделе **??**. □

Определение. Пусть $\mathbb{H} \subset \mathbb{G}$ – *подгруппа* группы $\mathbb{G}$, т.е. элементы $\mathbb{H}$ сами по себе образуют группу. При этом единица подгруппы $\mathbb{H}$ совпадает с единицей группы $\mathbb{G}$. Обозначим через $a\mathbb{H}$ и $\mathbb{H}a$ множество элементов вида ah и ha, где a – некоторый фиксированный элемент группы $\mathbb{G}$, а элемент $h \in \mathbb{H}$ пробегает всю подгруппу $\mathbb{H}$. Множества элементов $a\mathbb{H}$ и $\mathbb{H}a$, где $a \in \mathbb{G}$, называются *левым* и *правым смежными классами* группы $\mathbb{G}$ по подгруппе $\mathbb{H}$. Если $a \in \mathbb{H}$, то левый и правый смежный класс совпадает с $\mathbb{H}$. □

Нетрудно проверить, что два левых смежных класса по $\mathbb{H}$ либо совпадают, либо не имеют ни одного общего элемента. Поэтому любой элемент группы принадлежит одному и только одному левому смежному классу, и его можно рассматривать как представитель этого класса. Под произведением двух левых смежных классов $a\mathbb{H}$ и $b\mathbb{H}$, которое также является левым смежным классом по подгруппе $\mathbb{H}$, понимается множество всех элементов вида cd, где $c \in a\mathbb{H}$ и $d \in b\mathbb{H}$. Сказанное выше верно и для правых смежных классов.

Определение. Два элемента a' и a называются *сопряженными*, если они связаны преобразованием *подобия*

$$a' = bab^{-1},$$

где b – некоторый элемент из $\mathbb{G}$. Сопряженность элементов является отношением эквивалентности (см. раздел 2.4) и определяет разбиение группы $\mathbb{G}$ на классы сопряженных элементов. Подгруппа $\mathbb{H}' \subset \mathbb{G}$ называется *сопряженной* подгруппе $\mathbb{H}$, если

$$\mathbb{H}' = b\mathbb{H}b^{-1},$$

для некоторого $b \in \mathbb{G}$. Очевидно, что единичные элементы сопряженных подгрупп $\mathbb{H}$ и $\mathbb{H}'$ совпадают. Подгруппа $\mathbb{H}$ отображается на себя для всех $b \in \mathbb{G}$ тогда и только тогда, когда подгруппа $\mathbb{H}$ содержит все элементы, сопряженные с ее элементами, что можно записать в виде $b\mathbb{H} = \mathbb{H}\,b$. Такая подгруппа называется *нормальной* подгруппой или *нормальным делителем*. Ее также называют *инвариантной подгруппой*. Левые и правые смежные классы по нормальному делителю совпадают и образуют группу по отношению к операции умножения смежных классов. Действительно, пусть $\mathbb{H}$ – нормальный делитель группы $\mathbb{G}$. Тогда для любых смежных классов $a\mathbb{H}$ и $b\mathbb{H}$ определено умножение

$$a\mathbb{H}b\mathbb{H} = ab\mathbb{H}\mathbb{H} = ab\mathbb{H}.$$

То есть множество смежных классов образует группу само по себе, при этом роль единицы выполняет нормальная подгруппа $\mathbb{H}$. Эта группа называется *факторгруппой* и обозначается $\mathbb{G}/\mathbb{H}$. □

Пример 2.2.8. Если в качестве подгруппы выбрать саму группу, $\mathbb{H} = \mathbb{G}$, то факторгруппа состоит из одного элемента – единицы, $\mathbb{G}/\mathbb{G} = e$. □

Определение. Множество всех элементов группы $\mathbb{G}$, перестановочных с любым элементом этой группы, является нормальным делителем и называется *центром* группы. Центр всегда является абелевой подгруппой группы $\mathbb{G}$.

Для абелевых групп композицию двух элементов часто называют *сложением* и пишут $a + b$. Тогда группу $\mathbb{G}$ называют *аддитивной группой* или *модулем*. Вместо единичного элемента здесь фигурирует *нулевой элемент*:

$$0 + a = a, \qquad \forall a \in \mathbb{G},$$

а обратный элемент обозначают $-a$:

$$-a + a = 0.$$ □

Очевидно, что любая подгруппа абелевой группы является нормальным делителем.

Пример 2.2.9. Множество натуральных чисел $\mathbb{N}$ группы по отношению к сложению не образует, т.к. не содержит ни нулевого, ни обратных элементов. □

Пример 2.2.10. Множества целых $\mathbb{Z}$ и вещественных чисел $\mathbb{R}$ образуют аддитивные группы по отношению к сложению. По сути дела, отсюда и пошло название "сложение" для групповой композиции в абелевых группах. □

Пример 2.2.11. По отношению к умножению вещественные числа группу не образуют, т.к. у нуля обратного элемента не существует. В то же время множество вещественных чисел без нуля образует абелеву группу по отношению к умножению. Эта группа является одномерной группой Ли и состоит из двух несвязных компонент: отрицательных и положительных чисел чисел $\mathbb{R}_- \cup \mathbb{R}_+$. При этом $e = 1 \in \mathbb{R}_+$. □

Пример 2.2.12. Множество целых чисел с операцией сложения и отношением эквивалентности $n + p \sim n$, где натуральное число $p \in \mathbb{N}$ фиксировано, и $n \in \mathbb{Z}$ любое, представляет собой конечную циклическую группу $\mathbb{Z}_p$, состоящую из p элементов. В этом случае мы пишем $n + p = n \mod p$.

Группу $\mathbb{Z}_p$ можно описать также следующим образом. Обозначим через $p\mathbb{Z}$ подгруппу целых чисел, состоящую из тех чисел, которые делятся на p без остатка. Множество $p\mathbb{Z}$ является нормальной подгруппой в $\mathbb{Z}$. Тогда множества левых и правых смежных классов совпадают и образуют фактор группу

$$\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}.$$

При $p = 2$ эта группа изоморфна группе, рассмотренной в примере 2.2.5. □

Определение. *Прямым произведением* двух групп $\mathbb{G}_1$ и $\mathbb{G}_2$ называется группа $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$, образованная всеми упорядоченными парами (a_1, a_2), где $a_1 \in \mathbb{G}_1$ и $a_2 \in \mathbb{G}_2$, с умножением, определяемым формулой

$$(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_2 b_2),$$

для всех $a_1, b_1 \in \mathbb{G}_1$ и $a_2, b_2 \in \mathbb{G}_2$. □

Если группы $\mathbb{G}_1$ и $\mathbb{G}_2$ конечны, то порядок группы $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$ равен произведению порядков групп $\mathbb{G}_1$ и $\mathbb{G}_2$. Единицей группы $\mathbb{G}$ является пара (e_1, e_2), где e_1 и e_2 – единицы соответственно в группах $\mathbb{G}_1$ и $\mathbb{G}_2$. При этом элементы вида (a_1, e_2), где a_1 пробегает всю группу $\mathbb{G}_1$, образуют подгруппу в прямом произведении $\mathbb{G}_1 \times \mathbb{G}_2$, которая изоморфна $\mathbb{G}_1$. Аналогично, элементы вида (e_1, a_2) образуют подгруппу в $\mathbb{G}_1 \times \mathbb{G}_2$, изоморфную $\mathbb{G}_2$.

Если у групп $\mathbb{G}_1$ и $\mathbb{G}_2$ нет общих элементов, то элемент их прямого произведения (a_1, a_2) можно записывать просто $a_1 a_2$, причем $a_1 e_2 = a_1$ и $e_1 a_2 = a_2$.

Пример 2.2.13. Любая размерная величина в физике является прямым произведением числа и единицы измерения. Выражения вида сила = масса × ускорение также являются прямыми произведениями. □

Вещественные числа $\mathbb{R}$ представляют собой множество с двумя бинарными операциями. А именно, для двух произвольных чисел $a, b \in \mathbb{R}$ однозначно определена их сумма $a + b$ и произведение ab. Вещественные числа представляют собой частный случай множества, которое называется в алгебре кольцом.

Определение. Непустое множество с двумя бинарными операциями называется *кольцом*, если выполняются следующие условия.

Законы сложения.

1) Закон ассоциативности: $a + (b + c) = (a + b) + c$.

2) Закон коммутативности: $a + b = b + a$.

3) Разрешимость уравнения $a + x = b$ для всех a и b.

Закон умножения.

1) Закон ассоциативности: $a(bc) = (ab)c$.

Законы дистрибутивности.

1) $a(b + c) = ab + ac$.

2) $(b + c)a = ba + ca$. Если умножение в кольце коммутативно, то говорят о *коммутативном кольце*. □

Три закона сложения означают в совокупности, что элементы кольца образуют абелеву группу (модуль) по отношению к сложению. Если кольцо обладает правым и левым единичным элементом одновременно, $ea = ae = a$, то он называется *единицей*, и говорят о *кольце с единицей*. В общем случае левые и правые единицы у кольца могут различаться, а единиц может быть несколько.

Пример 2.2.14. Минимальное конечное кольцо состоит из двух элементов 0 и 1 со следующей таблицей сложения и умножения

$$\begin{array}{lll} 0 + 0 = 0, & 0 + 1 = 1, & 1 + 1 = 0, \\ 0 \cdot 0 = 0, & 0 \cdot 1 = 0, & 1 \cdot 1 = 1. \end{array}$$

Это кольцо коммутативно, т.к.

$$1 \cdot 0 = 1 \cdot (1 + 1) = 1 \cdot 1 + 1 \cdot 1 = 1 + 1 = 0.$$ □

Пример 2.2.15. Множество целых чисел с операцией сложения и умножения образуют коммутативное кольцо с единицей. □

Определение. Кольцо, для каждого элемента которого $a^2 = a$, называется *булевым*. □

Определение. Кольцо называется *телом*, если:

1) в нем есть по крайней мере один элемент, отличный от нуля;

2) уравнения $ax = b$ и $xa = b$ при $a \neq 0$ разрешимы.

Коммутативное тело называется *полем* или *рациональным кольцом*. □

Замечание. Тело, как и кольцо, является аддитивной группой (модулем) по отношению к сложению, однако не является группой по отношению к умножению, т.к. у нуля в общем случае нет обратного элемента. Операции сложения и умножения в поле связаны дистрибутивными законами. □

Пример 2.2.16. Множества рациональных $\mathbb{Q}$ и действительных $\mathbb{R}$ с естественными операциями сложения и умножения являются полями. При этом $\mathbb{Q} \subset \mathbb{R}$. Целые числа образуют подкольцо поля рациональных чисел $\mathbb{Z} \subset \mathbb{Q}$. □

Следующий пример показывает, что из одного поля можно построить другие поля.

Пример 2.2.17. В дальнейшем мы будем рассматривать не только вещественные многообразия, но и комплексные. Поэтому напомним основные определения для комплексных чисел. Рассмотрим упорядоченную пару вещественных чисел (a, b), где $a, b \in \mathbb{R}$. Будем считать две пары (a, b) и (c, d) равными, если $a = c$ и $b = c$. Введем на множестве пар операции сложения и умножения следующим образом:

$$(a, b) + (c, d) := (a + b, c + d),$$
$$(a, b)(c, d) := (ac - bd, ad + bc).$$

Нетрудно проверить, что все аксиомы поля для множества пар выполнены.

Определение. Множество пар (a, b) с введенными выше операциями сложения и умножения называется *полем комплексных чисел* и обозначается $z := (a, b) \in \mathbb{C}$. □

Множество комплексных чисел вида $(a, 0) \in \mathbb{C}$ можно отождествить с множеством вещественных чисел $a \in \mathbb{R}$. Тогда поле вещественных чисел является подполем поля комплексных чисел, $\mathbb{R} \subset \mathbb{C}$.

Определим умножение пар $(a, b) \in \mathbb{C}$ на вещественные числа:

$$c(a, b) := (a, b)c := (ca, cb), \qquad c \in \mathbb{R}.$$

Множество комплексных чисел с операцией сложения и умножения на вещественные числа образует двумерное векторное пространство. В качестве базиса этого векторного пространства выберем пары:

$$e_1 := (1, 0), \qquad e_2 := (0, 1).$$

Теперь произвольное комплексное число можно представить в виде

$$(a, b) = ae_1 + be_2.$$

Базисный вектор e_1 при умножении ведет себя как единица: $ze_1 = e_1z = z$. Поэтому его можно отождествить с единицей. Для второго элемента базиса принято обозначение $i := (0, 1)$. Его называют *мнимой единицей*. Нетрудно проверить, что $i^2 = -1$. Обычно комплексные числа записывают в виде

$$z = a + ib, \qquad a, b \in \mathbb{R}, \quad z \in \mathbb{C}.$$

Определение. Множество комплексных чисел $z = (a, b)$ отождествляется с евклидовой плоскостью $\mathbb{C} \approx \mathbb{R}^2$, где вещественные числа a и b рассматриваются в качестве декартовых координат. В этом случае множество комплексных чисел называется *комплексной плоскостью*. При этом ось абсцисс называется *действительной осью*, а ось ординат – *мнимой осью*. □

Линейная структура евклидовой плоскости $\mathbb{R}^2$, рассматриваемой как векторное пространство (см. раздел 2.3.3), совпадает с линейной структурой комплексных чисел $\mathbb{C}$. Однако скалярное умножение векторов в $\mathbb{R}^2$ не имеет никакого отношения к умножению комплексных чисел.

Пусть $z = a + ib \in \mathbb{C}$. Тогда $\sqrt{a^2 + b^2}$ – это евклидово расстояние от начала координат до точки z. Пусть φ – угол между положительным направлением оси абсцисс и радиус-вектором точки z определяется уравнениями:

$$\cos\varphi = \frac{a}{\sqrt{a^2 + b^2}}, \qquad \sin\varphi = \frac{b}{\sqrt{a^2 + b^2}}, \qquad |z| \neq 0, \qquad \varphi \in [0, 2\pi).$$

Тогда приняты следующие обозначения и названия:

$$\begin{aligned} a &:= \mathsf{re}\, z \ \text{ – действительная часть } z, \\ b &:= \mathsf{im}\, z \ \text{ – мнимая часть } z, \\ \bar z &:= a - ib \ \text{ – число, комплексно сопряженное к } z, \\ |z| &:= \sqrt{a^2+b^2} \ \text{ – модуль } z, \\ \mathsf{arg}\, z &:= \varphi + 2\pi k,\ k \in \mathbb{Z}, \ \text{ – аргумент } z. \end{aligned} \tag{2.4}$$

Функция $\mathsf{arg}\, z$ определена с точностью до прибавления целого кратного 2π.

Для комплексных чисел часто используют тригонометрическую запись:

$$z = r\, \mathsf{e}^{i\varphi} = r\cos\varphi + ir\sin\varphi, \qquad r := |z|.$$

При этом

$$\bar z = r\, \mathsf{e}^{-i\varphi} = r\cos\varphi - ir\sin\varphi.$$

Определение. В теории функций комплексного переменного часто используют *расширенную комплексную плоскость* $\overline{\mathbb{C}}$, которая получается из комплексной плоскости $\mathbb{C}$ добавлением бесконечно удаленной точки $z = \infty$. На этом множестве вводится следующая топология. Если подмножество $\mathbb{U} \subset \overline{\mathbb{C}}$ не содержит бесконечно удаленной точки ∞, то оно считается открытым, если оно открыто в $\mathbb{C}$. Если подмножество $\mathbb{U}$ содержит точку ∞, то оно считается открытым в $\overline{\mathbb{C}}$, если его дополнение является компактом в $\mathbb{C}$. Как правило, открытыми окрестностями точки ∞ мы будем считать дополнения $\overline{\mathbb{C}} \setminus \overline{\mathbb{S}}_R$ замкнутых дисков $\overline{\mathbb{S}}_R$ радиуса R до всей расширенной комплексной плоскости $\overline{\mathbb{C}}$. Расширенную комплексную плоскость называют также *комплексной сферой* или *сферой Римана*. □

Комплексная плоскость $\mathbb{C}$ с естественной топологией евклидова пространства является некомпактным многообразием. При добавлении бесконечной точки мы меняем топологию комплексной плоскости (теперь, например, точки $(-\infty, 0)$ и $(\infty, 0)$ близки). В результате получаем компактное топологическое пространство – сферу Римана $\overline{\mathbb{C}}$ (см. раздел 2.3.2). Комплексная сфера является компактификацией комплексной плоскости (см. раздел 2.4.1), которая осуществляется с помощью стереографической проекции.

На сфере Римана $\overline{\mathbb{C}}$ мы определили топологию, которая отличается от топологии комплексной плоскости $\mathbb{C}$. Эта топология является метрической. Соответствующую метрику можно определить следующим образом. С помощью стереографической проекции (см. раздел 8.1) устанавливаем взаимно однозначное соответствие между точками расширенной комплексной плоскости и обычной сферой, вложенной в трехмерное евклидово пространство, $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Затем определяем расстояние между точками расширенной комплексной плоскости как обычное евклидово расстояние между точками сферы $\mathbb{S}^2$. Это расстояние определяет на расширенной комплексной плоскости метрику, которая называется *хордовой* и отличается от исходной евклидовой метрики. Явные выражения для хордовой метрики нам не понадобятся, и мы их не будем их приводить. (Их можно найти, например, в [**?**].) Заметим, что при работе с расширенной комплексной плоскостью в ТФКП используют, как правило, не хордову метрику, которая определяет топологию, а обычную евклидову метрику как более наглядную.

Очевидно, что сфера Римана не несет структуры поля или векторного пространства. Эти структуры можно определить только на подмножестве $\mathbb{C} \subset \overline{\mathbb{C}}$. Тем не менее

для каждой точки $p \in \mathbb{C}$ отображение $z \mapsto \frac{1}{z-p}$, определенное обычным образом при $z \in \mathbb{C} \setminus \{p\}$, можно продолжить до биекции множества $\overline{\mathbb{C}}$ на себя, если положить $\frac{1}{0} := \infty$ и $\frac{1}{\infty} := 0$. □

Определение. Говорят, что поле $\mathbb{F}$ имеет *характеристику* p, если существует такое простое число p, что выполнено равенство

$$\underbrace{a + \ldots + a}_{p} = 0, \qquad \forall a \in \mathbb{F}.$$

В этом случае поле $\mathbb{F}$ содержит подполе $\mathbb{Z}_p$. Если такого числа p не существует, то говорят, что поле $\mathbb{F}$ имеет характеристику нуль. В последнем случае оно содержит подполе, состоящее из рациональных чисел. □

Пример 2.2.18. Поля рациональных $\mathbb{Q}$, действительных $\mathbb{R}$ и комплексных $\mathbb{C}$ чисел имеют характеристику нуль. □

2.3 Евклидово пространство $\mathbb{R}^n$

Основным понятием дифференциально геометрии является дифференцируемое многообразие, которое является обобщением евклидова пространства. Евклидово пространство наделено различными структурами: евклидовой метрикой, топологией, структурой векторного и аффинного пространства. Поэтому в настоящем разделе мы изучим это пространство с различных точек зрения, чтобы в дальнейшем было ясно, что именно и как обобщается в дифференциальной геометрии.

2.3.1 $\mathbb{R}^n$ как метрическое пространство

Пусть $\mathbb{R}$ – поле вещественных чисел. В геометрии поле вещественных чисел наглядно изображается в виде прямой линии на рисунках. При этом каждая точка прямой $x \in \mathbb{R}$ находится во взаимно однозначном соответствии с вещественным числом, которое обозначается той же буквой x и называется координатой точки. На прямой выбирается произвольная точка – *начало отсчета*, которой ставится в соответствие число нуль $0 \in \mathbb{R}$. Обычно предполагают, что на рисунках координаты точек упорядочены и возрастают слева направо, что отмечают стрелкой.

Евклидово расстояние l между двумя точками $x, y \in \mathbb{R}$, по-определению, равно модулю разности двух вещественных чисел $l(x, y) = |x - y|$.

Определение. Обозначим через $\mathbb{R}^n$ прямое произведение n прямых:

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n}, \tag{2.5}$$

где $n \in \mathbb{N}$ – произвольное натуральное число, которое называется *размерностью* пространства $\mathbb{R}^n$. Точкой $x \in \mathbb{R}^n$ является упорядоченный набор n вещественных чисел $x^\alpha \in \mathbb{R}$, $\alpha = 1, \ldots, n$, которые называются *декартовыми координатами* данной точки. Мы записываем координаты точки в виде строки,

$$x = \{x^\alpha\} = (x^1, \ldots, x^n) \in \mathbb{R}^n.$$

Каждый из сомножителей, входящих в определение пространства $\mathbb{R}^n$ (2.5), называется *координатной прямой*, а точка с нулевыми координатами $(0, \ldots, 0) \in \mathbb{R}^n$ – *началом координат*.

Под 0-мерным пространством $\mathbb{R}^0$ понимают одну точку – число нуль. Одномерное пространство представляет собой прямую, $\mathbb{R}^1 = \mathbb{R}$. Двумерное пространство $\mathbb{R}^2$ называется *плоскостью*. □

Замечание. Номер координаты обозначается с помощью верхнего индекса так же, как и показатель степени. Как правило, различие в значении индексов ясно из контекста. □

В определении пространства $\mathbb{R}^n$ точку и ее координаты можно отождествить. Однако, определив таким образом $\mathbb{R}^n$, мы можем затем перейти в другую систему координат (см. раздел 2.5). Тогда той же точке пространства $\mathbb{R}^n$ будет соответствовать другой набор вещественных чисел. Поэтому следует различать точку пространства $\mathbb{R}^n$ и ее координаты, которые зависят от выбора системы координат.

Расстояние l, т.е. отображение $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, между двумя произвольными точками $p, q \in \mathbb{R}^n$ с декартовыми координатами p^α и q^α определяется следующей формулой

$$l(p,q) := |q-p| = \sqrt{(q^1-p^1)^2 + \cdots + (q^n-p^n)^2}. \tag{2.6}$$

Между двумя бесконечно близкими точками x^α и $x^\alpha + dx^\alpha$ расстояние задается *интервалом*, который представляет симметричная квадратичная форма,

$$ds^2 := l^2(x, x+dx) = dx^\alpha dx^\beta g_{\alpha\beta}, \tag{2.7}$$

где метрика $g_{\alpha\beta}$ не зависит от точки пространства $\mathbb{R}^n$ и является единичной матрицей, которую будем обозначать следующим образом

$$g_{\alpha\beta} := \delta_{\alpha\beta} = \mathsf{diag}\,(\underbrace{1,\ldots,1}_{n}). \tag{2.8}$$

В формуле (2.7) и в дальнейшем по повторяющимся индексам, один из которых пишется сверху, а другой – снизу, производится суммирование, если не оговорено противное. Это правило называется *правилом суммирования Эйнштейна*. Матрица (2.8) называется *евклидовой метрикой* и имеет специальное обозначение $\delta_{\alpha\beta}$.

В дифференциальной геометрии роль индексов чрезвычайно важна. Поэтому отметим ряд общих правил, которые всюду используются в дальнейшем. Эти правила связаны с группами преобразований, которые действуют на геометрические объекты.

1) Каждое слагаемое может содержать некоторый индекс один или два раза. В первом случае он называется *свободным*, а во втором случае – *немым*.
2) Если некоторое выражение состоит из суммы нескольких слагаемых, то каждое слагаемое должно содержать один и тот же набор свободных индексов. При этом значения этих индексов во всех слагаемых должно фиксироваться одновременно.
3) Немой индекс обязательно встречается один раз сверху и один раз снизу в каждом слагаемом. Значение этого индекса в каждом слагаемом не может быть зафиксировано, т.к. по нему проводится суммирование. В разных слагаемых немые индексы можно обозначать различными буквами, а число их пар может различаться.

Иногда мы все же будем писать повторяющиеся индексы как нижние или верхние. Например, запись $g_{\alpha\alpha}$ обозначает диагональный элемент метрики, стоящий на α-том месте. При этом суммирование не проводится.

Выше мы ввели евклидову метрику в декартовой системе координат, с помощью которой было определено пространство $\mathbb{R}^n$. С помощью расстояния, т.е. отображения

$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, понятие декартовой системы координат можно обобщить следующим образом. Существуют преобразования координат пространства $\mathbb{R}^n$, которые сохраняют вид расстояния (2.6) и, следовательно, интервала (2.7). Эти преобразования образуют неоднородную группу вращений $\mathbb{IO}(n)$, состоящую из вращений пространства $\mathbb{R}^n$,

$$x'^{\alpha} = x^{\beta} S_{\beta}{}^{\alpha}, \qquad S_{\beta}{}^{\alpha} \in \mathbb{O}(n),$$

и сдвигов,

$$x'^{\alpha} = x^{\alpha} + a^{\alpha}, \qquad a^{\alpha} = \mathsf{const}.$$

Поэтому в дальнейшем под декартовой системой координат мы будем понимать любую систему координат, в которой расстояние между точками пространства $\mathbb{R}^n$ имеет вид (2.8).

Определение. Система координат x^{α}, $\alpha = 1, \ldots, n$ пространства $\mathbb{R}^n$, в которой расстояние между двумя произвольными точками $a, b \in \mathbb{R}^n$ имеет вид (2.6), называется *декартовой системой координат.* □

После того, как пространство $\mathbb{R}^n$ определено, в нем можно строить произвольные криволинейные, например, сферические или цилиндрические системы координат в зависимости от специфики той или иной задачи. В таких системах координат метрика $g_{\alpha\beta}(x)$ в (2.7) будет зависеть от точки пространства $\mathbb{R}^n$.

Определение. *Кривой* $\gamma = x(t) = \{x^{\alpha}(t)\}$ в пространстве $\mathbb{R}^n$ называется отображение замкнутого единичного отрезка $[0, 1]$ в пространство $\mathbb{R}^n$,

$$\gamma : \quad [0, 1] \ni t \;\mapsto\; \{x^{\alpha}(t)\} \in \mathbb{R}^n, \tag{2.9}$$

где t – вещественный параметр вдоль кривой. Говорят, что кривая соединяет две точки p и q, где $x(0) = p$, $x(1) = q$. Если граничные точки кривой совпадают, $p = q$, то кривая называется *замкнутой*. Совокупность функций $\dot{x}^{\alpha}$, где точка обозначает дифференцирование по параметру t, определяет касательный вектор к кривой,

$$X_{(\gamma)} = \{X^{\alpha}_{(\gamma)} := \dot{x}^{\alpha}\},$$

который называется *вектором скорости* кривой. Кривая γ называется также *путем*, при этом точка p является началом, а q – концом пути. □

Не следует смешивать понятие кривой с множеством точек, через которые она проходит. Согласно определению понятие кривой включает также порядок прохождения точек данного множества.

В общем случае кривая может иметь точки самопересечения.

Определение. Кривая, не имеющая точек самопересечения, называется *простой кривой*. Совпадение начала и конца замкнутой кривой мы не будем рассматривать как точки самопересечения. Поэтому можно говорить о *простой замкнутой кривой.* □

Определение. Кривая γ называется *гладкой (дифференцируемой)*, если все координатные функции $x^{\alpha}(t)$ являются гладкими (дифференцируемыми). Мы предполагаем, что вектор скорости дифференцируемой кривой отличен от нуля, т.е. отлична от нуля по крайней мере одна из компонент $\dot{x}^{\alpha}$, что соответствует погружению отрезка $[0, 1]$ в $\mathbb{R}^n$. Кривая γ называется *кусочно гладкой (дифференцируемой)*, если ее можно представить в виде объединения конечного числа простых замкнутых гладких (дифференцируемых) кривых. □

Согласно данным определениям не всякая гладкая кривая является кусочно гладкой, т.к. может иметь бесконечное число точек самопересечения.

Единичный отрезок в определении кривой выбран для удобства определения произведения путей и фундаментальной группы многообразия (см. раздел 11.1). Выбор другого замкнутого интервала соответствует *перепараметризации кривой*. Под этим понимается замена параметра $t \mapsto t' = t'(t)$, где $t'(t)$ – произвольная достаточно гладкая монотонная функция такая, что $dt'/dt \neq 0$. При этом вектор скорости кривой преобразуется по-правилу дифференцирования сложных функций:

$$X^\alpha_{(\gamma)} = \frac{dx^\alpha}{dt} = \frac{dx^\alpha}{dt'}\frac{dt'}{dt}.$$

Определение. *Длиной* дифференцируемой кривой γ называется интеграл

$$l(\gamma) := \int_p^q ds = \int_0^1 dt\sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}}. \tag{2.10}$$

Если кривая является кусочно дифференцируемой, то ее длиной называется сумма длин каждой дифференцируемой части. В евклидовом пространстве $\mathbb{R}^n$ длину кривой $s(t)$ от начала p до текущей точки $x(t) \in \gamma$ всегда можно выбрать в качестве параметра вдоль кривой. В этом случае s называют *каноническим параметром*. Он однозначно определяется уравнением

$$ds = dt\sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}} = dt\sqrt{X^\alpha_{(\gamma)} X^\beta_{(\gamma)} g_{\alpha\beta}},$$

с начальным условием $s(0) = 0$. □

Из формулы (2.10) следует, что определение длины кривой не зависит от выбора ее параметризации.

Если кривая параметризована каноническим параметром s, то вектор скорости $X^\alpha_{(\gamma)} = dx^\alpha/ds$ имеет единичную длину:

$$X^\alpha_{(\gamma)} X^\beta_{(\gamma)} g_{\alpha\beta} = 1.$$

Можно доказать, что расстояние (2.6) является точной нижней гранью интегралов (2.10) по всем возможным путям, соединяющим точки p и q,

$$l(p,q) = \inf_\gamma \int_p^q ds. \tag{2.11}$$

Кривая, вдоль которой интеграл (2.11) принимает наименьшее значение называется отрезком *прямой линии*, соединяющим точки p и q.

Точная нижняя грань (2.11) по всем кривым, соединяющим две точки $p, q \in \mathbb{R}^n$, или явная формула (2.6) задают расстояние или метрику пространства $\mathbb{R}^n$ в топологическом смысле. Напомним определение метрики для произвольного множества.

Определение. *Метрикой* на множестве $\mathbb{M}$ называется функция $l(p,q)$, определенная для любых двух точек $p, q \in \mathbb{M}$, которая удовлетворяет следующим условиям:

1) $l(p,p) = 0; \quad l(p,q) > 0, \quad (p \neq q)$ – положительная определенность
2) $l(p,q) = l(q,p)$ – симметричность,
3) $l(p,q) \leq l(p,r) + l(r,q), \quad \forall r \in \mathbb{M}$ – неравенство треугольника.

Значение $l(p,q)$ называется *расстоянием* между точками p и q. Пара $(\mathbb{M}, l)$, т.е. множество $\mathbb{M}$ с заданной метрикой l, называется *метрическим пространством*. □

Замечание. Эту метрику мы будем называть *топологической метрикой*, чтобы отличать ее от дифференциально-геометрической метрики (2.8) и ее обобщения, которое будет сделано в разделе 5. □

Понятие топологической метрики очень важно и может быть использовано для определения сходимости последовательностей. Пусть $(\mathbb{M}, l)$ – метрическое пространство, и $\{x_i\}$, $i \in \mathbb{N}$ – последовательность точек в $\mathbb{M}$.

Определение. Назовем $\{x_i\}$ *фундаментальной последовательностью* или *последовательностью Коши*, если для любого $\epsilon > 0$ существует такое натуральное число $N_\epsilon \in \mathbb{N}$, что $l(x_i, x_k) < \epsilon$ при всех $i > N_\epsilon$ и $k > N_\epsilon$. Фундаментальная последовательность, по-определению, называется сходящейся к некоторой точке, которая называется *пределом* последовательности. Пространство $\mathbb{M}$ называется *метрически полным*, если любая фундаментальная последовательность в $(\mathbb{M}, l)$ сходится к некоторой точке из $\mathbb{M}$. □

Другими словами полнота означает, что предел любой последовательности из $\mathbb{M}$, если он существует, тоже принадлежит $\mathbb{M}$. Нетрудно показать, что если предел существует, то он единственен.

На одном и том же множестве $\mathbb{M}$ можно задавать различные топологические метрики. Метрика l на множестве $\mathbb{M}$ называется *полной*, если метрическое пространство $(\mathbb{M}, l)$ полное.

Пример 2.3.1. Все пространство $\mathbb{R}^n$, а также любое замкнутое подмножество в нем с метрикой (2.11) является метрически полным. Открытые подмножества $\mathbb{R}^n$, отличные от всего пространства, метрически неполны, т.к. фундаментальные последовательности могут сходиться к граничным точкам, которые не принадлежат данным подмножествам. □

В пространстве $\mathbb{R}^n$, используя понятие расстояния (топологической метрики) (2.6), можно задавать различные подмножества, которые широко используются в дальнейшем и играют большую роль в приложениях. Эти подмножества определены в декартовой системе координат. В дальнейшем мы также предполагаем, что топология на этих подмножествах индуцирована вложением в евклидово пространство.

Пример 2.3.2 (Шар). Подмножество $\mathbb{B}^n_r(p) \subset \mathbb{R}^n$, определяемое неравенством:

$$\mathbb{B}^n_r(p) := \{x \in \mathbb{R}^n : \ |x - p| < r\}, \qquad r = \mathsf{const} > 0, \tag{2.12}$$

называется n-мерным открытым *шаром* радиуса r с центром в точке p. В одномерном случае, $n = 1$, шар называется *интервалом*[1]. При $n = 2$ шар $\mathbb{B}^2_r$ называется *диском* или *кругом*.

Если пара $(\mathbb{M}, l)$ – произвольное метрическое пространство, то в нем также можно определить шар радиуса r с центром в точке p:

$$\mathbb{B}_r(p) := \{x \in \mathbb{M} : \ l(x, p) < r\}. \tag{2.13}$$

Последнее определение не зависит от того имеет ли множество точек $\mathbb{M}$ размерность или нет. □

[1]Этот термин употребляется в дифференциальной геометрии также для обозначения квадрата расстояния между двумя бесконечно близкими точками (2.7)

Пример 2.3.3 (Сфера). Подмножество $\mathbb{S}^{n-1}_r(p) \subset \mathbb{R}^n$, определяемое равенством:

$$\mathbb{S}^{n-1}_r(p) := \{x \in \mathbb{R}^n : \ |x-p| = r\}, \qquad r = \mathsf{const} > 0, \tag{2.14}$$

называется $(n-1)$-мерной *сферой* $\mathbb{S}^{n-1}_r$ радиуса r с центром в точке p. При $n = 1$ сфера вырождается в две точки, являющиеся концами интервала $(p-r, p+r) \subset \mathbb{R}$. В дальнейшем для сферы единичного радиуса нижний индекс мы будем опускать, $\mathbb{S}^n := \mathbb{S}^n_1$. □

Пример 2.3.4 (Куб). Подмножество $\mathbb{U}^n_a(p) \subset \mathbb{R}^n$, определяемое неравенством:

$$\mathbb{U}^n_a(p) := \{x \in \mathbb{R}^n : \ |x^\alpha - p^\alpha| < a/2, \ \alpha = 1, \ldots, n\}, \qquad a = \mathsf{const} > 0, \tag{2.15}$$

называется n-мерным открытым *кубом* со стороной a и центром в точке p. Куб представляет собой прямое произведение n интервалов $(-a/2, a/2) \subset \mathbb{R}$. □

Замечание. Условия (2.12)–(2.15) определяют только множества точек шара, сферы и куба, ничего не говоря о том, как устроена топология на этих множествах. Мы всегда предполагаем, что топология шара, сферы, куба, а также других подмножеств евклидова пространства индуцирована их вложением в $\mathbb{R}^n$ (см. следующий раздел). □

Пример 2.3.5. Обозначим через $\mathbb{R}^n_+ \subset \mathbb{R}^n$ подпространство в $\mathbb{R}^n$, определяемое условием $x^n > 0$. Его замыкание $\overline{\mathbb{R}}^n_+$ имеет край, который является гиперплоскостью $\mathbb{R}^{n-1}$ и определяется уравнением $x^n = 0$. Мы считаем, что топологии на $\mathbb{R}^n_+$ и $\overline{\mathbb{R}}^n_+$ индуцированы вложениями $\mathbb{R}^n_+, \overline{\mathbb{R}}^n_+ \hookrightarrow \mathbb{R}^n$. Заметим, что открытые подмножества $\overline{\mathbb{R}}^n_+$ могут содержать точки края. □

Определение. Подмножество $\mathbb{U} \subset \mathbb{R}^n$ называется *ограниченным*, если существует шар $\mathbb{B}^r_n$ конечного радиуса r, целиком содержащий $\mathbb{U}$. □

Пример 2.3.6. Все $\mathbb{R}^n$, а также $\mathbb{R}^n_+, \overline{\mathbb{R}}^n_+$ является неограниченными подмножествами в $\mathbb{R}^n$. □

2.3.2 $\mathbb{R}^n$ как топологическое пространство

В предыдущем разделе мы использовали понятия открытого и замкнутого множеств, не дав им определений. Ниже мы восполним этот пробел и приведем необходимые сведения из общей топологии, которые необходимы для определения основного понятия дифференциальной геометрии – многообразия.

Начнем с общего определения топологического пространства и топологии.

Определение. *Топологическим пространством* называется пара $(\mathbb{M}, \mathcal{I})$, состоящая из множества точек $x \in \mathbb{M}$ и некоторого семейства $\{\mathbb{U}_i\}_{i\in I} = \mathcal{I}$ своих подмножеств, которые удовлетворяют следующим условиям:

1) пустое множество $\emptyset$ и все множество $\mathbb{M}$ принадлежат $\mathcal{I}$;

2) пересечение *любой пары* подмножеств из $\mathcal{I}$ принадлежит $\mathcal{I}$;

3) объединение *любого семейства* подмножеств $\mathcal{I}$ принадлежит $\mathcal{I}$.

Элементы семейства $\mathcal{I}$ называются *открытыми множествами* пространства $\mathbb{M}$, а семейство открытых множеств $\mathcal{I}$ – *топологией* пространства $\mathbb{M}$. Некоторое семейство $\{\mathbb{B}_j\}_{j\in J} = \mathcal{J}$ открытых множеств называется *базой* топологии пространства $\mathbb{M}$, если каждое множество из $\mathcal{I}$ есть объединение каких-либо множеств из $\mathcal{J}$. □

Из условия 2) следует, что пересечение любого, но конечного числа открытых множеств также является открытым. В примере 2.3.14 показано, что пересечение бесконечного числа открытых множеств может не быть открытым. В условии 3) допускается объединение бесконечного числа открытых множеств.

Пустое множество $\emptyset$ – это множество, не содержащее ни одного элемента.

Одна и та же топология $\mathcal{I}$ на множестве $\mathbb{M}$ может иметь много баз.

На любом множестве можно задать топологию, и не одну. Исключение составляют пустое множество $\emptyset$ (топология состоит из одного открытого множества – самого $\emptyset$) и множество, состоящее из одного элемента. В последнем случае топология единственна и состоит из двух открытых множеств: пустого множества и самого элемента.

Пример 2.3.7. Пусть $\mathbb{M}$ – множество. Будем считать, что каждая точка $x \in \mathbb{M}$ является открытым множеством и их совокупность образует базу топологии $\mathbb{M}$. Такую топологию называют *дискретной*. □

Пример 2.3.8. Пусть $\mathbb{M}$ – множество. Будем считать, что вся топология $\mathbb{M}$ состоит из двух множеств: пустого множества и всего $\mathbb{M}$. Это – пример другой крайности, и поэтому такую топологию называют *антидискретной*. Антидискретную топологию называют также *тривиальной*, потому что она слишком бедна. □

Определение. Если на множестве $\mathbb{M}$ задано две топологии $\mathcal{I}_1$ и $\mathcal{I}_2$, причем $\mathcal{I}_1 \subset \mathcal{I}_2$, то говорят, что топология $\mathcal{I}_1$ *слабее (грубее)* топологии $\mathcal{I}_2$, или что топология $\mathcal{I}_2$ *сильнее (тоньше)* топологии $\mathcal{I}_1$. □

Очевидно, что дискретная топология является наиболее тонкой, а антидискретная топология – наиболее грубой.

Замечание. Явное описание всех открытых множеств, т.е. топологии на множестве точек $\mathbb{M}$ в общем случае является сложной задачей. Поэтому на практике, чтобы определить топологию в пространстве $\mathbb{M}$, сначала задают базу топологии, а затем объявляют все возможные объединения элементов базы открытыми множествами. Поэтому понятие базы топологии помогает конструктивно подойти к заданию топологии на множестве точек $\mathbb{M}$. □

Пример 2.3.9. Пусть $\mathbb{R}$ – вещественная прямая. Семейство всех открытых интервалов $(a, b) \subset \mathbb{R}$ является базой некоторой топологии на $\mathbb{R}$, которая занимает промежуточное положение между дискретной и антидискретной топологией. Такая топология на вещественной прямой называется *естественной*. Она лежит в основе математического анализа. □

Определение. Пусть $(\mathbb{M}, \mathcal{I})$ и $(\mathbb{N}, \mathcal{J})$ – два топологических пространства. На прямом произведении $\mathbb{M} \times \mathbb{N}$, образованном всеми упорядоченными парами (x, y), где $x \in \mathbb{M}$ и $y \in \mathbb{N}$, можно ввести топологию. А именно, множество всех пар подмножеств $(\mathbb{U}, \mathbb{V})$, где $\mathbb{U}$ и $\mathbb{V}$ открыты соответственно в $\mathbb{M}$ и $\mathbb{N}$, образует базу некоторой топологии в прямом произведении $\mathbb{M} \times \mathbb{V}$. Такое произведение называется *топологическим*. □

Определение. *Евклидовым* пространством $\mathbb{R}^n$ размерности n называется топологическое произведение n вещественных прямых $\mathbb{R}$:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n}, \tag{2.16}$$

каждая из которых снабжена естественной топологией. Топологию $\mathbb{R}^n$ также будем называть *естественной*. Будем писать $\dim \mathbb{R}^n = n$. □

Для топологического произведения мы будем использовать тот же символ, что и для прямого, подразумевая, что на прямом произведении задается топология, которая определяется топологией сомножителей. Поэтому для евклидова пространства мы сохранили общепринятое прежнее обозначение $\mathbb{R}^n$.

Базой естественной топологии в евклидовом пространстве $\mathbb{R}^n$ является множество всех кубов (2.15) с произвольными сторонами и центрами.

Аналогично определяется комплексное пространство

$$\mathbb{C}^n = \underbrace{\mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}}_{n}, \tag{2.17}$$

где $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ – комплексная плоскость с естественной топологией, совпадающей с топологией евклидовой плоскости $\mathbb{R}^2$. Очевидно, что, как топологическое пространство, комплексное пространство $\mathbb{C}^n$ диффеоморфно евклидову пространству вдвое большего числа измерений, $\mathbb{C}^n \approx \mathbb{R}^{2n}$.

Определение. Рассмотрим некоторое подмножество $\mathbb{D} \subset \mathbb{M}$ топологического пространства $(\mathbb{M}, \mathcal{I})$. Будем говорить, что на $\mathbb{D}$ задана *индуцированная топология* $(\mathbb{D}, \mathcal{J})$, если подмножество $\mathbb{V} \in \mathcal{J}$ открыто тогда и только тогда, когда $\mathbb{V} = \mathbb{U} \cap \mathbb{D}$, где $\mathbb{U} \in \mathcal{I}$ – некоторое открытое подмножество в $\mathbb{M}$. □

Пример 2.3.10. Пусть $\mathbb{D} \in \mathbb{R}^n$ – конечное множество точек пространства $\mathbb{R}^n$. Тогда топология на $\mathbb{D}$, индуцированная естественной топологией в $\mathbb{R}^n$, является дискретной. □

Замечание. При рассмотрении подмножеств евклидова пространства $\mathbb{R}^n$ как топологических пространств мы будем предполагать, что они снабжены индуцированной топологией, если не оговорено противное. В частности, говоря о шарах, сферах и других подмножествах в $\mathbb{R}^n$, мы предполагаем, что на них задана естественная топология, индуцированная вложением в евклидово пространство. □

Обсудим связь между метрическими пространствами, рассмотренными в предыдущем разделе, и топологическими пространствами. В произвольном метрическом пространстве $(\mathbb{M}, l)$ можно выбрать в качестве базы топологии семейство шаров, определенных формулой (2.12). Тем самым любое метрическое пространство превращается в топологическое. Обратное утверждение неверно. Не на всяком топологическом пространстве можно ввести метрику такую, чтобы она определяла заданную топологию.

Определение. Топологическое пространство $(\mathbb{M}, \mathcal{I})$ называется *метризуемым*, если на множестве его точек $\mathbb{M}$ можно ввести такую метрику $l(p, q)$, что множество всех шаров (2.13) является базой топологии $\mathcal{I}$. □

Пример 2.3.11. Топологическое пространство с дискретной топологией не является метризуемым. □

Если на множестве $\mathbb{M}$ заданы две различные топологические метрики (расстояния) l_1 и l_2, то они могут индуцировать одно и то же топологическое пространство $(\mathbb{M}, \mathcal{I})$.

Наличие евклидова расстояния (2.6) в евклидовом пространстве $\mathbb{R}^n$ позволяет определить шар (2.12). Совокупность всех шаров можно выбрать в качестве базы некоторой топологии евклидова пространства. Любой куб можно представить в виде объединения бесконечного числа шаров. Поэтому множество всех шаров также

является базой естественной топологии евклидова пространства $\mathbb{R}^n$. Таким образом, евклидово пространство является метризуемым. На нем определена евклидова метрика (2.6). В дальнейшем, мы, как правило, предполагаем, что на евклидовом пространстве задана также евклидова метрика.

Можно показать, что минимальная база естественной топологии евклидова пространства $\mathbb{R}^n$ состоит из шаров с рациональными радиусами. Эта база топологии является счетной.

В общем случае, если топологическое пространство $(\mathbb{M}, \mathcal{I})$ имеет счетную базу, то говорят, что $\mathbb{M}$ удовлетворяет *второй аксиоме счетности*.

В топологическом пространстве $\mathbb{M}$ *окрестностью* точки $x \in \mathbb{M}$ называют любое открытое множество в $\mathbb{M}$, содержащее x. Обычно в качестве окрестности точки евклидова пространства $x \in \mathbb{R}^n$ выбирается некоторый шар с центром в этой точке. Нетрудно показать, что подмножество $\mathbb{U} \in \mathbb{M}$ открыто тогда и только тогда, когда каждая точка $x \in \mathbb{U}$ имеет окрестность $\mathbb{U}_x$, целиком содержащуюся в $\mathbb{U}$. Точка x есть *предельная точка* подмножества $\mathbb{U} \subset \mathbb{M}$, если каждая окрестность точки x содержит точки множества $\mathbb{U}$, отличные от x.

Дадим два эквивалентных определения замкнутого множества.

Определение. Подмножество $\mathbb{U}$ топологического пространства $(\mathbb{M}, \mathcal{I})$ *замкнуто*, если

1) $\mathbb{U}$ есть дополнение в $\mathbb{M}$ некоторого открытого множества,

или

2) $\mathbb{U}$ содержит все свои предельные точки. □

Нетрудно доказать, что объединение двух и, следовательно, произвольного конечного числа замкнутых множеств является замкнутым множеством. Кроме того, пересечение любого семейства, в том числе бесконечного, замкнутых множеств есть замкнутое множество.

Пустое множество включается в определение топологического пространства и, по-определению, является открытым. Отсюда вытекает, что все топологическое пространство, как дополнение пустого множества, является одновременно открытым и замкнутым. Аналогично, само пустое множество также является открытым и замкнутым одновременно. Отдельная точка евклидова пространства $\mathbb{R}^n$ является замкнутым подмножеством как дополнение открытого подмножества.

Замечание. Топологию пространства можно определить в терминах замкнутых подмножеств. Определение аналогично приведенному выше, только открытые множества надо заменить на замкнутые, а операции пересечения и объединения поменять местами. □

Замечание. Отрезок $[0, 1] \subset \mathbb{R}$ является замкнутым подмножеством в $\mathbb{R}$. Однако, если рассматривать отрезок $[0, 1]$ как самостоятельное топологическое пространство с топологией, индуцированной из $\mathbb{R}$, то он будет одновременно и замкнутым, и открытым. □

Определение. *Дискретным подмножеством* топологического пространства $\mathbb{M}$ называется такое его подпространство $\mathbb{U} \subset \mathbb{M}$, каждое подмножество которого замкнуто в $\mathbb{M}$. □

Пример 2.3.12. Набор изолированных точек в евклидовом пространстве $\mathbb{R}^n$ является дискретным подмножеством. Он всегда конечен или счетен. □

Рассмотрим произвольное подмножество $\mathbb{U} \subset \mathbb{R}^n$, которое не совпадает со всем евклидовым пространством. В общем случае оно может не быть ни открытым, ни замкнутым. *Замыканием* множества называется объединение самого множества $\mathbb{U}$ и всех его предельных точек или, что эквивалентно, пересечение всех замкнутых подмножеств, содержащих $\mathbb{U}$. Оно обозначается через $\overline{\mathbb{U}}$. Из определения следует, что всегда $\mathbb{U} \subset \overline{\mathbb{U}}$. Можно доказать следующее

Предложение 2.3.1. *Точка $x \in \overline{\mathbb{U}}$ тогда и только тогда, когда любая окрестность $\mathbb{V}_x$ точки x имеет непустое пересечение с $\mathbb{U}$, $\mathbb{V}_x \bigcap \mathbb{U} \neq \emptyset$.*

Доказательство. См., например, [**?**]. □

Определение. Если некоторое множество совпадает со своим замыканием, $\mathbb{U} = \overline{\mathbb{U}}$, то оно является замкнутым. *Внутренностью* множества $\mathbb{U}$ называется наибольшее открытое множество $\mathsf{int}\,\mathbb{U}$, целиком содержащееся в $\mathbb{U}$, или, что эквивалентно, объединение всех открытых множеств, содержащихся в $\mathbb{U}$. Точка $x \in \mathbb{U}$ называется внутренней, если $x \in \mathsf{int}\,\mathbb{U}$. *Границей* множества $\mathbb{U}$ называется разность

$$\partial\mathbb{U} = \overline{\mathbb{U}} \setminus \mathsf{int}\,\mathbb{U}. \qquad \square$$

Пример 2.3.13. Шар, определенный равенством (2.12), является *открытым* шаром. Замыкание шара $\overline{\mathbb{B}}^n_r$ включает в себя точки $(n-1)$-мерной сферы (2.14). При этом сфера $\mathbb{S}^{n-1}_r$ является границей как открытого, так и замкнутого шара. □

Пример 2.3.14. Покажем, что пересечение бесконечного числа открытых множеств может быть замкнутым. Рассмотрим последовательность шаров переменного радиуса $\mathbb{B}^n_{r_i}(0)$, $i \in \mathbb{N}$, в евклидовом пространстве $\mathbb{R}^n$, $n \geq 2$, с центром в начале координат. Пусть все радиусы больше единицы, $r_i > 1$, и их последовательность сходится к единице. Тогда пересечением этой последовательности шаров является замкнутый шар единичного радиуса $\overline{\mathbb{B}}^n_1(0)$, включающий граничную окружность $\mathbb{S}^{n-1}_1(0)$. □

Пример 2.3.15. Объединение бесконечного числа замкнутых подмножеств может быть открыто. Рассмотрим последовательность замкнутых шаров переменного радиуса $\overline{\mathbb{B}}^n_{r_i}(0)$, $i \in \mathbb{N}$, в евклидовом пространстве $\mathbb{R}^n$, $n \geq 2$, с центром в начале координат. Пусть все радиусы ограничены, $0 < r_i < 1$, и их последовательность сходится к единице. Тогда объединением этой последовательности шаров является открытый шар единичного радиуса $\mathbb{B}^n_1(0)$. □

Предложение 2.3.2. *Граница $\partial\mathbb{U}$ любого подмножества $\mathbb{U} \subset \mathbb{M}$ является замкнутым подмножеством.*

Доказательство. Поскольку $\mathbb{M} \setminus \overline{\mathbb{U}}$ и $\mathsf{int}\,\mathbb{U}$ открыты в $\mathbb{M}$, то граница $\partial\mathbb{U}$ всегда замкнута, как дополнение открытых множеств. □

Дадим эквивалентное определение внутренних и граничных точек.

Определение. Пусть $\mathbb{U} \subset \mathbb{M}$ – собственное подмножество. Точка $x \in \mathbb{U}$ называется *внутренней*, если подмножество $\mathbb{U}$ содержит некоторую окрестность $\mathbb{U}_x$, содержащую точку x. Точка $y \in \mathbb{M}$ называется *внешней* по отношению к подмножеству $\mathbb{U} \subset \mathbb{M}$, если существует окрестность $\mathbb{U}_y \ni y$, которая не имеет общих точек с $\mathbb{U}$. Точка $x \in \mathbb{M}$ называется *граничной* точкой подмножества $\mathbb{U} \subset \mathbb{M}$, если любая окрестность $\mathbb{U}_x \ni x$ содержит как внутренние, так и внешние точки подмножества $\mathbb{U} \subset \mathbb{M}$. □

Дадим еще несколько менее наглядных определений.

Определение. Точка $x \in \mathbb{M}$ называется *изолированной*, если у нее есть окрестность, не содержащая других точек $\mathbb{M}$. Подмножество $\mathbb{U} \subset \mathbb{M}$ называется *всюду плотным* в $\mathbb{M}$, если его замыкание $\overline{\mathbb{U}}$ совпадает со всем $\mathbb{M}$. Эквивалентно, множество $\mathbb{U}$ *всюду плотно* в $\mathbb{M}$, если в каждом открытом множестве пространства $\mathbb{M}$ содержится хотя бы одна точка из $\mathbb{U}$. □

Пример 2.3.16. Множество рациональных чисел $\mathbb{Q}$ всюду плотно в $\mathbb{R}$. □

Определение. Семейство $\{\mathbb{U}_i\}_{i\in I}$ подмножеств пространства $\mathbb{M}$ называется *покрытием* множества $\mathbb{M}$, если

$$\mathbb{M} = \bigcup_{i\in I} \mathbb{U}_i.$$

При этом множество значений индекса $i \in I$ не обязано быть счетным. Если $\mathbb{M}$ – топологическое пространство, то покрытие называется *открытым*, если все множества $\mathbb{U}_i$ открыты. Мы говорим, что некоторое покрытие имеет *подпокрытие* $\{\mathbb{U}_i\}_{i\in J}$, где $J \subset I$, если оно является покрытием само по себе. Если множество значений индекса J является конечным, то подпокрытие будет конечным. Покрытие $\{\mathbb{V}_i\}_{i\in I}$ называется *измельчением* покрытия $\{\mathbb{U}_i\}_{i\in I}$, если для всех значений индекса $\mathbb{V}_i \subset \mathbb{U}_i$. При этом некоторые из множеств $\mathbb{V}_i$ могут быть пустыми. □

Пример 2.3.17. База топологии на $\mathbb{M}$ является покрытием. □

Пример 2.3.18. Совокупность интервалов $\mathbb{U}_1 = (0, 1)$, $\mathbb{U}_2 = (0, 3/4)$, $\mathbb{U}_3 = (1/4, 1)$ является покрытием единичного интервала $(0, 1)$. Это покрытие имеет много измельчений, в том числе $\mathbb{V}_1 = \emptyset$, $\mathbb{V}_2 = (0, 2/3)$, $\mathbb{V}_3 = (1/3, 1)$. □

Замечание. В дифференциальной геометрии важную роль играют счетные измельчения покрытий. □

Определение. Топологическое пространство $\mathbb{M}$ называется *компактным*, если каждое его открытое покрытие содержит конечное подпокрытие. Хаусдорфово компактное топологическое пространство называется *компактом*. Подмножество $\mathbb{U}$ топологического пространства $\mathbb{M}$ называется *компактным*, если оно компактно в индуцированной топологии. Подмножество $\mathbb{U}$ топологического пространства $\mathbb{M}$ называется *относительно компактным*, если его замыкание $\overline{\mathbb{U}}$ компактно.

Компактные топологические пространства обладают рядом замечательных свойств. Доказательства следующих пяти теорем можно найти, например, в [**?**].

Теорема 2.3.1. *Топологическое пространство компактно тогда и только тогда, когда всякая непрерывная функция на нем ограничена, т.е. достигает на нем своего минимального и максимального значения.*

Теорема 2.3.2. *Топологическое пространство компактно тогда и только тогда, когда каждое бесконечное его подмножество имеет предельную точку.*

Теорема 2.3.3. *Если $\mathbb{U}$ – компактное подмножество хаусдорфова пространства $\mathbb{M}$ и $x \in \mathbb{M} \setminus \mathbb{U}$, то у точки x и подмножества $\mathbb{U}$ существуют непересекающиеся окрестности.*

Следствие. Каждое компактное подмножество $\mathbb{U}$ хаусдорфова пространства $\mathbb{M}$ замкнуто: $\mathbb{U} = \overline{\mathbb{U}}$.

Теорема 2.3.4. *У любых двух непересекающихся компактных подмножеств $\mathbb{U}$ и $\mathbb{V}$ хаусдорфова пространства $\mathbb{M}$ существуют непересекающиеся окрестности.*

Теорема 2.3.5. *Пусть $\{\mathbb{U}_i\}$ – открытое покрытие компактного подмножества $\mathbb{V}$ метрического пространства $(\mathbb{M}, l)$. Тогда существует такое положительное число r, что открытый шар радиуса r с центром в произвольной точке подмножества $\mathbb{U}$ содержится в некотором элементе покрытия $\mathbb{U}_i$.*

Следствие (**Лемма Лебега).** Для любого открытого покрытия $\{\mathbb{U}_i\}$ замкнутого интервала вещественных чисел существует такое положительное число r, что если $|x - y| < r$, то в покрытии найдется такой элемент $\mathbb{U}_i$, который содержит обе точки x и y.

В предыдущем разделе мы определили сходимость последовательностей в метрических пространствах $(\mathbb{M}, l)$. Дадим определение сходимости последовательностей в топологических пространствах $(\mathbb{M}, \mathcal{I})$, которое не опирается на метрическую сходимость.

Определение. Последовательность точек $\{x_i\}$, $i \in \mathbb{N}$, топологического пространства $\mathbb{M}$ сходится к точке $x \in \mathbb{M}$, если для каждой окрестности $\mathbb{U}_x$ точки x существует такое число $N_{\mathbb{U}} \in \mathbb{N}$, что $x_i \in \mathbb{U}_x$ для всех $i > N_{\mathbb{U}}$. □

Теорема 2.3.6. *Если топологическое пространство $(\mathbb{M}, \mathcal{I})$ является метризуемым $(\mathbb{M}, l)$ и полно по этой метрике, то сходящаяся последовательность точек $\{x_i\}$ в топологии $\mathcal{I}$ сходится также по метрике l. Обратно. Пусть $(\mathbb{M}, l)$ – полное метрическое пространство и $\mathcal{I}$ – топология на $\mathbb{M}$, индуцированная метрикой l. Тогда любая последовательность $\{x_i\}$, сходящаяся по метрике l, сходится также в топологии $\mathcal{I}$.*

Доказательство. □

Из этой теоремы следует, что для полных метрических пространств определения сходимости по метрике и индуцированной топологии, по-существу, эквивалентны. В то же время определение сходимости последовательностей в топологических пространствах является более общим, т.к. не всякое топологическое пространство метризуемо.

Замечание. Понятие полноты было введено для метрических пространств. Для топологических неметризуемых пространств понятие полноты не определено. □

Теорема 2.3.7. *Подмножество евклидова пространства с индуцированной топологией является компактным тогда и только тогда, когда всякая бесконечная последовательность его точек имеет в нем хотя бы одну предельную точку.*

Доказательство. См., например, [**?**]. □

Пример 2.3.19. Все евклидово пространство является некомпактным, т.к. можно выбрать последовательность точек, которая не содержит ни одной предельной точки. Например, множество натуральных чисел, лежащих на любой из координатных прямых не содержит предельной точки. □

Пример 2.3.20. Открытый шар $\mathbb{B}^n_r$ является некомпактным множеством, т.к. последовательность точек, сходящихся к некоторой граничной точке, не имеет в нем предельной точки. В то же время замкнутый шар $\overline{\mathbb{B}}^n_r$ будет уже компактным множеством. □

Пример 2.3.21. Множество с дискретной топологией компактно тогда и только тогда, когда оно конечно. □

Пример 2.3.22. Произвольное замкнутое ограниченное подмножество евклидова пространства $\overline{\mathbb{U}} \subset \mathbb{R}^n$ является *компактом*. □

Теорема 2.3.8 (**Гейне–Борель–Лебег**). *Любой компакт является компактным множеством.*

Доказательство. См., например, [**?**]. □

Теорема 2.3.9 (**Тихонов**). *Топологическое произведение произвольного множества непустых топологических пространств является компактным тогда и только тогда, когда каждый сомножитель является компактным пространством.*

Доказательство. См. [**?**]. □

Замечание. Количество сомножителей в сформулированной теореме может быть несчетным. □

Теорема Тихонова позволяет дать критерий компактности подмножеств евклидова пространства с индуцированной топологией.

Теорема 2.3.10. *Подпространство $\mathbb{M}$ n-мерного евклидова пространства $\mathbb{R}^n$ является компактом в том и только в том случае, если множество $\mathbb{M}$ замкнуто и ограничено.*

Доказательство. См., например, [**?**]. □

Определение. Топологическое пространство $\mathbb{M}$ называется *локально компактным*, если каждая точка $x \in \mathbb{M}$ имеет окрестность, замыкание которой компактно. □

Конечно, каждое компактное топологическое пространство является локально компактным. Обратное утверждение неверно.

Пример 2.3.23. Евклидово $\mathbb{R}^n$ и комплексное $\mathbb{C}^n$ пространства некомпактны, но локально компактны. □

Пример 2.3.24. Бесконечномерное гильбертово пространство не является локально компактным. □

В дальнейшем при рассмотрении многообразий необходимо будет воспользоваться существованием разбиения единицы. Для этой цели нам понадобится понятие паракомпактности, которое является слабейшим требованием, достаточным для существования разбиения единицы.

Определение. Топологическое пространство $\mathbb{M}$ называется *паракомпактным*, если любое его покрытие открытыми множествами $\{\mathbb{U}_i\}_{i\in I}$ имеет локально конечное измельчение $\{\mathbb{V}_i\}_{i\in I}$. *Локальная конечность* означает, что для каждой точки $x \in \mathbb{M}$ существует окрестность $\mathbb{W}_x \subset \mathbb{M}$ такая, что $\mathbb{V}_i \cap \mathbb{W}_x \neq \emptyset$ только для конечного числа индексов. Или, любое компактное подмножество $\mathbb{M}$ пересекается с конечным числом открытых множеств $\mathbb{V}_i$. □

Пример 2.3.25. Любое метризуемое топологическое пространство паракомпактно. В частности, евклидово пространство $\mathbb{R}^n$ является паракомпактным. □

Перечислим некоторые топологические свойства $\mathbb{R}^n$. Часть этих свойств, например, сепарабельность и хаусдорфовость, наследуется всеми многообразиями. Другие же свойства, такие как связность и односвязность, различны для различных многообразий.

Определение. Топологическое пространство называется *сепарабельным*, если оно содержит счетное всюду плотное подмножество. □

Пример 2.3.26. Евклидово пространство $\mathbb{R}^n$ является сепарабельным, т.к. в качестве такого подмножества можно выбрать, например, множество точек с рациональными координатами, которое, как известно, счетно при конечном n и всюду плотно в $\mathbb{R}^n$. □

Теорема 2.3.11. *Пространство, топология которого обладает счетной базой, сепарабельно.*

Доказательство. См., например, [**?**]. □

Определение. Топологическое пространство $\mathbb{M}$ называется *хаусдорфовым*, если для любой пары различных точек $x, y \in \mathbb{M}$ существуют открытые подмножества $\mathbb{U}_x$ и $\mathbb{U}_y$, содержащие соответственно точки x и y, такие, что $\mathbb{U}_x \cap \mathbb{U}_y = \emptyset$. □

Пример 2.3.27. Евклидово пространство является хаусдорфовым пространством. □

Хаусдорфовость топологического пространства важна при определении предела последовательностей. Если мы хотим, чтобы у любой последовательности точек мог существовать не более, чем один предел, необходимо потребовать, чтобы пространство было хаусдорфовым.

Теорема 2.3.12. *Топологическое пространство является хаусдорфовым тогда и только тогда, когда никакая последовательность в этом пространстве не сходится к двум различным точкам.*

Доказательство. См., например, [**?**]. □

Пример 2.3.28. Дадим пример нехаусдорфова топологического пространства. Рассмотрим множество точек, состоящее из объединения вещественной прямой $y = 0$ и точки $(0, 1) \in \mathbb{R}^2$ на евклидовой плоскости, изображенных на рис.2.2, *a*. Определим

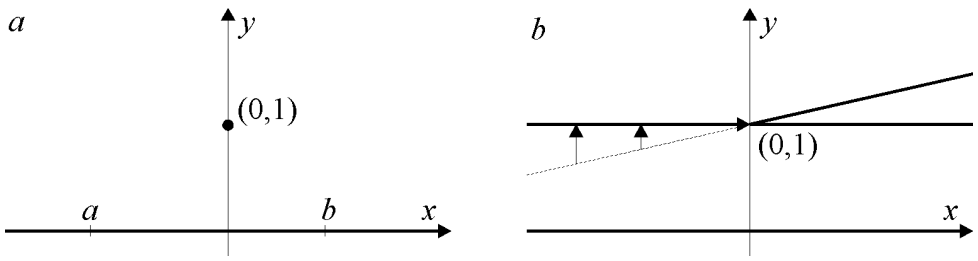


Рис. 2.2: Примеры 2.3.28 и 2.3.29 нехаусдорфовых топологических пространств.

топологию на рассматриваемом множестве следующим образом. Совокупность открытых интервалов (p, q) на оси x, а также множества, состоящие из объединения

точки $(0,1)$ с интервалами (p,q), $p<0$, $q>0$ оси x с выколотым началом координат, будем считать базой топологии. По-построению, любые окрестности начала координат $(0,0)\in\mathbb{R}^2$ и точки $(0,1)\in\mathbb{R}^2$ имеют непустое пересечение, и, следовательно, построенное топологическое пространство не является хаусдорфовым. Любая последовательность точек, сходящаяся к началу координат $(0,0)$, сходится также и к точке $(0,1)$, и наоборот. Другими словами, с точки зрения сходящихся последовательностей две точки $(0,0)$ и $(0,1)$ неразличимы. □

Пример 2.3.29. Рассмотрим две прямые на плоскости с естественной топологией, проходящие через точку $(0,1)\in\mathbb{R}^2$, и отождествим их точки с отрицательными абсциссами, $x<0$, как показано на рис.2.2,*b*. В результате получим некоторое топологическое пространство $\mathbb{M}$. Тогда одна точка плоскости $(0,1)\in\mathbb{R}^2$ соответствует двум различным точкам в $\mathbb{M}$, которые лежат на разных прямых и неотождествлены. Эти точки различны, и в тоже время не имеют непересекающихся окрестностей. □

Замечание. В приведенных примерах нехаусдорфовых пространств их топология не является топологией, индуцированной из $\mathbb{R}^2$. □

Теорема 2.3.13. *Любое метризуемое топологическое пространство* $\mathbb{M}$ *является хаусдорфовым.*

Доказательство. Выберем две произвольные различные точки $x\neq y$ в $\mathbb{M}$ и построим два непересекающихся шара с центрами в x и y. □

Дадим два эквивалентных определения связности топологических пространств.

Определение. Топологическое пространство $\mathbb{M}$ называется *связным*,
1) если его нельзя представить в виде объединения двух непересекающихся множеств, каждое из которых не содержит предельных точек другого, или
2) если единственными подмножествами $\mathbb{M}$, открытыми и замкнутыми одновременно, являются $\emptyset$ и $\mathbb{M}$.

Компонентой топологического пространства называется максимальное связное подпространство. □

Если пространство связно, то оно состоит из одной компоненты.

Предложение 2.3.3. *Если топологическое пространство является объединением двух непустых непересекающихся открытых подмножеств,* $\mathbb{M}=\mathbb{U}_1\cup\mathbb{U}_2$, *где* $\mathbb{U}_1\cap\mathbb{U}_2=\emptyset$, *то оно не является связным.*

Доказательство. От противного. Пусть $x\in\mathbb{U}_1$ – предельная точка для $\mathbb{U}_2$. Выберем окрестность $\mathbb{U}_x\subset\mathbb{U}_1$, содержащую точку x. Тогда она содержит точки множества $\mathbb{U}_2$, и, следовательно, подмножества $\mathbb{U}_1$ и $\mathbb{U}_2$ пересекаются. □

Пример 2.3.30. Евклидово пространство $\mathbb{R}^n$ является связным. Произвольное открытое связное и односвязное (см. раздел 11.3) подмножество $\mathbb{U}\subset\mathbb{R}^n$ называется *областью*. □

Пример 2.3.31. Рассмотрим сферу, вложенную в евклидово пространство, $\mathbb{S}^{n-1}\hookrightarrow\mathbb{R}^n$. Будем считать, что топология на сфере индуцирована вложением. Тогда сфера – связное топологическое пространство. □

Несвязные топологические пространства $\mathbb{M}$ представляют собой объединение конечного или бесконечного числа компонент. В последнем случае число компонент может быть счетным или несчетным. Каждая точка принадлежит только одной компоненте, которая называется компонентой данной точки. Всякая компонента является замкнутым подмножеством в $\mathbb{M}$.

Замечание. В общем случае компонента топологического пространства может не быть открытой. Экзотический пример описан в [**?**]. □

В дальнейшем будут использованы следующие два утверждения, доказательство которых приведено, например, в [**?**].

Теорема 2.3.14. *Если произвольное семейство $\{\mathbb{U}_i\}$, $i \in I$, связных подпространств топологического пространства $\mathbb{M}$ имеет непустое пересечение, то его объединение $\cup_{i\in I}\mathbb{U}_i$ связно.*

Теорема 2.3.15. *Топологическое произведение произвольного множества непустых топологических пространств связно в том и только в том случае, если все сомножители связны.*

Определение. Топологическое пространство $\mathbb{M}$ называется *локально связным*, если для каждой точки $x \in \mathbb{M}$ и любой ее окрестности $\mathbb{U}_x$ существует связное множество $\mathbb{V} \subset \mathbb{U}_x$, такое, что $x \in \mathsf{int}\,\mathbb{V}$. □

Пример 2.3.32. Евклидово пространство $\mathbb{R}^n$ локально связно. □

Очевидно, что всякое связное топологическое пространство является локально связным. Обратное утверждение неверно.

Пример 2.3.33. Рассмотрим подмножество точек плоскости $\mathbb{R}^2$, состоящее из прямых, проходящих через начало координат и заданных уравнениями $mx + ny = 0$ с целыми коэффициентами m и n. Будем считать, что топология этого подмножества индуцирована плоскостью. Это множество связно, как объединение связных множеств, имеющих общую точку. В то же время оно не является локально связным, т.к. любое связное подпространство обязано содержать начало координат. □

Теорема 2.3.16. *Топологическое пространство $\mathbb{M}$ локально связно в том, и только в том случае, если компоненты открытых множеств открыты. В частности, любая окрестность любой точки локально связного пространства содержит связную окрестность этой точки.*

Доказательство. См., например, [**?**]. □

Теорема 2.3.17. *В локально связном топологическом пространстве $\mathbb{M}$ каждая компонента является одновременно открытой и замкнутой.*

Доказательство. См., например, []. □

2.3.3 $\mathbb{R}^n$ как векторное пространство

Если на евклидовом пространстве $\mathbb{R}^n$ задана евклидова метрика и индуцированная топология, то этого достаточно для определения дифференцируемого многообразия. Более того, метрику можно и не задавать, понимая под евклидовым пространством просто топологическое произведение n прямых. Вместе с этим на евклидовом пространстве помимо метрики и топологии можно ввести дополнительную алгебраическую структуру – структуру векторного или линейного пространства. Она понадобится для определения касательного расслоения к многообразию. Кроме этого, структура векторного пространства играет большую роль в приложениях, т.к. пространство-время в нерелятивистских моделях и специальной теории относительности имеет структуру векторного пространства.

Пусть $\mathbb{R}^n$ – евклидово пространство с декартовыми координатами x^α, $\alpha = 1, \ldots, n$.

Определение. Назовем *вектором* $\overrightarrow{\mathrm{PQ}}$ евклидова пространства $\mathbb{R}^n$ с началом в точке $\mathrm{P} = \{x^\alpha_{\mathrm{P}}\}$ и концом в точке $\mathrm{Q} = \{x^\alpha_{\mathrm{Q}}\}$ упорядоченный набор чисел $\{x^\alpha_{\mathrm{Q}} - x^\alpha_{\mathrm{P}}\}$, которые называются компонентами вектора. □

В евклидовой геометрии вектор представляется в виде направленного отрезка прямой линии, соединяющего точки P и Q.

Рассмотрим множество векторов $\mathbb{V}$ с началом в начале системы координат. Для их обозначения будем использовать жирный наклонный шрифт

$$\boldsymbol{x} = \{x^\alpha\} \in \mathbb{V}.$$

Пространство векторов $\mathbb{V}$ находится во взаимно однозначном соответствии с точками евклидова пространства $\mathbb{R}^n$, при этом вектор $\boldsymbol{x} \in \mathbb{V}$ называется *радиусом-вектором* точки $\{x^\alpha\} \in \mathbb{R}^n$. Назовем *нулевым* вектором $\mathbf{0}$ вектор с нулевыми компонентами, $\mathbf{0} = (0, \ldots, 0) \in \mathbb{V}$. Вектор $-\boldsymbol{x}$ с компонентами $\{-x^\alpha\} \in \mathbb{V}$ назовем обратным вектором к вектору $\boldsymbol{x}$. Для любых двух векторов $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$ определим их сумму $\boldsymbol{x} + \boldsymbol{y}$ как вектор $\boldsymbol{z}$ с компонентами

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} = \{z^\alpha := x^\alpha + y^\alpha\} \in \mathbb{V}. \tag{2.18}$$

Из свойств сложения вещественных чисел следует, что операция сложения векторов коммутативна: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$. Пространство векторов $\mathbb{V}$ с введенной операцией сложения образует абелеву группу или модуль, поскольку сложение векторов коммутативно и выполнены необходимые групповые аксиомы:

$$\begin{aligned}
&1)\ \ \boldsymbol{x} + \boldsymbol{y} \in \mathbb{V} && \text{– замкнутость по отношению к сложению;}\\
&2)\ \ (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) && \text{– ассоциативность;}\\
&3)\ \ \boldsymbol{x} + \mathbf{0} = \boldsymbol{x} && \text{– существование единичного элемента;}\\
&4)\ \ \boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0} && \text{– существование обратного элемента.}
\end{aligned} \tag{2.19}$$

Напомним, что для абелевых групп групповую операцию принято называть сложением, а единичный элемент – нулем.

Замечание. Понятие вектора в дифференциальной геометрии является инвариантным и не зависит от выбора системы координат. Приведенные выше определения были сделаны в декартовой системе координат сознательно, т.к. линейная структура наиболее просто задается в декартовых координатах. Векторное пространство $\mathbb{V}$ можно было бы описать и в произвольных координатах. Однако в криволинейной системе координат операция сложения векторов выглядит довольно громоздко. □

Определим новую операцию на множестве векторов $\mathbb{V}$. А именно, введем операцию умножения векторов на действительные числа $a, b, \ldots \in \mathbb{R}$, которые в данном случае принято называть *скалярами*,

$$\mathbb{R} \times \mathbb{V} \ni a, \boldsymbol{x} \mapsto a\boldsymbol{x} = \{ax^\alpha\} \in \mathbb{V}. \tag{2.20}$$

Здесь каждая компонента вектора умножается на одно и то же число a. Эта операция обладает всеми свойствами, которые перечислены в следующем общем определении векторного пространства.

Определение. Абелева группа $\mathbb{V} = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \ldots\}$ с операцией сложения (2.19) и умножения на скаляры (2.20), которые удовлетворяют свойствам:

$$\begin{array}{lll}
1) & a\boldsymbol{x} \in \mathbb{V} & \text{– замкнутость по отношению к умножению на скаляры;} \\
2) & a(\boldsymbol{x} + \boldsymbol{y}) = a\boldsymbol{x} + a\boldsymbol{y} & \text{– дистрибутивность по отношению к сложению векторов;} \\
3) & (a + b)\boldsymbol{x} = a\boldsymbol{x} + b\boldsymbol{x} & \text{– дистрибутивность по отношению к сложению скаляров;} \\
4) & (ab)\boldsymbol{x} = a(b\boldsymbol{x}) & \text{– ассоциативность;} \\
5) & 1 \cdot \boldsymbol{x} = \boldsymbol{x} & \text{– умножение на единицу.}
\end{array} \tag{2.21}$$

называется *линейным пространством*, *линейным векторным пространством* или просто *векторным пространством*. Элементы пространства $\mathbb{V}$ называются *векторами*. Мы также говорим, что векторное пространство – это модуль над полем вещественных чисел. □

Замечание. В определении сложения векторов евклидова пространства (2.18) и их умножения на скаляры (2.20) мы складываем и умножаем компоненты как вещественные числа. В абстрактном подходе рассматривают множество векторов $\mathbb{V}$ как произвольную абелеву группу, в которой задано умножение на скаляры со свойствами (2.21). В этом случае можно забыть про исходное евклидово пространство и говорить про декартовы или криволинейные координаты не имеет смысла. Если в определении векторного пространства заменить вещественные числа на комплексные, то получим векторное пространство над полем комплексных чисел. □

Векторное пространство $\mathbb{V} \simeq \mathbb{R}^n$ было построено для векторов с началом в начале координат. Естественным образом можно построить векторное пространство в каждой точке евклидова пространства $\mathbb{V}_x$, $x \in \mathbb{R}^n$. Все эти векторные пространства изоморфны между собой. Если объединить все векторные пространства $\mathbb{V}_x$, то получим прямое произведение $\mathbb{R}^n \times \mathbb{V}$ размерности $2n$, где первый сомножитель рассматривается не как векторное, а как топологическое пространство. Такое объединение является *касательным расслоением* над $\mathbb{R}^n$ и обозначается

$$\mathbb{T}(\mathbb{R}^n) = \bigcup_{x \in \mathbb{R}^n} \mathbb{V}_x = \mathbb{R}^n \times \mathbb{V}.$$

При этом $\mathbb{R}^n$ называется базой расслоения, а $\mathbb{V}$ – типичным слоем расслоения или просто *слоем*. Об этом речь пойдет в разделе **??**.

В произвольном векторном пространстве можно ввести важное понятие *базиса*, как такой набор линейно независимых векторов $\{\boldsymbol{e}_\alpha\}$, что произвольный вектор можно представить в виде линейной комбинации базисных векторов:

$$\boldsymbol{x} = x^\alpha \boldsymbol{e}_\alpha. \tag{2.22}$$

В этом разложении числа $x^\alpha \in \mathbb{R}$ называются компонентами вектора $\boldsymbol{x}$ по отношению к базису $\boldsymbol{e}_\alpha$. Мы будем говорить, что векторное пространство $\mathbb{V}$ натянуто на векторы $\boldsymbol{e}_\alpha$, т.е. состоит из всех линейных комбинаций вида (2.22). Базис векторного пространства определен неоднозначно, но число базисных векторов от выбора базиса не зависит. Это число называется *размерностью* векторного пространства, и мы пишем

$$\mathsf{dim}\,\mathbb{V} = n, \qquad n \in \mathbb{N}.$$

Другими словами, размерностью линейного пространства называется максимальное число линейно независимых векторов. Эти векторы можно выбрать в качестве базиса векторного пространства. В общем случае размерность n может быть бесконечна, однако в настоящей монографии мы рассматриваем, как правило, конечномерные пространства. По построению, базисные векторы имеют следующие компоненты:

$$\boldsymbol{e}_\alpha = (\underbrace{0,\ldots,0}_{\alpha-1},1,0,\ldots,0), \tag{2.23}$$

где на α-том месте стоит единица, а остальные компоненты равны нулю.

Пусть задано произвольное векторное пространство $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, с фиксированным базисом. Тогда каждый вектор $\boldsymbol{x} \in \mathbb{V}$ взаимно однозначно задается своими компонентами в этом базисе, $\boldsymbol{x} \leftrightarrow \{x^\alpha\}$, $\alpha = 1,\ldots,n$. Сложение векторов и умножение на скаляры в компонентах задается при этом точно так же, как и для векторов евклидова пространства $\mathbb{R}^n$. Поэтому произвольное векторное пространство $\mathbb{V}$ с фиксированным базисом изоморфно евклидову пространству $\mathbb{R}^n$ той же размерности. Тем самым на любом векторном пространстве можно ввести структуру многообразия, при этом оно становится тривиальным многообразием, диффеоморфном $\mathbb{R}^n$.

Пример 2.3.34. Произвольное поле $\mathbb{F}$ можно рассматривать как одномерное векторное пространство с базисом $\boldsymbol{e}$. Базисным вектором в этом случае может являться любой элемент $\boldsymbol{e} \in \mathbb{F}$, отличный от нуля. □

Пример 2.3.35. Естественным базисом евклидова пространства $\mathbb{R}^n$, которое рассматривается, как векторное пространство, является набор единичных векторов, направленных вдоль декартовых координатных осей. □

Пример 2.3.36. Множество абсолютно интегрируемых функций на отрезке $[-\pi,\pi]$ с поточечным сложением и умножением на действительные числа является примером бесконечномерного векторного пространства. При этом разложение в ряд Фурье представляет собой разложение по счетному базису, состоящему из тригонометрических функций. □

В приложениях часто используются также комплексные векторные пространства. Если задано вещественное векторное пространство, то оно определяет также комплексное векторное пространство.

Определение. *Комплексификацией* вещественного векторного пространства $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, называется комплексное векторное пространство ${}^{\mathbb{C}}\mathbb{V}$, комплексной размерности $\mathsf{dim}_{\,\mathbb{C}}\mathbb{V} = n$, состоящее из пар $(\boldsymbol{x},\boldsymbol{y})$, которые обозначаются $\boldsymbol{z} := \boldsymbol{x} + i\boldsymbol{y}$, где $\boldsymbol{x},\boldsymbol{y} \in \mathbb{V}$, с обычными операциями сложения и умножения на комплексные числа. Комплексификацией линейного оператора $A:\ \mathbb{V} \to \mathbb{V}$ называется линейный оператор

$${}^{\mathbb{C}}A:\quad {}^{\mathbb{C}}\mathbb{V} \ni \boldsymbol{x} + i\boldsymbol{y} \mapsto {}^{\mathbb{C}}A(\boldsymbol{x} + i\boldsymbol{y}) := A\boldsymbol{x} + iA\boldsymbol{y} \in {}^{\mathbb{C}}\mathbb{V}. \qquad \square$$

Определение. Отображение двух векторных пространств

$$f: \quad \mathbb{V} \ni \boldsymbol{x} \mapsto \boldsymbol{y} \in \mathbb{W}$$

называется *линейным*, если

$$f(a\boldsymbol{x}_1 + b\boldsymbol{x}_2) = af(\boldsymbol{x}_1) + bf(\boldsymbol{x}_2), \qquad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{V}, \quad a, b \in \mathbb{R}.$$

Отображение f называется также *линейным оператором.* □

Если в пространствах $\mathbb{V}$ и $\mathbb{W}$ заданы базисы $\boldsymbol{e}_\alpha$, $\alpha = 1, \ldots, \mathsf{dim}\,\mathbb{V}$, и $\boldsymbol{e}_i$, $i = 1, \ldots, \mathsf{dim}\,\mathbb{W}$, то произвольный линейный оператор задается матрицей ${A_\alpha}^i$:

$$f: \quad \mathbb{V} \ni \boldsymbol{x} = x^\alpha \boldsymbol{e}_\alpha \mapsto \boldsymbol{y} = y^i \boldsymbol{e}_i \in \mathbb{W}, \qquad y^i = x^\alpha {A_\alpha}^i,$$

где $f(\boldsymbol{e}_\alpha) := {A_\alpha}^i \boldsymbol{e}_i$.

Если компоненты всех векторов (2.22) преобразовать с помощью невырожденной матрицы

$$x^\alpha \mapsto x^\beta {A_\beta}^\alpha, \qquad \mathsf{det}\, A \neq 0, \tag{2.24}$$

то получим взаимно однозначное отображение векторного пространства на себя, при котором все линейные свойства пространства сохраняются. То есть каждая невырожденная матрица задает некоторый автоморфизм векторного пространства. Верно и обратное утверждение. Любой автоморфизм векторного пространства в компонентах задается некоторой невырожденной матрицей. Совокупность матриц A образует группу, которая называется группой линейных однородных преобразований векторного пространства и обозначается $\mathbb{GL}(n, \mathbb{R})$. При преобразовании координат (2.24) с произвольной матрицей $A \in \mathbb{GL}(n, \mathbb{R})$ нулевой вектор не меняется. Поэтому группа общих линейных преобразований действует эффективно, но не свободно (см. раздел 10.1).

Если некоторое подмножество векторного пространства $\mathbb{V}_1 \subset \mathbb{V}$ само является векторным пространством, то оно называется *линейным подпространством.* Пусть $\mathbb{V}_1$ и $\mathbb{V}_2$ – два подпространства векторного пространства $\mathbb{V}$. Нетрудно проверить, что их пересечение $\mathbb{V}_1 \cap \mathbb{V}_2$ также является векторным пространством. В то же время их объединение в общем случае векторным пространством не будет. Тем не менее мы будем писать $\mathbb{V}_1 \cup \mathbb{V}_2$, когда понадобиться необходимость сказать, что некоторый элемент принадлежит либо пространству $\mathbb{V}_1$, либо пространству $\mathbb{V}_2$.

Пример 2.3.37. Пусть $\mathbb{V} = \mathbb{R}^3$. Рассмотрим две плоскости $\mathbb{V}_1$ и $\mathbb{V}_2$ натянутые на базисные векторы $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ и $(\boldsymbol{e}_2, \boldsymbol{e}_3)$ соответственно. Тогда их пересечение $\mathbb{V}_1 \cap \mathbb{V}_2$ является прямой линией с направляющим вектором $\boldsymbol{e}_2$. Тем самым пересечение является одномерным векторным пространством. Объединение $\mathbb{V}_1 \cup \mathbb{V}_2$ представляет собой объединение двух перпендикулярных плоскостей, что, конечно, не совпадает с трехмерным евклидовым пространством $\mathbb{R}^3$. □

Очевидно, что

$$\mathsf{dim}\,(\mathbb{V}_1 \cap \mathbb{V}_2) \leq \mathsf{min}\,\{\,\mathsf{dim}\,\mathbb{V}_1,\, \mathsf{dim}\,\mathbb{V}_2\}.$$

Определение. Два векторных подпространства $\mathbb{V}_1$ и $\mathbb{V}_2$ конечномерного векторного пространства $\mathbb{V}$ называются *трансверсальными* друг к другу, если они порождают все $\mathbb{V}$, т.е.

$$\mathsf{dim}\,(\mathbb{V}_1 \cap \mathbb{V}_2) + \mathsf{dim}\,\mathbb{V} = \mathsf{dim}\,\mathbb{V}_1 + \mathsf{dim}\,\mathbb{V}_2.$$ □

В частном случае, если подпространства не пересекаются, $\mathbb{V}_1 \cap \mathbb{V}_2 = \emptyset$, то $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$.

Векторное пространство $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, над полем вещественных чисел образует абелеву группу по отношению к сложению или модуль. Рассмотрим его подпространство $\mathbb{H} \subset \mathbb{V}$, состоящее из всех линейных комбинаций

$$\sum_{i=1}^{p} a^i \boldsymbol{x}_i \in \mathbb{H}, \quad a^i \in \mathbb{R}, \quad \boldsymbol{x}_i \in \mathbb{V},$$

где $p < n$. Не ограничивая общности, можно считать, что все векторы $\boldsymbol{x}_i$ линейно независимы. Тем самым мы предполагаем, что $\mathsf{dim}\,\mathbb{H} = p$. Подпространство $\mathbb{H}$ образует подгруппу векторного пространства $\mathbb{V}$ или подмодуль. Эта подгруппа является нормальной, поскольку группа абелева. Отсюда следует, что существует факторгруппа или факторпространство $\mathbb{V}/\mathbb{H}$. Элементами факторпространства являются все линейные комбинации вида

$$\boldsymbol{x} + \sum_{i=1}^{p} a^i \boldsymbol{x}_i \in \boldsymbol{x} + \mathbb{H}, \quad \text{где } \boldsymbol{x} \in \mathbb{V},$$

со следующим отношением эквивалентности. Два элемента факторпространства $\boldsymbol{x} + \mathbb{H}$ и $\boldsymbol{y} + \mathbb{H}$ совпадают, если их разность лежит в $\mathbb{H}$:

$$\boldsymbol{y} - \boldsymbol{x} \in \mathbb{H}.$$

В факторпространстве $\mathbb{V}/\mathbb{H}$ естественным образом вводится умножение на вещественные числа, что превращает его само в векторное пространство. Очевидно, что $\mathsf{dim}\,\mathbb{V}/\mathbb{H} = \mathsf{dim}\,\mathbb{V} - \mathsf{dim}\,\mathbb{H} = n - p$.

Пример 2.3.38. Пусть подпространство $\mathbb{H} \subset \mathbb{V}$ натянуто на $p < n$ первых базисных векторов векторного пространства $\mathbb{V}$

$$\mathbb{H} = \{\boldsymbol{x} \in \mathbb{V} : \quad \boldsymbol{x} = \sum_{i=1}^{p} a^i \boldsymbol{e}_i\}.$$

Тогда факторпространство $\mathbb{V}/\mathbb{H}$ изоморфно векторному пространству, натянутому на оставшиеся базисные векторы:

$$\frac{\mathbb{V}}{\mathbb{H}} \simeq \{\boldsymbol{x} \in \mathbb{V} : \quad \boldsymbol{x} = \sum_{k=p+1}^{n} a^k \boldsymbol{e}_k\}. \qquad \square$$

Определение. *Прямой линией*, проходящей через две точки векторного пространства $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$ называется множество векторов

$$\{\boldsymbol{z} \in \mathbb{V} : \quad \boldsymbol{z} = t\boldsymbol{x} + (1-t)\boldsymbol{y}, \quad \forall t \in \mathbb{R}\}. \tag{2.25}$$

Подмножество векторного пространства $\mathbb{U} \subset \mathbb{V}$ называется *плоским*, если вместе с любыми двумя точками, оно содержит прямую, проходящую через эти точки. Если параметр пробегает единичный отрезок, $t \in [0, 1]$, то мы имеем отрезок, соединяющий точки $\boldsymbol{x}$ и $\boldsymbol{y}$. Подмножество точек $\mathbb{U}$ векторного пространства называется *выпуклым*, если вместе с двумя произвольными точками $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{U}$ оно содержит и отрезок, соединяющий эти точки. $\square$

Прямая сумма

Если задано два векторных пространств $\mathbb{V}_1$ и $\mathbb{V}_2$, то с их помощью можно построить новые векторные пространства: прямую сумму и тензорное произведение.

Определение. *Прямой суммой* двух векторных пространств называется векторное пространство $\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2$ образованное всеми парами $\boldsymbol{x}_1 \oplus \boldsymbol{x}_2$, где $\boldsymbol{x}_1 \in \mathbb{V}_1$ и $\boldsymbol{x}_2 \in \mathbb{V}_2$, с векторным сложением и умножением на скаляры, определяемыми формулами:

$$\boldsymbol{x}_1 \oplus \boldsymbol{x}_2 + \boldsymbol{y}_1 \oplus \boldsymbol{y}_2 = (\boldsymbol{x}_1 + \boldsymbol{y}_1) \oplus (\boldsymbol{x}_2 + \boldsymbol{y}_2),$$
$$a(\boldsymbol{x}_1 \oplus \boldsymbol{x}_2) = (a\boldsymbol{x}_1) \oplus (a\boldsymbol{x}_2),$$

где $\boldsymbol{y}_1 \oplus \boldsymbol{y}_2$ также принадлежит $\mathbb{V}_1 \oplus \mathbb{V}_2$. □

Нетрудно проверить, что все аксиомы векторного пространства для прямой суммы выполнены.

Размерность прямой суммы векторных пространств равна сумме размерностей пространств $\mathbb{V}_1$ и $\mathbb{V}_2$,

$$\mathsf{dim}\,(\mathbb{V}_1 \oplus \mathbb{V}_2) = \mathsf{dim}\,\mathbb{V}_1 \oplus \mathsf{dim}\,\mathbb{V}_2.$$

Пусть в векторных пространствах заданы базисы $\boldsymbol{e}_{\alpha_1}$, $\alpha_1 = 1, \ldots, \mathsf{dim}\,\mathbb{V}_1$, и $\boldsymbol{e}_{\alpha_2}$, $\alpha_2 = 1, \ldots, \mathsf{dim}\,\mathbb{V}_2$. Тогда прямые суммы $\boldsymbol{e}_{\alpha_1} \oplus \boldsymbol{e}_{\alpha_2}$ образуют базис в $\mathbb{V}_1 \oplus \mathbb{V}_2$. Компонентами прямой суммы двух векторов в этом базисе является упорядоченный набор $\mathsf{dim}\,\mathbb{V}_1 + \mathsf{dim}\,\mathbb{V}_2$ чисел, состоящий из компонент первого и второго вектора:

$$\boldsymbol{x}_1 \oplus \boldsymbol{x}_2 = \{x_1^{\alpha_1}, x_2^{\alpha_2}\}.$$

Пример 2.3.39. Вещественную прямую $\mathbb{R}$ можно естественным образом рассматривать, как одномерное векторное пространство. Тогда векторное пространство $\mathbb{R}^n$ есть прямая сумма n одномерных векторных пространств.

$$\mathbb{R}^n = \underbrace{\mathbb{R} \oplus \ldots \oplus \mathbb{R}}_{n}.$$ □

Векторные пространства $\mathbb{V}_1$ и $\mathbb{V}_2$ можно рассматривать, как подпространства $\mathbb{V}_1 \oplus \mathbf{0}$ и $\mathbf{0} \oplus \mathbb{V}_2$ в прямой сумме $\mathbb{V}_1 \oplus \mathbb{V}_2$. Поэтому они изоморфны следующим факторпространствам

$$\mathbb{V}_1 \simeq \frac{\mathbb{V}_1 \oplus \mathbb{V}_2}{\mathbb{V}_2} \quad \text{и} \quad \mathbb{V}_2 \simeq \frac{\mathbb{V}_1 \oplus \mathbb{V}_2}{\mathbb{V}_1}$$

Обобщением разложения вектора по базису (2.22) служит градуировка векторного пространства.

Определение. *Градуированным векторным (линейным) пространством* называется векторное пространство $\mathbb{V}$ вместе с его разложением в прямую сумму подпространств

$$\mathbb{V} = \bigoplus_{i=1}^{\infty} \mathbb{V}_i, \tag{2.26}$$

где некоторые подпространства могут быть пустыми. Суммарная размерность подпространств в сумме должна быть равна размерности $\mathbb{V}$. В общем случае бесконечномерных векторных пространств эта сумма бесконечна, но каждый отдельный вектор $\boldsymbol{x} \in \mathbb{V}$ однозначно представляется в виде конечной суммы

$$\boldsymbol{x} = \sum_{i=1}^{\infty} \boldsymbol{x}_i, \qquad \boldsymbol{x}_i \in \mathbb{V}_i.$$

где все $\boldsymbol{x}_i$ кроме конечного числа равны нулю. Вектор $\boldsymbol{x}_i$ называется *однородной компонентой* вектора $\boldsymbol{x}$ степени i. Если $x \in \mathbb{V}_i$, то вектор $\boldsymbol{x}$ называется *однородным элементом* степени i. □

Пример 2.3.40. Евклидово векторное пространство $\mathbb{R}^n$ представляет собой прямую сумму n прямых, что можно рассматривать как градуировку. При этом количество подпространств в разложении (2.26) является конечным. На одном и том же векторном пространстве можно задать несколько различных градуировок. Например, евклидово пространство можно представить в виде суммы двух подпространств:

$$\mathbb{R}^n = \mathbb{R}^m \oplus \mathbb{R}^{n-m}, \quad 1 \leq m \leq n-1. \qquad \square$$

Определение. Одним из способов изучения множеств $\mathbb{M}$ с заданными алгебраическими структурами состоит в выделении в них последовательности подмножеств $\mathbb{M}_0 \subset \mathbb{M}_1 \subset \mathbb{M}_2 \subset \ldots$ или $\mathbb{M}_0 \supset \mathbb{M}_1 \supset \mathbb{M}_2 \supset \ldots$ таким образом, что переход от одного подмножества к другому устроен к каком то смысле просто. Общее название таких последовательностей – *фильтрации (возрастающая и убывающая* соответственно). В теории линейных пространств строго возрастающая последовательность подпространств $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \ldots \subset \mathbb{V}_n$ пространства $\mathbb{V}$ называется *флагом*. Число n называется длиной флага. Флаг $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \ldots \subset \mathbb{V}_i \subset \ldots$ называется *максимальным*, если $\mathbb{V}_0 = \{0\}$, $\cup_i \mathbb{V}_i = \mathbb{V}$ и между $\mathbb{V}_i$ и $\mathbb{V}_{i+1}$ нельзя вставить подпространство, т.е. если $\mathbb{V}_i \subset \mathbb{U} \subset \mathbb{V}_{i+1}$, то либо $\mathbb{U} = \mathbb{V}_i$, либо $\mathbb{U} = \mathbb{V}_{i+1}$. □

Замечание. Мотивировка названия: флаг {точка 0} $\subset$ {прямая} $\subset$ {плоскость} – это “гвоздь”, “древко” и “полотнище”. □

По всякому базису $\boldsymbol{e}_\alpha$, $\alpha = 1, \ldots, n$, векторного пространства $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, можно построить флаг следующим образом. Положим $\mathbb{V}_0 := \{0\}$ и пусть $\mathbb{V}_i$ – линейная оболочка первых i базисных векторов $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i\}$ при $1 \leq i \leq n$. Нетрудно проверить, что построенный флаг максимален.

Предложение 2.3.4. *Размерность векторного пространства $\mathbb{V}$ равна длине любого его максимального флага.*

Доказательство. См., например, [**?**]. □

В конечномерном линейном пространстве $\mathbb{V}$ любой флаг можно дополнить до максимального, и поэтому его длина всегда $\leq \mathsf{dim}\,\mathbb{V}$.

Тензорная алгебра

Перейдем к описанию более сложного понятия тензорного произведения векторных пространств. Обозначим через $\mathbb{V}_1 \times \mathbb{V}_2$ прямое произведение векторных пространств. Оно не снабжено структурой линейного пространства и просто обозначает множество упорядоченных пар элементов $(\boldsymbol{x}_1, \boldsymbol{x}_2)$.

Определение. Пусть $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ – векторное пространство над полем вещественных чисел, свободно порожденное элементами вида $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathbb{V}_1 \times \mathbb{V}_2$. То есть $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ состоит из всех конечных линейных комбинаций пар $(\boldsymbol{x}_1, \boldsymbol{x}_2)$. А именно, мы рассматриваем все конечные линейные комбинации

$$\sum_{ij} a_{ij}(\boldsymbol{x}_{1i}, \boldsymbol{x}_{2j}), \quad \boldsymbol{x}_{1i} \in \mathbb{V}_1, \quad \boldsymbol{x}_{2j} \in \mathbb{V}_2, \quad a_{ij} \in \mathbb{R}.$$

Построенное векторное пространство $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ является бесконечномерным, т.к. базисом этого пространства являются все элементы вида $(\boldsymbol{x}_1, \boldsymbol{x}_2)$, которых бесконечно много.

Пусть $\mathbb{H}(\mathbb{V}_1, \mathbb{V}_2)$ – линейное подпространство в $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$, свободно порожденное всеми элементами вида

$$\begin{aligned} (\boldsymbol{x}_1, \boldsymbol{x}_2 + \boldsymbol{y}_2) - (\boldsymbol{x}_1, \boldsymbol{x}_2) - (\boldsymbol{x}_1, \boldsymbol{y}_2), \\ (\boldsymbol{x}_1 + \boldsymbol{y}_1, \boldsymbol{x}_2) - (\boldsymbol{x}_1, \boldsymbol{x}_2) - (\boldsymbol{y}_1, \boldsymbol{x}_2), \\ (a\boldsymbol{x}_1, \boldsymbol{x}_2) - a(\boldsymbol{x}_1, \boldsymbol{x}_2), \\ (\boldsymbol{x}_1, a\boldsymbol{x}_2) - a(\boldsymbol{x}_1, \boldsymbol{x}_2), \end{aligned} \tag{2.27}$$

где $a \in \mathbb{R}$. Это пространство также бесконечномерно. Тогда факторпространство

$$\mathbb{V}_1 \otimes \mathbb{V}_2 = \frac{\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)}{\mathbb{H}(\mathbb{V}_1, \mathbb{V}_2)}$$

называется *тензорным произведением* векторных пространств. При этом каждой паре элементов $\boldsymbol{x}_1 \in \mathbb{V}_1$ и $\boldsymbol{x}_2 \in \mathbb{V}_2$ ставится в соответствие их *тензорное произведение* $\boldsymbol{x}_1 \otimes \boldsymbol{x}_2 \in \mathbb{V}_1 \otimes \mathbb{V}_2$. □

Тензорное произведение некоммутативно, т.к. порядок сомножителей важен и фиксирован.

Из определения следует, что тензорное произведение векторов *билинейно*, т.е. линейно по каждому из сомножителей при фиксированном другом:

$$\begin{aligned} \boldsymbol{x}_1 \otimes (\boldsymbol{x}_2 + \boldsymbol{y}_2) &= \boldsymbol{x}_1 \otimes \boldsymbol{x}_2 + \boldsymbol{x}_1 \otimes \boldsymbol{y}_2, \\ (\boldsymbol{x}_1 + \boldsymbol{y}_1) \otimes \boldsymbol{x}_2 &= \boldsymbol{x}_1 \otimes \boldsymbol{x}_2 + \boldsymbol{y}_1 \otimes \boldsymbol{x}_2, \\ (a\boldsymbol{x}_1) \otimes \boldsymbol{x}_2 &= a(\boldsymbol{x}_1 \otimes \boldsymbol{x}_2), \\ \boldsymbol{x}_1 \otimes (a\boldsymbol{x}_2) &= a(\boldsymbol{x}_1 \otimes \boldsymbol{x}_2), \end{aligned} \tag{2.28}$$

так как элементы $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ вида (2.27) отождествлены.

Пусть $\mathbb{V}_1$ и $\mathbb{V}_2$ – конечномерные векторные пространства с заданными базисами $\boldsymbol{e}_\alpha$, $\alpha = 1, \ldots, \mathsf{dim}\, \mathbb{V}_1$ и $\boldsymbol{e}_i$, $i = 1, \ldots, \mathsf{dim}\, \mathbb{V}_2$. Тогда их элементы имеют вид $\boldsymbol{x}_1 = x_1^\alpha \boldsymbol{e}_\alpha$ и $\boldsymbol{x}_2 = x_2^i \boldsymbol{e}_i$. Из билинейности (2.28) вытекает, что тензорное произведение произвольных векторов имеет вид

$$\boldsymbol{x}_1 \otimes \boldsymbol{x}_2 = x^{\alpha i} \boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i,$$

где $x^{\alpha i} = x_1^\alpha x_2^i$. Это означает, что векторы $\{\boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i\}$ образуют базис тензорного пространства $\mathbb{V}_1 \otimes \mathbb{V}_2$, и компонентами тензорного произведения векторов $\boldsymbol{x}_1 \otimes \boldsymbol{x}_2$ относительно этого базиса являются простые произведения компонент сомножителей $\{x_1^\alpha x_2^i\}$. Отсюда следует, что размерность тензорного произведения конечна и равна произведению размерностей векторных пространств:

$$\mathsf{dim}\, (\mathbb{V}_1 \otimes \mathbb{V}_2) = \mathsf{dim}\, \mathbb{V}_1 \times \mathsf{dim}\, \mathbb{V}_2.$$

Заметим, что тензорное произведение различных пар векторов может совпадать.

Пример 2.3.41.

$$\boldsymbol{x}_1 \otimes \boldsymbol{x}_2 = (a\boldsymbol{x}_1) \otimes \left(\frac{\boldsymbol{x}_2}{a}\right), \qquad \forall a \neq 0.$$ □

Определение. Элементы тензорного произведения векторных пространств $\mathbb{V}_1 \otimes \mathbb{V}_2$ называются *тензорами*, а элементы вида $\boldsymbol{x}_1 \otimes \boldsymbol{x}_2$ – *разложимыми тензорами.* □

Поскольку элементы вида $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ составляют базис $\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$, то разложимые тензоры $\boldsymbol{x}_1 \otimes \boldsymbol{x}_2$ порождают все тензорное произведение $\mathbb{V}_1 \otimes \mathbb{V}_2$. Однако они не являются базисом, т.к. между ними существует много линейных зависимостей.

Определение. Каждой паре элементов $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathbb{V}_1 \times \mathbb{V}_2 \subset \mathbb{F}(\mathbb{V}_1, \mathbb{V}_2)$ мы поставили в соответствие их тензорное произведение $\boldsymbol{x}_1 \otimes \boldsymbol{x}_2$, которое определяется естественной проекцией

$$\mathbb{F}(\mathbb{V}_1, \mathbb{V}_2) \to \mathbb{V}_1 \otimes \mathbb{V}_2.$$

Это отображение называется *каноническим билинейным отображением* векторных пространств и обозначается следующим образом

$$\mathbb{V}_1 \times \mathbb{V}_2 \ni (\boldsymbol{x}_1, \boldsymbol{x}_2) \mapsto \varphi(\boldsymbol{x}_1, \boldsymbol{x}_2) := \boldsymbol{x}_1 \otimes \boldsymbol{x}_2 \in \mathbb{V}_1 \otimes \mathbb{V}_2. \qquad \square$$

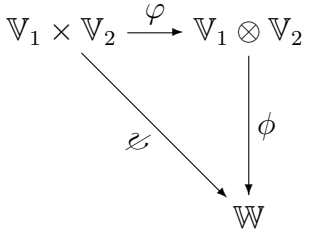
Тензорное произведение обладает следующим свойством универсальности. Пусть $\mathbb{W}$ – некоторое векторное пространство и ψ – билинейное (2.28) отображение векторных пространств $\mathbb{V}_1$ и $\mathbb{V}_2$:

$$\psi : \quad \mathbb{V}_1 \times \mathbb{V}_2 \to \mathbb{W}.$$

Тогда для любых $\mathbb{W}$ и ψ существует единственное линейное отображение

$$\phi : \quad \mathbb{V}_1 \otimes \mathbb{V}_2 \to \mathbb{W}$$

такое, что $\psi = \phi \circ \varphi$, где φ – каноническое билинейное отображение. Это означает коммутативность диаграммы



Последовательно умножая векторные пространства можно получить тензорное произведение произвольного конечного числа сомножителей. При этом выполняется закон ассоциативности $(\mathbb{V}_1 \otimes \mathbb{V}_2) \otimes \mathbb{V}_3 = \mathbb{V}_1 \otimes (\mathbb{V}_2 \otimes \mathbb{V}_3)$ и можно просто писать $\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \mathbb{V}_3$.

Два вектора из одного векторного пространства $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$ можно рассматривать как элементы двух одинаковых векторных пространств, $\mathbb{V}_1 = \mathbb{V}_2 = \mathbb{V}$, и построить их тензорное произведение $\boldsymbol{x} \otimes \boldsymbol{y}$. Оно, как было отмечено выше, некоммутативно,

$$\boldsymbol{x} \otimes \boldsymbol{y} \neq \boldsymbol{y} \otimes \boldsymbol{x},$$

поскольку в тензорном произведении важен порядок сомножителей.

Определение. Рассмотрим векторное пространство $\mathbb{V}$ и его тензорное произведение на себя

$$\mathbb{V}^r = \underbrace{\mathbb{V} \otimes \ldots \otimes \mathbb{V}}_{r}.$$

Элемент $\boldsymbol{x} \in \mathbb{V}^r$ называется *тензором* ранга r. Прямая сумма всех тензорных пространств

$$\otimes \mathbb{V} = \bigoplus_{r=0}^{\infty} \mathbb{V}^r, \tag{2.29}$$

где $\mathbb{V}^0$ – поле вещественных чисел, с операциями сложения, умножения на числа и тензорным умножением называется *тензорной алгеброй.* □

В тензорной алгебре определены три операции: сложение, умножение на числа (элементы из $\mathbb{V}^0$), а также тензорное произведение. При этом сложение двух тензоров одного ранга понимается, как сложение в соответствующем векторном пространстве $\mathbb{V}^r$. В то же время для тензоров разного ранга используется прямая сумма.

Тензорная алгебра некоммутативна и ассоциативна. Если $\boldsymbol{x} \in \mathbb{V}^r$, то мы пишем $\mathsf{deg}\, \boldsymbol{x} = r$. Тензорная алгебра (2.29) бесконечномерна, т.к. ранг тензоров неограничен. По построению она обладает естественной $\mathbb{Z}$-градуировкой. Гомоморфизм $\mathbb{Z} \to \mathbb{Z}_2$, делящий все целые числа на четные и нечетные, индуцирует $\mathbb{Z}_2$-градуировку тензорной алгебры.

Напомним общее определение алгебры.

Определение. *Ассоциативной алгеброй* над полем вещественных чисел называется кольцо $\mathbb{A}$, которое является векторным пространством над полем вещественных чисел $\mathbb{R}$ и удовлетворяет условию

$$a(fg) = (af)g = f(ag), \qquad \forall f, g \in \mathbb{A}, \quad \forall a \in \mathbb{R}.$$

Если в определении кольца исключить условие ассоциативности, то получим общее определение алгебры. Алгебра называется *коммутативной*, если $fg = gf$. *Размерностью* алгебры называется ее размерность как векторного пространства. □

Таким образом в алгебре определено три операции: сложение, умножение и умножение на числа (скаляры).

Отметим, что коммутативная алгебра может не быть ассоциативной.

Пример 2.3.42. Рассмотрим множество квадратных $n \times n$-матриц, элементами которых являются вещественные $\mathbb{R}$ или комплексные числа $\mathbb{C}$. В этих множествах определены обычные операции сложения и умножения матриц. Вместе с умножением матриц на скаляры, в роли которых выступают соответственно вещественные или комплексные числа, мы получаем ассоциативные некоммутативные (при $n \geq 2$) алгебры над полем $\mathbb{R}$ или $\mathbb{C}$. Эти алгебры обозначаются соответственно $\mathfrak{gl}(n, \mathbb{R})$ и $\mathfrak{gl}(n, \mathbb{C})$. Отметим, что умножение матриц не является групповой операцией в алгебре, т.к. среди матриц есть вырожденные (с нулевым определителем), которые не имеют обратных. Размерность алгебры $\mathfrak{gl}(n, \mathbb{R})$ равна числу независимых вещественных параметров, однозначно определяющих соответствующие матрицы,

$$\mathsf{dim}\, \mathfrak{gl}(n, \mathbb{R}) = n^2.$$

Комплексная размерность алгебры $\mathfrak{gl}(n, \mathbb{C})$ также равна n^2. Алгебру комплексных матриц $\mathfrak{gl}(n, \mathbb{C})$ можно также рассматривать над полем вещественных чисел. В этом случае ее вещественная размерность равна $2n^2$. Подмножества обратимых матриц в алгебрах $\mathfrak{gl}(n, \mathbb{R})$ и $\mathfrak{gl}(n, \mathbb{C})$ представляют собой общие линейные группы $\mathbb{GL}(n, \mathbb{R})$, $\mathbb{GL}(n, \mathbb{C})$. В них сохраняется только одна бинарная операция – умножение матриц. □

Для любых двух элементов f, g ассоциативной алгебры $\mathbb{A}$ можно ввести их *коммутатор*

$$[f, g] = fg - gf.$$

Ассоциативность необходима для того, чтобы выполнялись *тождества Якоби*

$$\Big[f, [g, h]\Big] + \Big[g, [h, f]\Big] + \Big[h, [f, g]\Big] = 0, \qquad f, g, h \in \mathbb{A},$$

которые проверяются прямой проверкой. Это означает, что на любой ассоциативной алгебре можно ввести структуру алгебры Ли (см. раздел 3.6.7). Обратное утверждение неверно. Не всякую алгебру Ли можно рассматривать, как ассоциативную алгебру.

Предложение 2.3.5. *На алгебре Ли векторных полей нельзя определить структуру ассоциативной алгебры.*

Напомним определение идеала.

Определение. Подмножество $\mathbb{I}$ кольца $\mathbb{R}$ называется *идеалом*, если выполнены два условия:

1) $\mathbb{I}$ есть подгруппа $\mathbb{R}$ по отношению к сложению;
2) $\mathbb{I}$ содержит все произведения ab (*левый идеал*), или все произведения ba (*правый идеал*), или все произведения ab и ba (*двусторонний идеал*), где a – любой элемент кольца $\mathbb{R}$ и b – любой элемент идеала $\mathbb{I}$. □

Определение. Алгебра $\mathbb{A}$ называется *фильтрованной*, если в ней выделены подпространства $\mathbb{A}_k$, индексированные элементами линейно упорядоченной группы $\mathbb{G}$ таким образом, что $\mathbb{A}_k \subseteq \mathbb{A}_l$ при $k < l$ и $\mathbb{A}_k\mathbb{A}_l \subseteq \mathbb{A}_{k+l}$.

С каждой фильтрацией данной алгебры ассоциируется *градуированная алгебра*

$$\mathsf{gr}\,\mathbb{A} = \oplus_k \bar{\mathbb{A}}_k,$$

где

$$\bar{\mathbb{A}}_k := \frac{\mathbb{A}_k}{\sum_{l<k} \mathbb{A}_l},$$

а произведение элементов $\bar{f} \in \bar{\mathbb{A}}_k$ и $\bar{g} \in \bar{\mathbb{A}}_l$ определяется по формуле $\bar{f}\bar{g} := \overline{fg}$, где f и g – представители смежных классов $\bar{f}$ и $\bar{g}$, а $\overline{fg} \in \bar{\mathbb{A}}_{k+l}$ – смежный класс, порожденный элементом $fg \in \mathbb{A}_{k+l}$. □

Замечание. Чаще всего в качестве линейно упорядоченной группы выступает группа целых чисел по сложению $\mathbb{Z}$. В этом случае $\bar{\mathbb{A}}_k = \mathbb{A}_k/\mathbb{A}_{k-1}$. □

Если в алгебре $\mathbb{A}$ выполняется какое либо полилинейное тождество (например, коммутативность, ассоциативность или тождество Якоби), то в градуированной алгебре $\mathsf{gr}\,\mathbb{A}$ также выполняется это тождество.

Пример 2.3.43. Рассмотрим алгебру Клиффорда $\mathbb{CL}(\mathbb{V}, Q)$ с образующими e^a, $a = 1, \ldots, n$ (см. раздел **??**). Обозначим через $\mathbb{A}_k$, $k \in \mathbb{Z}$, множество элементов алгебры Клиффорда, представимых в виде многочленов степени $\leq k$ от образующих. Множества $\mathbb{A}_k$ при $k < 0$ являются пустыми. Тогда подмножества $\mathbb{A}_k \subset \mathbb{CL}(\mathbb{V}, Q)$ задают фильтрацию алгебры Клиффорда. Ассоциированной градуированной алгеброй будет внешняя алгебра $\Lambda_n(\mathbb{V})$. □

Роль двусторонних идеалов в алгебрах аналогична роли нормальных подгрупп в теории групп. В частности, можно проверить, что фактор пространство $\mathbb{A}/\mathbb{I}$ является алгеброй.

Пример 2.3.44. ??? □

Обертывающая алгебра ???

Нормированные векторные пространства

Векторное евклидово пространство $\mathbb{R}^n$ можно снабдить естественной топологической метрикой. Определим расстояние между двумя векторами $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ с началом в начале системы координат по следующей формуле

$$l(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x^1 - y^1)^2 + \ldots + (x^n - y^n)^2}, \tag{2.30}$$

что совпадает с евклидовым расстоянием (2.6) между точками в $\mathbb{R}^n$, соответствующими концам векторов. При этом все свойства метрики, очевидно, выполнены. Таким образом векторное пространство $\mathbb{R}^n$ становится метрическим пространством, и, следовательно, на нем можно определить топологию. Можно доказать, что любая топология в евклидовом пространстве, совместимая со структурой векторного пространства, является естественной (с. 55, **??**).

Метрика (2.30) обладает двумя важными свойствами:

1) $l(\boldsymbol{x}, \boldsymbol{y}) = l(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + \boldsymbol{z})$ – инвариантность относительно сдвига на вектор $\boldsymbol{z} \in \mathbb{R}^n$;
2) $l(a\boldsymbol{x}, a\boldsymbol{y}) = |a| l(\boldsymbol{x}, \boldsymbol{y})$, – умножение на скаляр увеличивает расстояние в $|a|$ раз.

Из свойств метрики (2.30) вытекают следующие свойства нормы:

1) $\|\mathbf{0}\| = 0, \quad \|\boldsymbol{x}\| > 0, \quad \forall \boldsymbol{x} \neq \mathbf{0}$ – положительная определенность;
2) $\|\boldsymbol{x}\| = \|-\boldsymbol{x}\|$ – четность;
3) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ – неравенство треугольника;
4) $\|a\boldsymbol{x}\| = |a| \|\boldsymbol{x}\|$ – умножение на скаляры.

Свойство 2) является следствием 4) при $a = -1$. Оно выделено, потому что является также следствием инвариантности метрики относительно сдвига и симметрии относительно перестановки аргументов.

Определение. Векторное пространство $\mathbb{V}$ называется *нормированным*, если оно снабжено нормой,

$$\|\cdot\| : \quad \mathbb{V} \to \mathbb{R}, \tag{2.31}$$

удовлетворяющей условиям 1)– 4). При этом число $\|\boldsymbol{x}\|$ называется *нормой* или *длиной* вектора $\boldsymbol{x} \in \mathbb{V}$. Полное (см. раздел 2.3.1) нормированное векторное пространство называется *банаховым*. □

Если норма порождена метрикой, то $\|\boldsymbol{x}\| = l(\boldsymbol{x}, \mathbf{0})$.

Пример 2.3.45. Евклидово пространство $\mathbb{R}^n$, рассматриваемое как векторное пространство с естественной нормой (**??**), является банаховым. □

Предложение 2.3.6. *Каждое конечномерное нормированное векторное пространство является полным.*

Доказательство. При фиксированном базисе любое векторное пространство $\mathbb{V}$ размерности $\dim \mathbb{V} = n$ изоморфно евклидову пространству $\mathbb{R}^n$, которое полно. □

Пример 2.3.46. В евклидовом пространстве наиболее часто используются следующие нормы:

$$\begin{aligned} \|\boldsymbol{x}\|_1 &= |x^1| + |x^2| + \ldots + |x^n|, && L_1 - \text{норма}, \\ \|\boldsymbol{x}\|_2 &= \sqrt{|x^1|^2 + |x^2|^2 + \ldots + |x^n|^2}, && L_2 - \text{норма}, \\ \|\boldsymbol{x}\|_p &= \left(\sum_{\alpha=1}^{n} |x^\alpha|^p \right)^{1/p}, \quad p \geq 1 && L_p - \text{норма}, \\ \|\boldsymbol{x}\|_\infty &= \max\{|x^1|, |x^2|, \ldots, |x^n|\}, && L_\infty - \text{норма}. \end{aligned} \tag{2.32}$$

Все выражения, как нетрудно проверить, удовлетворяют свойствам 1)-4) норм. L_1-норма не порождается никаким скалярным произведением. Обычная длина вектора в евклидовом пространстве (**??**) есть ни что иное, как L_2-норма, которая порождается евклидовым скалярным произведением. L_p-норма называется также нормой *Гёльдера* с показателем $p \geq 1$. Она порождается скалярным произведением

$$l(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{\alpha=1}^{n} |x^\alpha - y^\alpha|^p \right)^{1/p}.$$

При $0 < p < 1$ выражение $\|\boldsymbol{x}\|_p$ удовлетворяет всем свойствам нормы кроме неравенства треугольника. L_∞-норма является пределом L_p-нормы. А именно,

$$\|\boldsymbol{x}\|_\infty = \lim_{p \to \infty} \|\boldsymbol{x}\|_p, \qquad \forall \boldsymbol{x} \in \mathbb{R}^n.$$

Все приведенные выше нормы прямо обобщаются на комплексное пространство $\mathbb{C}^n$ □

Рассмотренный пример показывает, что в одном и том же векторном пространстве можно задавать различные нормы и скалярные произведения. Выбор той или иной нормы зависит от конкретной задачи.

Пример 2.3.47. Понятие нормы можно ввести и на бесконечномерном векторном пространстве. На множестве $\mathcal{C}[a, b]$ всех непрерывных функций $f(x)$ на отрезке $[a, b]$, которое является бесконечномерным векторным пространством, обычно рассматривают следующие нормы:

$$\begin{aligned} \|f\|_1 &= \int_a^b dx |f(x)|, && L_1 - \text{норма}, \\ \|f\|_2 &= \left[\int_a^b dx |f(x)|^2 \right]^{1/2}, && L_2 - \text{норма}, \\ \|f\|_p &= \left[\int_a^b dx |f(x)|^p \right]^{1/p}, \quad p \geq 1 && L_p - \text{норма}, \\ \|f\|_\infty &= \max\{|f(x)| : \ x \in [a, b]\}, && L_\infty - \text{норма}. \end{aligned} \tag{2.33}$$

При этом, как и в предыдущем примере, выполнено равенство

$$\|f\|_\infty = \lim_{p \to \infty} \|f\|_p, \qquad \forall f \in \mathcal{C}[a, b].$$ □

Определение. Метрику (2.30) (а также любую другую метрику) можно использовать для определения базы топологии векторного пространства. В результате будет построено *топологическое векторное пространство*. Топологическое векторное пространство $\mathbb{V}$ называется *метризуемым*, если его топология может быть индуцирована трансляционно инвариантной метрикой $l(\boldsymbol{x},\boldsymbol{y})=l(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z})$, $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{V}$. При этом выполнения других свойств не требуется. В общем случае топологическое векторное пространство, которое является метризуемым и полным, называется *пространством Фреше*. □

Пример 2.3.48. Евклидово пространство $\mathbb{R}^n$, рассматриваемое как векторное пространство с метрикой (2.30), является пространством Фреше. □

Определение. Пусть $f:\ \mathbb{V}\to\mathbb{W}$ – отображение, возможно, нелинейное, двух нормированных векторных пространств, и пусть $\boldsymbol{x}\in\mathbb{V}$ – некоторая точка. Тогда, если отображение f вблизи точки $\boldsymbol{x}$ можно представить в виде

$$f(\boldsymbol{x}+\boldsymbol{h})=f(\boldsymbol{x})+f'_{\boldsymbol{x}}(\boldsymbol{h})+\mathsf{o}(\boldsymbol{h}),$$

где $f'_{\boldsymbol{x}}$ – линейное отображение $\mathbb{V}\to\mathbb{W}$ и $\mathsf{o}(\boldsymbol{h})\in\mathbb{W}$ – некоторый вектор, такой, что

$$\lim_{\|\boldsymbol{h}\|\to 0}\frac{\|\mathsf{o}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|}=0, \tag{2.34}$$

то отображение $f'_{\boldsymbol{x}}:\ \mathbb{V}\to\mathbb{W}$ называется *производной по Фреше* отображения f в точке $\boldsymbol{x}$. □

В частности, если $\mathbb{W}=\mathbb{R}$ – вещественная прямая, то мы имеем производную по Фреше от функционала на нормированном векторном пространстве.

Нетрудно показать, что если производная по Фреше существует, то она единственна.

Пример 2.3.49. Пусть в векторных пространствах $\mathbb{V}$ и $\mathbb{W}$ заданы базисы $\boldsymbol{e}_\alpha$, $\alpha=1,\ldots,n$, и $\boldsymbol{e}_i$, $i=1,\ldots,m$, соответственно. Тогда отображение f задается n функциями $f^i(\boldsymbol{x})$ от n переменных, которые в общем случае могут быть нелинейными. Если функции f^i дифференцируемы по всем аргументам, то $n\times m$-матрица Якоби, составленная из частных производных $\{\partial_\alpha f^i|_{\boldsymbol{x}}\}$ является производной по Фреше отображения f в точке $\boldsymbol{x}\in\mathbb{V}$. □

Для производной по Фреше справедлива теорема о дифференцировании сложной функции.

Если отображение f двух банаховых пространств непрерывно дифференцируемо по Фреше в некоторой окрестности точки $\boldsymbol{x}$ и в этой точке производная Фреше $f'_{\boldsymbol{x}}$ является гомеоморфизмом, то в окрестности данной точки существует обратное отображение, которое также является гомеоморфизмом.

Наряду с производной Фреше отображений двух линейных пространств используется также производная Гато, для которой необходимо только наличие топологии. Пусть $f:\ \mathbb{V}\to\mathbb{W}$ – отображение двух топологических векторных пространств, и пусть $\boldsymbol{x}\in\mathbb{V}$ некоторая точка.

Определение. Если отображение f двух линейных пространств вблизи точки $\boldsymbol{x}$ можно представить в виде

$$f(\boldsymbol{x}+\boldsymbol{h})=f(\boldsymbol{x})+f'_{G,\boldsymbol{x}}(\boldsymbol{h})+\mathsf{o}(\boldsymbol{h}),$$

где $f'_{G,\boldsymbol{x}}(\boldsymbol{h})$ – линейное отображение $\mathbb{V} \to \mathbb{W}$ и

$$\lim_{t\to 0} \frac{\mathsf{o}(t\boldsymbol{h})}{t} = 0, \qquad t \in \mathbb{R}, \tag{2.35}$$

где сходимость определяется топологией пространства $\mathbb{W}$, то линейное отображение $f'_{G,\boldsymbol{x}}: \ \mathbb{V} \to \mathbb{W}$ называется *производной по Гато* отображения f в точке $\boldsymbol{x}$. □

В частности, если $\mathbb{W} = \mathbb{R}$ – вещественная прямая, то мы имеем производную по Гато от функционала на топологическом векторном пространстве.

Отличие производной Гато от производной Фреше сводится только к определению сходимости остаточных членов (2.34) и (2.35).

Поскольку каждое нормированное векторное пространство является топологическим с естественной топологией, то производная по Фреше является одновременно и производной по Гато. Обратное утверждение в общем случае неверно, т.к. наличие топологии в векторном пространстве совсем не означает наличие нормы. По этой причине производную Фреше называют также *сильной производной*, а производную Гато – *слабой*.

Для производной Гато теорема о дифференцируемости сложной функции, вообще говоря, не верна.

Скалярное произведение в векторном пространстве

Определение. Скалярным произведением $(\boldsymbol{x}, \boldsymbol{y})$ двух элементов векторного пространства $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$ называется билинейное отображение $\mathbb{V}\times\mathbb{V} \to \mathbb{R}$, которое обладает следующими свойствами:

1) $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$ – симметричность,
2) $(\boldsymbol{x}, \boldsymbol{x}) > 0, \ \forall \boldsymbol{x} \neq \boldsymbol{0}; \quad (\boldsymbol{0}, \boldsymbol{0}) = 0$ – положительная определенность.

Если скалярное произведение двух векторов равно нулю $(\boldsymbol{x}, \boldsymbol{y}) = 0$, то они называются *ортогональными*. □

В компонентах скалярное произведение задается положительно определенной квадратичной формой

$$(\boldsymbol{x}, \boldsymbol{y}) = x^\alpha y^\beta g_{\alpha\beta}, \tag{2.36}$$

где $g_{\alpha\beta}$ – произвольная симметричная положительно определенная матрица.

Пример 2.3.50. В евклидовом пространстве $\mathbb{R}^n$ естественное скалярное произведение векторов задается в декартовых координатах с помощью евклидовой метрики (2.8) Это скалярное произведение симметрично и положительно определено. При этом норму (длину) вектора можно записать в виде $\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$. □

Отметим, что скалярное произведение различных векторов $(\boldsymbol{x}, \boldsymbol{y})$ может быть отрицательно. Нулевой вектор ортогонален самому себе и всем другим векторам.

Предложение 2.3.7. *Пусть $\mathbb{V}$ – произвольное векторное пространство с положительно определенным скалярным произведением (2.36). Тогда на $\mathbb{V}$ можно задать норму или* длину вектора

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}. \tag{2.37}$$

Доказательство. Из линейности и положительной определенности скалярного произведения немедленно следуют свойства 1), 2) и 4) в определении нормы. Для доказательства неравенства треугольника воспользуемся неравенством Шварца (2.38), которое будет доказано чуть позже:

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|^2 = (\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}) &\le \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2|(\boldsymbol{x},\boldsymbol{y})| \\ &\le \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| \\ &= \Big(\|\boldsymbol{x}\| + \|\boldsymbol{y}\|\Big)^2. \end{aligned}$$

□

Обсудим некоторые свойства векторных пространств с положительно определенной метрикой.

Предложение 2.3.8 (**Правило параллелограмма**). *Пусть* $\mathbb{V}$ *– векторное пространство с положительно определенным скалярным произведением. Тогда справедлива формула*

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2,$$

где $\boldsymbol{x},\boldsymbol{y} \in \mathbb{V}$ *– произвольные векторы.*

Доказательство. Прямая проверка. □

Предложение 2.3.9 (**Теорема Пифагора**). *Пусть* $\mathbb{V}$ *– векторное пространство с положительно определенным скалярным произведением. Тогда справедлива формула*

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2,$$

где $\boldsymbol{x},\boldsymbol{y} \in \mathbb{V}$ *– произвольные ортогональные векторы,* $(\boldsymbol{x},\boldsymbol{y}) = 0$.

Доказательство. Прямая проверка. □

Допустим, что в векторном пространстве $\mathbb{V}$ задан произвольный набор взаимно ортогональных векторов $\boldsymbol{x}_i$, $i = 1,\dots,\mathrm{N} \le n$:

$$(\boldsymbol{x}_i,\boldsymbol{x}_j) = 0, \quad i \ne j.$$

Тогда из теоремы Пифагора по индукции следует равенство

$$\left\|\sum_{i=1}^{\mathrm{N}} \boldsymbol{x}_i\right\|^2 = \sum_{i=1}^{\mathrm{N}} \|\boldsymbol{x}_i\|^2.$$

В произвольном векторном пространстве $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, со скалярным произведением можно выбрать *ортонормальный базис* $\boldsymbol{e}_\alpha$, $\alpha = 1,\dots,n$, состоящий из взаимно ортогональных векторов единичной длины:

$$(\boldsymbol{e}_\alpha,\boldsymbol{e}_\beta) = \delta_{\alpha\beta},$$

где $\delta_{\alpha\beta}$ – евклидова метрика (2.8). Произвольный вектор $\boldsymbol{x} \in \mathbb{V}$ можно разложить по ортонормальному базису

$$\boldsymbol{x} = x^\alpha \boldsymbol{e}_\alpha,$$

где $x^\alpha := \delta^{\alpha\beta}(\boldsymbol{x},\boldsymbol{e}_\beta)$ – компоненты данного вектора в ортонормальном базисе.

Предложение 2.3.10 (Неравенство Бесселя). *Пусть $\boldsymbol{x}_i \in \mathbb{V}$, $i = 1, \ldots, \mathrm{N} \leq n$, – произвольный набор ортонормальных векторов. Тогда справедливо неравенство*

$$\sum_{i=1}^{\mathrm{N}} |(\boldsymbol{x}, \boldsymbol{x}_i)|^2 \leq \|\boldsymbol{x}\|^2,$$

где $\boldsymbol{x} \in \mathbb{V}$ – произвольный вектор.

Доказательство. Из неравенства

$$0 \leq \left\| \boldsymbol{x} - \sum_{i=1}^{\mathrm{N}} (\boldsymbol{x}, \boldsymbol{x}_\alpha)\boldsymbol{x}_\alpha \right\|^2 = \|\boldsymbol{x}\|^2 - 2\sum_{i=1}^{\mathrm{N}} (\boldsymbol{x}, \boldsymbol{x}_i)^2 + \sum_{i=1}^{\mathrm{N}}\sum_{j=1}^{\mathrm{N}} (\boldsymbol{x}, \boldsymbol{x}_i)(\boldsymbol{x}, \boldsymbol{x}_j)(\boldsymbol{x}_i, \boldsymbol{x}_j).$$

вытекает неравенство Бесселя, т.к. векторы $\boldsymbol{x}_i$ ортонормальны, $(\boldsymbol{x}_i, \boldsymbol{x}_j) = \delta_{ij}$. □

Знак равенства в неравенстве Бесселя достигается тогда и только тогда, когда набор векторов $\boldsymbol{x}_i$ образует ортонормальный базис в $\mathbb{V}$, т.е. $\mathrm{N} = n$.

Предложение 2.3.11. *Пусть $\boldsymbol{e}_\alpha \in \mathbb{V}$, $\alpha = 1, \ldots, n$, – ортонормальный базис векторного пространства. Тогда справедливо равенство*

$$\sum_{\alpha=1}^{n} |(\boldsymbol{x}, \boldsymbol{e}_\alpha)|^2 = \|\boldsymbol{x}\|^2,$$

где $\boldsymbol{x} \in \mathbb{V}$ – произвольный вектор. То есть квадрат длины произвольного вектора равен сумме квадратов компонент относительно ортонормального базиса.

Доказательство. В доказательстве неравенства Бесселя надо заменить знак неравенства на равенство. □

Предложение 2.3.12 (Неравенство Коши–Буняковского). *Пусть $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$ – произвольные векторы. Тогда справедливо неравенство*

$$|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|. \tag{2.38}$$

Неравенство Коши–Буняковского называют также *неравенством Шварца.*

Доказательство. Если $\boldsymbol{y} = 0$, то неравенство Шварца (2.38) выполнено. Допустим, что $\boldsymbol{y} \neq 0$. В этом случае можно построить единичный вектор $\boldsymbol{y}_0 := \boldsymbol{y}/\|\boldsymbol{y}\|$, $\|\boldsymbol{y}_0\| = 1$. Тогда из неравенства Бесселя следует неравенство

$$|(\boldsymbol{x}, \boldsymbol{y}_0)| \leq \|\boldsymbol{x}\|.$$

Поскольку

$$|(\boldsymbol{x}, \boldsymbol{y}_0)| = \frac{(\boldsymbol{x}, \boldsymbol{y})}{\|\boldsymbol{y}\|},$$

то отсюда вытекает неравенство Шварца. □

В моделях математической физики, например, при рассмотрении лоренцевых многообразий, под скалярным произведением понимают произвольное симметричное билинейное отображение, отбрасывая требование положительной определенности. В этом случае два вектора также называются ортогональными, если их скалярное произведение равно нулю. При этом отличный от нуля вектор $\boldsymbol{x} \neq 0$ может оказаться

ортогональным самому себе, если $(\boldsymbol{x}, \boldsymbol{x}) = 0$. Такой вектор называется *нулевым*. Если метрика имеет лоренцеву сигнатуру $\mathsf{sign}\, g_{\alpha\beta} = (+ - \ldots -)$, то нулевой вектор называется также *светоподобным*.

Из симметрии матрицы $g_{\alpha\beta}$ следует, что все ее собственные числа вещественны. Для того, чтобы матрица была положительно определена необходимо и достаточно, чтобы все ее собственные числа были вещественны и положительны.

Если в векторном пространстве $\mathbb{V}$ задано скалярное произведение, то в нем всегда можно выбрать базис $\boldsymbol{e}_\alpha$, состоящий из взаимно ортогональных векторов. Эти векторы будут собственными векторами матрицы $g_{\alpha\beta}$:

$$\sum_{\beta=1}^{n} g_{\alpha\beta}\boldsymbol{e}_\beta = \lambda_\alpha \boldsymbol{e}_\alpha,$$

где проводится суммирование по индексу β, но отсутствует суммирование по α, и $\lambda_\alpha > 0$ – собственные значения матрицы $g_{\alpha\beta}$. В общем случае часть собственных значений может совпадать. Для метрики евклидова пространства $\mathbb{R}^n$ в декартовых координатах все собственные значения совпадают и равны единице.

Рассмотрим два векторных пространства $\mathbb{V}_1$ и $\mathbb{V}_2$ с элементами $\boldsymbol{x} = x^\alpha \boldsymbol{e}_\alpha \in \mathbb{V}_1$, $\alpha = 1, \ldots, m$ и $\boldsymbol{y} = y^i \boldsymbol{e}_i \in \mathbb{V}_2$, $i = 1, \ldots, n$. Выберем в каждом пространстве базис, состоящий из ортогональных векторов. Пусть в них заданы скалярные произведения матрицами $g_{\alpha\beta}$ и g_{ij} соответственно. Тогда в тензорном произведении $\mathbb{V}_1 \otimes \mathbb{V}_2$ можно определить скалярное произведение. Действительно, пусть задано два вектора

$$X = X^{\alpha i}\boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i, \qquad Y = Y^{\alpha a}\boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i.$$

Тогда билинейное отображение

$$(X, Y) = X^{\alpha i} Y^{\beta j} g_{\alpha\beta} g_{ij}$$

симметрично и положительно определено. Последнее утверждение следует из того, что векторы базиса $\boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i$ являются собственными векторами матрицы $g_{(\alpha i)(\beta j)} = g_{\alpha\beta} g_{ij}$, определяющей скалярное произведение,

$$\sum_{\beta,j} g_{\alpha\beta} g_{ij} \boldsymbol{e}_\beta \otimes \boldsymbol{e}_j = \lambda_\alpha \lambda_i \boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i,$$

с положительными собственными числами $\lambda_\alpha \lambda_i > 0$.

Как уже отмечалось, векторные пространства можно рассматривать над полем комплексных чисел. В этом случае говорят о комплексном векторном пространстве $\mathbb{H}$.

Определение. *Скалярным произведением* в комплексном векторном пространстве $\mathbb{H}$ называется отображение

$$\mathbb{H} \times \mathbb{H} \ni x_1, x_2 \mapsto (x_1, x_2) \in \mathbb{C},$$

которое линейно по второму аргументу,

$$(x_1, a x_2 + b x_3) = a(x_1, x_2) + b(x_1, x_3),$$

и удовлетворяет равенству

$$(x_1, x_2)^\dagger = (x_2, x_1),$$

где символ † обозначает комплексное сопряжение. Если скалярное произведение положительно определено, т.е. $(x, x) \geq 0$ для всех $x \in \mathbb{H}$, то оно задает норму в $\mathbb{H}$:

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}.$$

Эта норма называется строго положительной, если из условия $\|\boldsymbol{x}\| = 0$ следует $\boldsymbol{x} = 0$. Полное комплексное векторное пространство со строго положительной нормой называется *гильбертовым*. □

Гильбертовы пространства лежат в основе квантовой механики, где каждое состояние физической системы отождествляется с вектором соответствующего гильбертова пространства. Гильбертовы пространства в квантовой механике обычно бесконечномерны.

Пример 2.3.51. В нерелятивистской квантовой механике в качестве гильбертова пространства, как правило, рассматривается пространство $\mathbb{H} = \mathcal{L}_2(\mathbb{R}^n)$ достаточно гладких квадратично интегрируемых функций в евклидовом пространстве

$$\phi : \quad \mathbb{R}^n \to \mathbb{C}.$$

Скалярное произведение в этом пространстве определяется интегралом

$$\mathbb{H} \times \mathbb{H} \ni \phi_1, \phi_2 \mapsto \int_{\mathbb{R}^n} d^n x \phi_1^\dagger \phi_2.$$

Можно доказать, что это пространство является полным. □

Сопряженные пространства

Определение. *Линейным функционалом* или *линейной формой* $\boldsymbol{f}$ на векторном пространстве $\mathbb{V}$ называется линейное отображение

$$\boldsymbol{f} : \quad \mathbb{V} \to \mathbb{R},$$

обладающее свойствами:

$$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{y}), \qquad \boldsymbol{f}(a\boldsymbol{x}) = a\boldsymbol{f}(\boldsymbol{x}).$$

Множество $\mathbb{V}^*$ всех линейных функционалов на $\mathbb{V}$ снабжено естественной структурой векторного пространства над полем вещественных чисел относительно сложения и умножения на числа:

$$(\boldsymbol{f}_1 + \boldsymbol{f}_2)(\boldsymbol{x}) := \boldsymbol{f}_1(\boldsymbol{x}) + \boldsymbol{f}_2(\boldsymbol{x}),$$
$$(a\boldsymbol{f})(\boldsymbol{x}) := a\boldsymbol{f}(\boldsymbol{x}).$$

Векторное пространство $\mathbb{V}^*$ называется *сопряженным*, *дуальным* или *двойственным* пространством к $\mathbb{V}$. □

Можно показать, что сопряженное пространство имеет ту же размерность, что и векторное пространство $\mathbb{V}$. Для конечномерных векторных пространств взятие сопряжения дважды приводит к векторному пространству, которое изоморфно исходному, $\mathbb{V}^{**} = \mathbb{V}$.

Определение. Пусть $\mathbb{U} \subset \mathbb{V}$, тогда *ортогональным дополнением* или *аннулятором* множества $\mathbb{U}$ (не обязательно подпространства) называется подмножество

$$\mathbb{U}^\perp = \{\boldsymbol{f} \in \mathbb{V}^* : \quad \boldsymbol{f}(\boldsymbol{x}) = 0, \quad \forall \boldsymbol{x} \in \mathbb{U}\}. \quad \square \tag{2.39}$$

Предложение 2.3.13. *Если $\mathbb{U}$ – линейное подпространство в $\mathbb{V}$ размерности m, $1 \le m \le n$, то его ортогональное дополнение $\mathbb{U}^\perp$ является линейным подпространством в $\mathbb{V}^*$ и имеет размерность $n - m$.*

Доказательство. Выберем в подпространстве $\mathbb{U}$ m линейно независимых векторов $\boldsymbol{x}_i$, $i = 1, \ldots, m$. Тогда система линейных уравнений $\boldsymbol{f}(\boldsymbol{x}_i) = 0$ будет иметь $n - m$ линейно независимых решений, которые можно выбрать в качестве базиса в $\mathbb{U}^\perp$. $\square$

Аналогично определяется ортогональное дополнение некоторого подмножества $\mathbb{U}^* \subset \mathbb{V}^*$:

$$\mathbb{U}^{*\perp} = \{\boldsymbol{x} \in \mathbb{V} : \quad \boldsymbol{f}(\boldsymbol{x}) = 0, \quad \forall \boldsymbol{f} \in \mathbb{U}^*\}. \tag{2.40}$$

Это означает, что операцию ортогонального дополнения можно применить дважды. Тогда для конечномерных векторных пространств справедливо равенство

$$(\mathbb{U}^\perp)^\perp = \mathbb{U}.$$

Если $\mathbb{V}_1, \mathbb{V}_2 \subset \mathbb{V}$ – два подпространства векторного пространства, то верна формула

$$\begin{aligned} (\mathbb{V}_1 \cap \mathbb{V}_2)^\perp &= \mathbb{V}_1^\perp \oplus \left[\mathbb{V}_2^\perp \setminus (\mathbb{V}_1^\perp \cap \mathbb{V}_2^\perp)\right] \\ &= \left[\mathbb{V}_1^\perp \setminus (\mathbb{V}_1^\perp \cap \mathbb{V}_2^\perp)\right] \oplus \mathbb{V}_2^\perp. \end{aligned} \tag{2.41}$$

Здесь из одного из слагаемых вычтено пересечение ортогональных дополнений, чтобы не учитывать его дважды.

Определение. *Гиперплоскостью* в векторном пространстве $\mathbb{V}$, ортогональной некоторому фиксированному элементу из сопряженного пространства $\boldsymbol{f} \in \mathbb{V}^*$, называется множество точек

$$\{\boldsymbol{x} \in \mathbb{V} : \quad \boldsymbol{f}(\boldsymbol{x}) = a, \quad \boldsymbol{f} \neq 0, \ \boldsymbol{f} \in \mathbb{V}^*, \ \ a \in \mathbb{R}\}. \quad \square \tag{2.42}$$

Гиперплоскости являются подпространствами размерности $n - 1$ и определены для всех ненулевых линейных функционалов.

Пример 2.3.52. В евклидовом пространстве $\mathbb{R}^n$ скалярное произведение векторов при одном фиксированном векторе задает линейный функционал. Верно и обратное утверждение: любой линейный функционал можно задать, как скалярное произведение с некоторым вектором. Это означает, что дуальное пространство $\mathbb{R}^{n*}$ совпадает с ним самим:

$$\mathbb{R}^{n*} = \mathbb{R}^n.$$

Поэтому понятия ортогонального дополнения и гиперплоскости, ортогональной заданному вектору, в евклидовом пространстве имеют наглядный геометрический смысл.

$\square$

Если $\boldsymbol{e}_\alpha$ – базис векторного пространства $\mathbb{V}$, то множество функционалов $\boldsymbol{e}^\alpha$, определенное соотношением

$$\boldsymbol{e}^\alpha(\boldsymbol{e}_\beta) = \delta^\alpha_\beta, \tag{2.43}$$

определяет единственный *сопряженный (дуальный) базис* сопряженного пространства. При этом любой функционал взаимно однозначно представим в виде

$$\boldsymbol{f} = \boldsymbol{e}^\alpha f_\alpha, \tag{2.44}$$

с некоторыми числовыми коэффициентами f_α, а значение функционала на векторе равно

$$\boldsymbol{f}(\boldsymbol{x}) = x^\alpha f_\alpha. \tag{2.45}$$

Используя понятие линейного функционала, можно дать новое эквивалентное определение тензорного произведения векторных пространств.

Определение. Пусть $\mathbb{V}_1^*$ и $\mathbb{V}_2^*$ – векторные пространства, дуальные к $\mathbb{V}_1$ и $\mathbb{V}_2$. Тогда *тензорным произведением* векторных пространств $\mathbb{V}_1 \otimes \mathbb{V}_2$ называется множество всех билинейных отображений упорядоченной пары дуальных пространств в поле вещественных чисел

$$\mathbb{V}_1 \otimes \mathbb{V}_2 : \quad (\mathbb{V}_1^*, \mathbb{V}_2^*) \to \mathbb{R}. \quad \square \tag{2.46}$$

В компонентах отображение (2.46) записывается следующим образом. Если $\boldsymbol{f}_1 = f_{1\alpha}\boldsymbol{e}^\alpha \in \mathbb{V}_1$ и $\boldsymbol{f}_2 = f_{2i}\boldsymbol{e}^i \in \mathbb{V}_2$, то элемент тензорного произведения $x^{\alpha i}\boldsymbol{e}_\alpha \otimes \boldsymbol{e}_i \in \mathbb{V}_1 \otimes \mathbb{V}_2$ задает отображение $x^{\alpha i} f_{1\alpha} f_{2i} \in \mathbb{R}$.

Метрика в евклидовом пространстве $\mathbb{R}^n$ задает двойственное пространство $\mathbb{R}^{n*}$, элементы которого определяются компонентами $x_\alpha = x^\beta g_{\beta\alpha}$. Наличие метрики позволяет установить изоморфизм $\mathbb{R}^n$ и его дуального пространства $\mathbb{R}^{n*}$, рассматриваемых как векторные пространства.

Пусть задано векторное пространство $\mathbb{V}$ и сопряженное к нему пространство $\mathbb{V}^*$. Тензором типа (r, s) мы будем называть элемент тензорного произведения

$$\underbrace{\mathbb{V} \otimes \ldots \otimes \mathbb{V}}_{r} \otimes \underbrace{\mathbb{V}^* \otimes \ldots \otimes \mathbb{V}^*}_{s},$$

где пространства $\mathbb{V}$ и $\mathbb{V}^*$ встречаются соответственно r и s раз. Для определенности векторное пространство $\mathbb{V}$ выбрано в качестве первых r сомножителей, хотя возможен и любой другой порядок множителей. Если в пространствах $\mathbb{V}$ и $\mathbb{V}^*$ заданы базисы $\boldsymbol{e}_\alpha$ и $\boldsymbol{e}^\alpha$, то тензор типа (r, s) будет иметь r верхних и s нижних индексов

$$X = X^{\alpha_1 \ldots \alpha_r}{}_{\beta_1 \ldots \beta_s} \boldsymbol{e}_{\alpha_1} \otimes \ldots \otimes \boldsymbol{e}_{\alpha_r} \otimes \boldsymbol{e}^{\beta_1} \otimes \ldots \otimes \boldsymbol{e}^{\beta_s},$$

которые называются *контравариантными* и *ковариантными*, соответственно. Поскольку тензорное произведение некоммутативно, то порядок индексов является существенным. Так $X^{\alpha\beta} \neq X^{\beta\alpha}$ и $X^\alpha{}_\beta \neq X_\beta{}^\alpha$.

Определение. Пусть задано линейное отображение двух векторных пространств

$$\varphi : \quad \mathbb{V} \;\to\; \mathbb{W},$$

которые могут иметь разные размерности: $\mathsf{dim}\,\mathbb{V} = n$, $\mathsf{dim}\,\mathbb{W} = m$. В компонентах это отображение задается $n \times m$-матрицей $A = (A_\alpha{}^i)$

$$\varphi : \quad \mathbb{V} \ni \{x^\alpha\} \;\mapsto\; \{y^i := x^\alpha A_\alpha{}^i\} \in \mathbb{W}.$$

Пусть $\boldsymbol{f}$ и $\boldsymbol{g}$ – линейные формы на $\mathbb{V}$ и $\mathbb{W}$, соответственно. Тогда отображение φ индуцирует *возврат отображения* φ^*, действующий в обратную сторону

$$\varphi^* : \quad \mathbb{W}^* \ \to \ \mathbb{V}^*$$

по следующему правилу

$$(\boldsymbol{f} = \varphi^* \boldsymbol{g})(\boldsymbol{x}) := \boldsymbol{g}\Big(\varphi(\boldsymbol{x})\Big).$$

В компонентах возврат отображения имеет вид

$$\varphi^* : \quad \mathbb{W}^* \ni \{g_i\} \ \mapsto \ \{f_\alpha := A_\alpha{}^i g_i\} \in \mathbb{V}^*. \qquad \square$$

Замечание. Матрица $A_\alpha{}^i$ в общем случае является прямоугольной, и говорить об обратной матрице не имеет смысла. $\square$

2.3.4 $\mathbb{R}^n$ как аффинное пространство

Прямая в евклидовом пространстве обладает двумя важными и независимыми свойствами. Во-первых, отрезок, соединяющий две точки евклидова пространства, имеет наименьшую длину среди всех кривых, соединяющих эти точки (метрическое свойство.) Во-вторых, касательный вектор к прямой остается касательным при параллельном переносе вдоль прямой. В настоящем разделе мы придадим смысл последнему утверждению. Это важно, т.к. параллельный перенос векторов в дифференциальной геометрии является нетривиальным обобщением параллельного переноса векторов в аффинном пространстве.

В предыдущем разделе было установлено, что точки евклидова пространства $\mathbb{R}^n$ находятся во взаимно однозначном соответствии с множеством векторов $\mathbb{V}$ евклидова пространства, имеющих начало в начале системы координат. Компоненты этих векторов совпадают с координатами точек – концов векторов – и поэтому их можно складывать с координатами произвольных точек из $\mathbb{R}^n$, которые сами уже не рассматриваются, как элементы векторного пространства. Введение этой операции приводит к понятию аффинного пространства. Чтобы подчеркнуть разницу между точками и векторами дадим сначала общее определение.

Определение. *Аффинным пространством* над полем вещественных чисел называется тройка $(\mathbb{A}, \mathbb{V}, +)$, состоящая из множества $\mathbb{A}$, элементы которого называются точками, ассоциированного векторного пространства $\mathbb{V}$ и отображения

$$\mathbb{A} \times \mathbb{V} \ni (\mathrm{P}, \boldsymbol{x}) \mapsto \mathrm{P} + \boldsymbol{x} \in \mathbb{A}, \tag{2.47}$$

удовлетворяющего следующим аксиомам:

1) $(\mathrm{P} + \boldsymbol{x}) + \boldsymbol{y} = \mathrm{P} + (\boldsymbol{x} + \boldsymbol{y}) \quad \forall \mathrm{P} \in \mathbb{A}, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{V},$
2) $\mathrm{P} + \boldsymbol{0} = \mathrm{P}, \quad \forall \mathrm{P} \in \mathbb{A},$
3) для двух произвольных точек $\mathrm{P}, \mathrm{Q} \in \mathbb{A}$ уравнение $\mathrm{Q} = \mathrm{P} + \boldsymbol{x}$ на вектор $\boldsymbol{x}$ имеет единственное решение. $\square$

В рассматриваемом случае множество точек представляет собой множество точек евклидова пространства $\mathbb{A} = \mathbb{R}^n$. Векторное пространство $\mathbb{V}$ также совпадает с евклидовым пространством $\mathbb{V} = \mathbb{R}^n$, которое рассматривается как векторное пространство. При этом отображение (2.47) называется *сдвигом* аффинного пространства на вектор $\boldsymbol{x}$. Сдвиг на противоположный вектор $-\boldsymbol{x}$ обозначается знаком минус, $\mathrm{P} - \boldsymbol{x}$.

Замечание. Отметим, что в определении аффинного пространства не фигурирует явно структура умножения на скаляры. Вообще говоря, для точек аффинного пространства не определено понятие суммы $\mathrm{P}+\mathrm{Q}$ и умножения на числа $a\mathrm{P}$. □

Тот единственный вектор $\boldsymbol{x}\in\mathbb{V}$, для которого $\mathrm{Q}=\mathrm{P}+\boldsymbol{x}$, удобно обозначать $\mathrm{Q}-\mathrm{P}$. Эта операция вычитания

$$\mathbb{A}\times\mathbb{A}\ni(\mathrm{Q},\mathrm{P})\mapsto\mathrm{Q}-\mathrm{P}\in\mathbb{V}$$

обладает следующими свойствами:

1) $(\mathrm{R}-\mathrm{Q})+(\mathrm{Q}-\mathrm{P})=\mathrm{R}-\mathrm{P}$;

2) $\mathrm{P}-\mathrm{P}=\mathbf{0}$;

3) $(\mathrm{P}+\boldsymbol{x})-(\mathrm{Q}+\boldsymbol{y})=(\mathrm{P}-\mathrm{Q})+(\boldsymbol{x}-\boldsymbol{y})$.

Свойство 1), справедливое для трех произвольных точек аффинного пространства, называется *соотношением Шаля* и его можно записать в эквивалентной форме

$$(\mathrm{R}-\mathrm{Q})+(\mathrm{Q}-\mathrm{P})+(\mathrm{P}-\mathrm{R})=\mathbf{0}.$$

Интуитивно аффинное пространство $(\mathbb{A},\mathbb{V},+)$ можно представлять себе как векторное пространство $\mathbb{V}$ с забытым началом координат $\mathbf{0}$, в котором введена операция сдвига на вектор.

В аффинном пространстве можно ввести аффинные координаты, состоящие из фиксированной точки $\mathrm{P}_0\in\mathbb{A}$ (начала координат) и базиса $\boldsymbol{e}_\alpha$ соответствующего векторного пространства. Это является следствием третьего свойства в определении аффинного пространства. Произвольная точка $\mathrm{P}\in\mathbb{A}$ имеет единственное представление

$$\mathrm{P}=\mathrm{P}_0+x^\alpha_{\mathrm{P}}\boldsymbol{e}_\alpha, \tag{2.48}$$

где числа x^α_{P} называются координатами точки P в данной системе аффинных координат. Если предположить, что точка P_0 совпадает с началом координат в евклидовом векторном пространстве $\mathbb{R}^n$, то произвольная точка P определяется только координатами соответствующего вектора $\mathrm{P}=\{x^\alpha_{\mathrm{P}}\}$. Таким образом, мы имеем взаимно однозначное соответствие между точками аффинного пространства и векторами из $\mathbb{V}$. При этом сдвиг точки P (2.47) на вектор $\boldsymbol{x}=\{x^\alpha\}$ записывается в виде

$$x^\alpha_{\mathrm{P}}\mapsto x^\alpha_{\mathrm{P}}+x^\alpha.$$

После отождествления точек аффинного и векторного пространства можно считать, что в общем случае точка P имеет координаты

$$\mathrm{P}=\{x^\alpha_0+x^\alpha_{\mathrm{P}}\},$$

где x^α_0 – координаты точки P_0 в разложении (2.48). Поскольку количество аффинных координат точек аффинного пространства равно n, то *размерностью* аффинного пространства называется размерность ассоциированного векторного пространства.

Пусть аффинное пространство имеет конечную размерность n. Если координаты всех точек аффинного пространства преобразовать по правилу

$$x^\alpha\mapsto x'^\alpha=x^\beta A_\beta{}^\alpha+a^\alpha, \tag{2.49}$$

где $A\in\mathbb{GL}(n,\mathbb{R})$ – произвольная невырожденная матрица и $\boldsymbol{a}=a^\alpha\boldsymbol{e}_\alpha\in\mathbb{V}$ – произвольный фиксированный вектор, то получим взаимно однозначное отображение аффинного пространства на себя. Последовательное выполнение двух преобразований (2.49) с параметрами $(A,\boldsymbol{a})$ и $(B,\boldsymbol{b})$ имеет вид

$$x''^\alpha=x^\gamma A_\gamma{}^\beta B_\beta{}^\alpha+a^\beta B_\beta{}^\alpha+b^\alpha.$$

Определение. Совокупность преобразований вида (2.49) с законом композиции

$$(B, \boldsymbol{b}) \circ (A, \boldsymbol{a}) = (AB, \boldsymbol{a}B + \boldsymbol{b})$$

образует группу Ли *аффинных преобразований* $\mathbb{A}(n)$ аффинного пространства $\mathbb{A}$. Преобразования (2.49) при $A_\beta{}^\alpha = \delta^\alpha_\beta$ называются *трансляциями* на вектор $\boldsymbol{a} \in \mathbb{V}$. □

Множество трансляций образует инвариантную (нормальную) подгруппу аффинной группы, а сама аффинная группа представляет собой полупрямое произведение, которое будем обозначать символом $\ltimes$, группы невырожденных матриц $\mathbb{GL}(n, \mathbb{R})$ на подгруппу трансляций, которая естественным образом отождествляется с векторным пространством $\mathbb{V}$. Таким образом группа аффинных преобразований имеет вид

$$\mathbb{A}(n) = \mathbb{GL}(n, \mathbb{R}) \ltimes \mathbb{V}.$$

Подгруппа $\mathbb{A}(n)$, сохраняющая начало координат, совпадает с $\mathbb{GL}(n, \mathbb{R})$.

Аффинная группа является неабелевой и имеет размерность $n^2 + n$. Она действует в аффинном пространстве транзитивно и эффективно (см. раздел 10.1). Аффинные преобразования называются также *движениями аффинного пространства*. Если $\det A > 0$ или $\det A < 0$, то движения аффинного пространства (2.49) называются соответственно *собственными* или *несобственными*.

Замечание. В векторном пространстве действует только группа $\mathbb{GL}(n, \mathbb{R})$, т.к. в нем отсутствует понятие сдвига (трансляции). □

Определение. Подгруппа сдвигов действует на $\mathbb{R}^n$ свободно. Рассмотрим подгруппу $\mathbb{G} \subset \mathbb{V} = \mathbb{R}^n$, где $\mathbb{V}$ рассматривается, как группа трансляций, состоящую из всех сдвигов

$$\sum_{i=1}^{m} n^i \boldsymbol{a}_i, \qquad n^i \in \mathbb{Z} \tag{2.50}$$

на линейно независимые векторы $\boldsymbol{a}_i$, $i = 1, \dots m \leq n$ с целыми коэффициентами. Эта подгруппа является дискретной группой преобразований (см. раздел 10.1). Поэтому, отождествляя точки евклидова пространства, связанные трансляциями (2.50), получим важный класс многообразий, называемых *цилиндрами*. Они являются фактор пространствами $\mathbb{R}^n/\mathbb{G}$. При $m = n$ цилиндр компактен и называется *тором* $\mathbb{T}^n$. □

Пример 2.3.53. В частном случае, сдвиги евклидовой плоскости $\mathbb{R}^2$ с декартовыми координатами x, y на постоянный вектор $\boldsymbol{a} = \{a, 0\}$, $a \neq 0$, вдоль оси абсцисс,

$$x \mapsto x + a, \qquad y \mapsto y,$$

приводит к простейшему двумерному цилиндру, хорошо известному из курса элементарной геометрии. □

В одномерном случае, $n = 1$, тор совпадает с окружностью $\mathbb{T}^1 = \mathbb{S}^1$. В общем случае $1 \leq m < n$ векторное пространство разлагается в прямую сумму $\mathbb{V} = \mathbb{U} \oplus \mathbb{W}$, где $\mathbb{W}$ – подпространство, натянутое на векторы $\boldsymbol{a}_i$. Отсюда следует

$$\frac{\mathbb{R}^n}{\mathbb{G}} = \frac{\mathbb{R}^{n-m} \times \mathbb{R}^m}{\mathbb{G}} = \mathbb{R}^{n-m} \times \frac{\mathbb{R}^m}{\mathbb{G}}.$$

То есть цилиндры представляют собой прямое произведение $(n-m)$-мерного аффинного пространства и m-мерного тора.

Вернемся к аффинным преобразованиям (2.49).

Теорема 2.3.18. *Всякое аффинное преобразование n-мерного евклидова пространства $\mathbb{A} = \mathbb{R}^n$ может быть представлено в виде композиции трех отображений:*
1) n растяжений с положительными коэффициентами вдоль n попарно ортогональных осей, проходящих через некоторую точку $P_0 \in \mathbb{A}$;
2) движения, оставляющего неподвижной точку P_0;
3) сдвига.

Доказательство. Первые два преобразования вытекают из полярного разложения матриц, теорема 28.1.6. Остаются еще сдвиги. Детали доказательства приведены, например, в [**?**]. □

Определение. В аффинном пространстве *прямой линией*, проходящей через точку $\mathrm{P} \in \mathbb{A}$ в направлении вектора $\boldsymbol{x} \in \mathbb{V}$, называется множество точек вида

$$\mathrm{P} + a\boldsymbol{x}, \qquad a \in \mathbb{R}. \tag{2.51}$$

Подмножество $\mathbb{B} \subset \mathbb{A}$ называется *аффинным подпространством*, если множество векторов $(\mathrm{Q} - \mathrm{P})$, $\forall \mathrm{Q}, \mathrm{P} \in \mathbb{B}$ образует подпространство $\mathbb{U} \subset \mathbb{V}$ в ассоциированном векторном пространстве. □

Ясно, что каждое аффинное подпространство $\mathbb{B} \subset \mathbb{A}$ состоит из точек вида

$$\mathbb{B} := \{\mathrm{P} + \boldsymbol{x} : \quad \boldsymbol{x} \in \mathbb{U}\},$$

где P – произвольная точка из $\mathbb{B}$.

Пример 2.3.54. Прямая линия является одномерным аффинным подпространством в $\mathbb{A}$. □

При аффинном преобразовании прямые линии переходят в прямые, причем пересекающиеся линии переходят в пересекающиеся, а параллельные – в параллельные.

Определение. Пусть $\mathrm{P}_1, \ldots \mathrm{P}_k$ – произвольные точки аффинного пространства, и числа $a_1, \ldots a_k$ удовлетворяют условию $\sum_{i=1}^k a_i = 1$. Определим сумму

$$\mathrm{P} = \sum_{i=1}^k a_i \mathrm{P}_i := \mathrm{P}_0 + \sum_{i=1}^k a_i(\mathrm{P}_i - \mathrm{P}_0), \tag{2.52}$$

где P_0 – произвольная точка из $\mathbb{A}$. Поскольку $(\mathrm{P} - \mathrm{P}_0) \in \mathbb{V}$ и выражение (2.52) не зависит от выбора P_0, то оно корректно определяет точку аффинного пространства $\mathrm{P} \in \mathbb{A}$. Эта точка называется *барицентрической комбинацией* точек $\mathrm{P}_1 \ldots \mathrm{P}_k$ с коэффициентами $a_1 \ldots a_k$. □

Пример 2.3.55. В евклидовом аффинном пространстве барицентрическая сумма

$$\sum_{i=1}^k \frac{m_i}{M} \mathrm{P}_i, \qquad M = \sum_{i=1}^k m_i$$

представляет собой положение центра масс системы масс m_i, расположенных в точках P_i. □

Из определения (2.52) следует, что система

$$\{\mathrm{P}_0,\ (\mathrm{P}_1-\mathrm{P}_0),\dots,(\mathrm{P}_n-\mathrm{P}_0)\},$$

состоящая из точки P_0 и множества векторов $(\mathrm{P}_i-\mathrm{P}_0)$, образует систему аффинных координат тогда и только тогда, когда любая точка $\mathrm{P}\in\mathbb{A}$ представима в виде барицентрической суммы

$$\mathrm{P}=\sum_{i=1}^{n}x_i\mathrm{P}_i,\qquad \sum_{i=1}^{n}x_i=1.$$

Если это условие выполнено, то множество точек $\{\mathrm{P}_0,\dots,\mathrm{P}_n\}$ называется *барицентрической* системой координат в аффинном пространстве, а числа $x_i\in\mathbb{R}$ – барицентрическими координатами точки $\mathrm{P}\in\mathbb{A}$.

Определение. Пусть точки P_i, $i=1,\dots n$ имеют аффинные координаты

$$\mathrm{P}_i=\{\underbrace{0\dots0}_{i-1}10\dots0\}.$$

Эти точки вместе с началом системы координат образуют барицентрическую систему координат в аффинном пространстве. Рассмотрим пересечение множества точек с барицентрическими координатами, не превосходящими единицы, $0\le x_i\le 1$, с положительным 2^n-тантом. Это множество называется *стандартным* $(n-1)$*-мерным симплексом*. В общем случае множество точек

$$\left\{\sum_{i=1}^{n}x_i\mathrm{P}_i:\ \ \sum_{i=1}^{n}x_i=1,\ \ 0\le x_i\le 1\right\} \tag{2.53}$$

называется *замкнутым симплексом* с вершинами в точках P_i в аффинном пространстве $\mathbb{A}$. Симплекс называется *вырожденным*, если векторы $(\mathrm{P}_2-\mathrm{P}_1),\dots,(\mathrm{P}_n-\mathrm{P}_1)$ линейно зависимы. □

Пример 2.3.56. Одномерный симплекс – это отрезок прямой, двумерный – треугольник, трехмерный – тетраэдр (см. рис.2.3). □

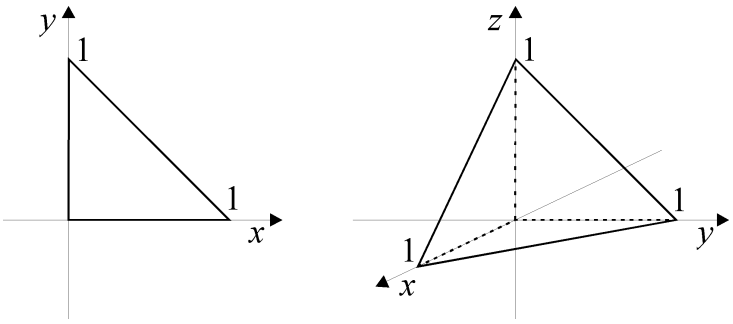


Рис. 2.3: Симплексы в двух и трех измерениях.

Определим теперь параллельный перенос векторов в аффинном пространстве.

Определение. Рассмотрим произвольную кривую $\gamma=x(t)$, соединяющую две произвольно выбранные точки P и Q в аффинном пространстве. С каждой точкой аффинного пространства и, следовательно, с каждой точкой кривой γ ассоциировано векторное пространство $\mathbb{V}_{x(t)}$. Будем говорить, что вектор $\boldsymbol{x}\in\mathbb{V}_{x(t)}$ *параллельно переносится* вдоль кривой γ, если его компоненты относительно некоторого фиксированного базиса не меняются при переходе от точки к точке. □

Замечание. В данном определении важно различать точки аффинного и векторного пространств. □

Очевидно, что параллельный перенос вектора из точки P в точку Q не зависит от кривой, соединяющей эти точки.

В настоящем и предыдущем разделах векторы обозначались жирным шрифтом. В дальнейшем мы упростим обозначения. Если у символа индекс присутствует, то это значит, что рассматривается соответствующая компонента вектора. Например, x^α обозначает компоненту вектора с индексом α. В то же время, если индекс опущен или используются фигурные скобки, то подразумевается весь вектор $x = \{x^\alpha\} = (x^1, \dots, x^n)$, и мы, как правило, не будем использовать жирный шрифт.

2.4 Отображения

Определение. Рассмотрим два множества $\mathbb{M}_1$ и $\mathbb{M}_2$ и их *отображение*

$$f: \quad \mathbb{M}_1 \to \mathbb{M}_2 \qquad \text{или} \qquad \mathbb{M}_1 \xrightarrow{f} \mathbb{M}_2, \tag{2.54}$$

которое ставит каждому элементу $x \in \mathbb{M}_1$ единственный элемент $f(x) \in \mathbb{M}_2$. Отображение f называется также *функцией*, а переменная $x \in \mathbb{M}_1$ –*аргументом*. Вообще говоря, функция может отображать различные элементы из $\mathbb{M}_1$ в один и тот же элемент множества $\mathbb{M}_2$. Ситуация, когда один элемент из $\mathbb{M}_1$ отображаются сразу в несколько элементов множества $\mathbb{M}_2$, не допускается. Пространство $\mathbb{M}_2$ в (2.54) называется *пространством-мишенью*. *Тождественное* отображение

$$\mathsf{id}\,(\mathbb{M}): \quad \mathbb{M} \to \mathbb{M}$$

определяется равенством $\mathsf{id}\,(x) = x$ для всех $x \in \mathbb{M}$. Отображение f называется *взаимно однозначным* или *инъективным*, если различные точки $x_1 \neq x_2$ из $\mathbb{M}_1$ отображаются в различные точки $f(x_1) \neq f(x_2)$ из $\mathbb{M}_2$. Отображение f является отображением $\mathbb{M}_1$ *на* $\mathbb{M}_2$ или *сюръективным*, если $f(\mathbb{M}_1) = \mathbb{M}_2$. Если отображение f инъективно и сюръективно одновременно, то оно называется *биективным отображением* или *биекцией*. Каждое отображение порождает некоторое подмножество в прямом произведении $\mathbb{M}_1 \times \mathbb{M}_2$,

$$\mathsf{Gr}_f := \{x, f(x): \quad x \in \mathbb{M}_1\} \subset,$$

которое называется *графиком* отображения f. □

Пример 2.4.1. Рассмотрим дифференцируемую функцию одной переменной $f(x) \in \mathcal{C}^1(\mathbb{R})$. Если эта функция монотонно возрастает, $f' > 0$, или монотонно убывает, $f' < 0$, на интервале $x \in (p, q)$, то она задает биективное отображение интервалов $f: \ (p, q) \to \Big(f(p), f(q)\Big)$. □

Пример 2.4.2. Преобразование координат с невырожденным якобианом является биекцией двух областей евклидова пространства (теорема 2.5.1). □

Сопоставление аргументу $x \in \mathbb{M}_1$ в (2.54) значения функции $f(x) \in \mathbb{M}_2$ будем обозначать ограниченной стрелкой:

$$f: \quad \mathbb{M}_1 \ni x \mapsto f(x) \in \mathbb{M}_2.$$

Определение. Множество $f(\mathbb{M}_1) = \mathbb{U} \subset \mathbb{M}_2$, которое может и не совпадать со всем $\mathbb{M}_2$, называется *образом* множества $\mathbb{M}_1$ или *областью значений* функции f. Множество $\mathbb{M}_1$ при этом называется *прообразом* $\mathbb{U}$ или *областью определения* f и обозначается $\mathbb{M}_1 = f^{-1}(\mathbb{U})$. Вообще, прообразом произвольного подмножества $\mathbb{V} \subset \mathbb{U} \subset \mathbb{M}_2$ называется множество тех точек $\mathbb{W} \subset \mathbb{M}_1$, для которых $f(\mathbb{W}) = \mathbb{V}$. При этом символ f^{-1} является отображением

$$f^{-1}: \quad \mathbb{U} \ni f(x) \mapsto x \in \mathbb{M}_1$$

тогда и только тогда, когда отображение f является взаимно однозначным. В этом случае оно единственно и называется *обратным отображением.* □

Для биективного отображения $f: \mathbb{M}_1 \to \mathbb{M}_2$ всегда существует обратное отображение

$$f^{-1}: \quad \mathbb{M}_2 \to \mathbb{M}_1.$$

При этом отображение f^{-1} также биективно и выполнено равенство

$$(f^{-1})^{-1} = f.$$

Если отображение f не является взаимно однозначным, то прообразом одной точки $f^{-1}(x)$, где $x \in \mathbb{U}$, может быть несколько элементов из $\mathbb{M}_1$ или их бесконечное число. В этом случае символ f^{-1} не является отображением.

Пример 2.4.3. Рассмотрим три отображения:

$$f: \mathbb{R} \to \mathbb{R}, \qquad g: \mathbb{R} \to \mathbb{R}_+ \cup \{0\}, \qquad h: \mathbb{R}_+ \to \mathbb{R}_+,$$

которые определены одним и тем же правилом $x \mapsto x^2$. Эти отображения различны: f не сюръективно, не инъективно; g сюръективно, но не инъективно; h биективно. Этот пример показывает, что задание области определения и области значений является существенной частью определения отображения (функции). □

Определение. Пусть f – функция (2.54) и $\mathbb{U}_1$ – некоторое подмножество $\mathbb{M}_1$, тогда отображение

$$f|_{\mathbb{U}_1}: \quad \mathbb{U}_1 \to \mathbb{M}_2,$$

называется *сужением* функции (или отображения) на $\mathbb{U}_1 \subset \mathbb{M}_1$. Функция f является *продолжением* функции

$$g: \quad \mathbb{U}_1 \to \mathbb{M}_2,$$

заданной на некотором подмножестве $\mathbb{U}_1 \subset \mathbb{M}_1$, если ее сужение на $\mathbb{U}_1$ совпадает с g, $f|_{\mathbb{U}_1} = g$. □

Определение. *Композицией*, (*произведением* или *суперпозицией*) двух отображений $f: \mathbb{M}_1 \to \mathbb{M}_2$ и $g: \mathbb{M}_2 \to \mathbb{M}_3$ называется отображение

$$g \circ f: \quad \mathbb{M}_1 \to \mathbb{M}_3,$$

которое состоит в последовательном применении этих отображений:

$$(g \circ f)(x) = g\Big(f(x)\Big).$$ □

Замечание. Произведение определено не для всех отображений, а только для тех, у которых множество $\mathbb{M}_2$ в предыдущих обозначениях общее. Если рассматриваются отображения некоторого множества в себя, то для них композиция всегда определена. □

Композиция отображений есть не что иное, как сложная функция.

Если задано отображение $f:\ \mathbb{M}_1 \to \mathbb{M}_2$, то

$$f \circ \mathsf{id}\,(\mathbb{M}_1) = f \quad \text{и} \quad \mathsf{id}\,(\mathbb{M}_2) f = f.$$

Для биективного отображения выполнены равенства:

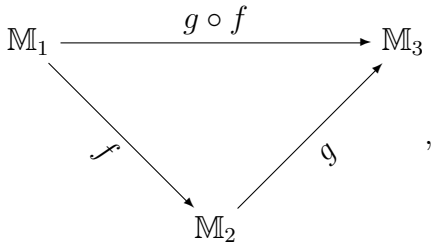
$$f^{-1} \circ f = \mathsf{id}\,(\mathbb{M}_1) \quad \text{и} \quad f \circ f^{-1} = \mathsf{id}\,(\mathbb{M}_2).$$

Если заданы два биективных отображения f и g и определена их композиция, т.е. область значений f совпадает с областью определения g, то обратное отображение для композиции существует и задается формулой

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

В дальнейшем вместо $g \circ f$ мы часто будем писать просто gf.

Произведение отображений играет огромную роль в математике, и его изображают в виде диаграммы



,

которая называется *коммутативной*. Это значит, что результат перехода от $\mathbb{M}_1$ к $\mathbb{M}_3$ не зависит от того, по какому пути мы движемся: то ли вдоль стрелки $g \circ f$, то ли последовательно вдоль стрелок f и g. Изображение произведений отображений в виде коммутативных диаграмм наглядно и полезно. Это будет продемонстрировано в следующем утверждении.

Предложение 2.4.1. *Пусть определено произведение трех отображений:*

$$f:\ \mathbb{M}_1 \to \mathbb{M}_2, \qquad g:\ \mathbb{M}_2 \to \mathbb{M}_3, \qquad h:\ \mathbb{M}_3 \to \mathbb{M}_4.$$

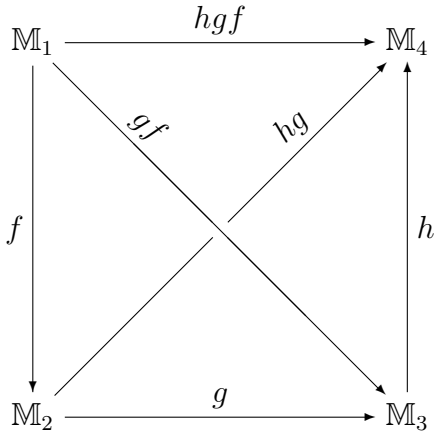
Тогда произведение этих отображений ассоциативно:

$$h(gf) = (hg)f$$

и его обозначают $hgf:\ \mathbb{M}_1 \to \mathbb{M}_4$.

Доказательство. Все необходимые рассуждения содержатся в следующей диаграм-

ме:



Для доказательства необходимо проследить куда отобразится произвольный элемент $x \in \mathbb{M}_1$ под действием композиций $h(gf)$ и $(hg)f$. Формально доказательство выглядит следующим образом:

$$\Big(h(gf)\Big)x = h\Big((gf)x\Big) = h\Big(g(fx)\Big) = (hg)(hx) = \Big((hg)f\Big)x. \qquad \square$$

Отображения конечных множеств обладают простыми свойствами.

Предложение 2.4.2. *Если отображение $f : \mathbb{M} \to \mathbb{M}$ конечного множества $\mathbb{M}$ в себя инъективно или сюръективно, то оно биективно.*

Доказательство. Пусть отображение f инъективно. Покажем, что оно является также сюръективным. Для произвольного элемента множества $x \in \mathbb{M}$ определим последовательность

$$f^0(x) = x, \quad \text{и} \quad f^k(x) = \underbrace{f\Big(\dots f}_{k}(x)\dots\Big) = f\Big(f^{k-1}(x)\Big), \quad k = 1, 2, \dots .$$

Из конечности множества $\mathbb{M}$ вытекает, что в этой последовательности есть повторения. Допустим, скажем, что $f^m(x) = f^n(x)$ при некоторых $m > n$. Если $n > 0$, то из равенства $f\Big(f^{m-1}(x)\Big) = f\Big(f^{n-1}(x)\Big)$ и инъективности f следует $f^{m-1}(x) = f^{n-1}(x)$. Сократив отображение f n раз, приходим к равенству

$$f^{m-n}(x) = x.$$

В этом случае для любого $x \in \mathbb{M}$ найдется такое $x' = f^{m-n-1}(x)$, что $f(x') = x$, т.е. отображение f сюръективно.

Второе утверждение предложения докажем от противного. Пусть отображение f сюръективно, но не инъективно. Тогда существует такой элемент $x \in \mathbb{M}$, что его прообраз состоит не менее, чем из двух элементов. Пусть множество $\mathbb{M}$ содержит N элементов. Тогда f должно отобразить оставшиеся элементы, которых не более, чем $N - 2$, на $N - 1$ элемент. Для сюръективного отображения это невозможно. $\square$

В приведенном доказательстве существенна конечность множества $\mathbb{M}$.

Пример 2.4.4. *Отображение следования*, определенное для натуральных чисел сдвигом на единицу,

$$\sigma : \quad \mathbb{N} \ni n \mapsto n + 1 \in \mathbb{N},$$

является инъективным, но не сюръективным, поскольку единица не принадлежит образу $\sigma(\mathbb{N})$. $\square$

Определение. Мы говорим, что множества $\mathbb{M}_1$ и $\mathbb{M}_2$ имеют одинаковую *мощность*, если существует биективное отображение $f: \mathbb{M}_1 \to \mathbb{M}_2$. Множества, имеющие ту же мощность, что и множество натуральных чисел $\mathbb{N}$ называется *счетным*. □

2.4.1 Отображения топологических пространств

При отображении топологических пространств важную роль играют отображения, согласованные с их топологией.

Определение. Отображение (2.54) двух топологических пространств называется *непрерывным* в точке $x \in \mathbb{M}_1$, если для каждой окрестности $\mathbb{U}_2$ точки $f(x) \in \mathbb{M}_2$ существует такая окрестность $\mathbb{U}_1$ точки x, что ее образ $f(\mathbb{U}_1)$ содержится в $\mathbb{U}_2$. Отображение называется непрерывным в области, если оно непрерывно в каждой точке этой области. □

Замечание. В общем случае окрестность $\mathbb{U}_2 \subset \mathbb{M}$ может содержать точки, не лежащие в образе $f(\mathbb{M}_1)$ и образ $f(\mathbb{U}_1)$ совсем не обязан совпадать с $\mathbb{U}_2$. □

Определение. Говорят, что два топологических пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ *гомеоморфны*, если существуют взаимно обратные непрерывные отображения $f: \mathbb{M}_1 \to \mathbb{M}_2$ и $g: \mathbb{M}_2 \to \mathbb{M}_1$ такие, что $g \circ f = \mathsf{id}\,(\mathbb{M}_1)$, $f \circ g = \mathsf{id}\,(\mathbb{M}_2)$. Тогда отображения f и g называются *гомеоморфизмами*. □

Ясно, что любой гомеоморфизм является биекцией. Обратное утверждение неверно.

Пример 2.4.5. Рассмотрим функцию

$$f(x) = \begin{cases} x, & x \le -1 \text{ или } x \ge 1, \\ -x, & x \in (-1, 1), \end{cases}$$

показанную на рис.2.4. Эта функция задает отображение $\mathbb{R} \to \mathbb{R}$, которое биективно, но не задает гомеоморфизм. □

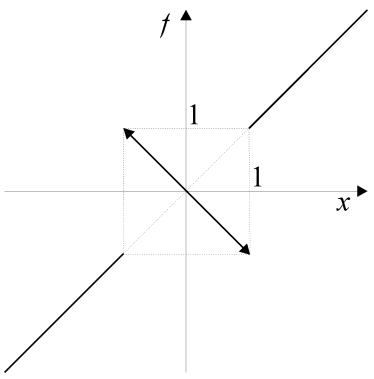


Рис. 2.4: Биективное отображение прямых, которое не является гомеоморфизмом.

Справедливы две фундаментальные теоремы, которые приведем без доказательства.

Теорема 2.4.1 (Инвариантность размерности). *Открытое подмножество в* $\overline{\mathbb{R}}_+^m$ *не может быть гомеоморфно никакому открытому подмножеству в* $\overline{\mathbb{R}}_+^n$, *если* $m \neq n$.

В этой теореме замкнутое полупространство $\overline{\mathbb{R}}^n_+$ можно заменить на все евклидово пространство $\mathbb{R}^n$.

Теорема 2.4.2 (Инвариантность края). *Пусть $\mathbb{U}$ и $\mathbb{V}$ – два открытых подмножества $\overline{\mathbb{R}}^n_+$ и f – гомеоморфизм $\mathbb{U}$ на $\mathbb{V}$, тогда $f(\mathbb{U}\bigcap\mathbb{R}^{n-1}) = \mathbb{V}\bigcap\mathbb{R}^{n-1}$, где $\mathbb{R}^{n-1}$ – гиперплоскость (край), определяемая условием $x^n = 0$. Другими словами, при гомеоморфизмах точки края отображаются в край.*

Приведем пять критериев (необходимых и достаточных условий) непрерывности отображения, которые будут использоваться в дальнейшем.

Теорема 2.4.3. *Если отображение $f:\ \mathbb{M}_1 \to \mathbb{M}_2$ сюръективно, то оно непрерывно тогда и только тогда, когда выполнено одно из следующих условий:*

1) прообраз любого замкнутого подмножества $\mathbb{M}_2$ замкнут в $\mathbb{M}_1$;

2) прообраз любого открытого подмножества $\mathbb{M}_2$ открыт в $\mathbb{M}_1$;

3) $\forall\mathbb{U}_1 \subset \mathbb{M}_1$ имеем $f(\overline{\mathbb{U}}_1) \subset \overline{f(\mathbb{U}_1)}$;

4) $\forall\mathbb{U}_2 \subset \mathbb{M}_2$ имеем $\overline{f^{-1}(\mathbb{U}_2)} \subset f^{-1}(\overline{\mathbb{U}}_2)$;

5) $\forall\mathbb{U}_2 \subset \mathbb{M}_2$ имеем $f^{-1}(\,\mathsf{int}\,\mathbb{U}_2) \subset\ \mathsf{int}\, f^{-1}(\mathbb{U}_2)$.

Доказательство. См., например, [**?**]. □

Определение и критерии непрерывности отображений говорят о топологических свойствах прообраза, исходя из свойств образа. При этом нельзя сделать однозначного утверждения относительно свойств образа при заданных свойствах прообраза. Действительно, из критерия 3) следует, что образ замкнутого множества, вообще говоря, не является замкнутым. Критерий 5), записанный в виде

$$\mathsf{int}\,\mathbb{U}_2 \subset f\Big(\,\mathsf{int}\, f^{-1}(\mathbb{U}_2)\Big),$$

говорит о том, что образ открытого множества может не быть открытым.

Определение. Непрерывное отображение (2.54) называется *открытым* (*замкнутым*) если образ любого открытого (замкнутого) подмножества $\mathbb{M}_1$ открыт (замкнут) в $\mathbb{M}_2$. □

Пример 2.4.6. Гомеоморфизм двух топологических пространств является открытым и замкнутым одновременно, т.е. открытое множество отображается в открытое, а замкнутое – в замкнутое. □

Пример 2.4.7. Приведем пример замкнутого, но не открытого непрерывного сюръективного отображения $f:\ \mathbb{R} \to \mathbb{R}$ (см. рис.2.5), которое задано следующим образом

$$f(x) = \begin{cases} x+1, & x \le -2, \\ -1, & -2 \le x \le -1, \\ x, & -1 \le x \le 1, \\ 1, & 1 \le x \le 2, \\ x-1, & 2 \le x. \end{cases}$$

При таком отображении открытый интервал (a,b), где $a \in (-2,-1)$ и $b \in (1,2)$ отображается в замкнутый отрезок $[-1,1]$. □

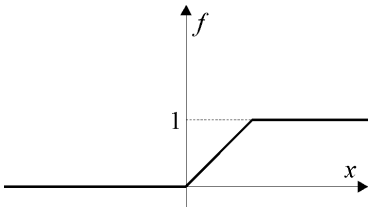


Рис. 2.5: Пример замкнутого, но не открытого непрерывного сюръективного отображения.

Рассмотрим отображение $f: \mathbb{M}_1 \to \mathbb{M}_2$ топологического пространства $\mathbb{M}_1$ в некоторое множество $\mathbb{M}_2$. Если отображение f является сюръективным, т.е. отображением на, то на множестве $\mathbb{M}_2$ можно определить *топологию идентификации* или *фактортопологию*. А именно, назовем подмножество $\mathbb{U} \subset \mathbb{M}_2$ открытым, если его прообраз $f^{-1}(\mathbb{U}) \subset \mathbb{M}_1$ открыт в $\mathbb{M}_1$. Это самая тонкая топология на $\mathbb{M}_2$, в которой отображение f является непрерывным.

Пример 2.4.8. Пусть $\mathbb{M}$ – топологическое пространство, на котором задано некоторое отношение эквивалентности $x_1 \sim x_2$. Пусть $\mathbb{M}/\sim$ – множество классов эквивалентности, и $f: \mathbb{M} \to \mathbb{M}/\sim$ – каноническая проекция элемента $x \in \mathbb{M}$ на его класс эквивалентности. Тогда отображение f сюръективно и на фактор пространстве $\mathbb{M}/\sim$ можно ввести топологию идентификации. □

На самом деле, этот пример является общим. Если задано сюръективное отображение $f: \mathbb{M}_1 \to \mathbb{M}_2$, то на множестве $\mathbb{M}_1$ всегда можно определить отношение эквивалентности: будем считать точки $x_1, x_2 \in \mathbb{M}_1$ эквивалентными, если $f(x_1) = f(x_2)$. Другими словами, в множестве $\mathbb{M}_1$ мы идентифицируем те точки, которые отображаются в одну точку пространства-мишени $\mathbb{M}_2$. Этим объясняется название “топология идентификации”.

Определение. Если некоторое свойство P топологического пространства сохраняется при гомеоморфизме, то оно называется *топологическим инвариантом* или *свойством*. Другими словами, топологическое пространство $\mathbb{M}$ обладает свойством P в том и только в том случае, если этим свойством обладает любое другое гомеоморфное ему топологическое пространство. □

Так как гомеоморфизм устанавливает взаимно однозначное соответствие не только между точками, но и между открытыми множествами, то любое свойство, определенное в терминах открытых множеств, является топологическим инвариантом.

Пример 2.4.9. Число связных компонент топологического пространства является топологическим инвариантом. □

Если отображение (2.54) непрерывно, но не является гомеоморфизмом, то число связных компонент может уменьшаться, но не увеличиваться.

Пример 2.4.10. Если пространство $\mathbb{M}_1 = \mathbb{I}_1 \cup \mathbb{I}_2$ является объединением двух непересекающихся открытых интервалов $\mathbb{I}_1$ и $\mathbb{I}_2$, то его можно непрерывно отобразить на произвольный открытый интервал $\mathbb{I}_3$. При этом $\mathbb{I}_1 \to \mathbb{I}_3$ и $\mathbb{I}_2 \to \mathbb{I}_3$, т.е. каждый из интервалов $\mathbb{I}_1$ и $\mathbb{I}_2$ отображается на весь интервал $\mathbb{I}_3$. Обратное утверждение неверно. Один интервал невозможно непрерывно отобразить на два. Действительно, если существует непрерывное отображение

$$g: \quad \mathbb{I}_3 \to \mathbb{I}_1 \cup \mathbb{I}_2,$$

то в силу непрерывности g прообразы $\mathbb{I}_1$ и $\mathbb{I}_2$ являются открытыми и замкнутыми непустыми подмножествами в $\mathbb{I}_3$ одновременно. Поскольку интервал $\mathbb{I}_3$ связен, то $\mathbb{I}_3 = g^{-1}(\mathbb{I}_1) = g^{-1}(\mathbb{I}_2)$. Это означает, что каждой точке интервала $\mathbb{I}_3$ соответствует две точки в объединении $\mathbb{I}_1 \bigcup \mathbb{I}_2$, что противоречит определению отображения (2.54). □

Определение. *Компактификацией* топологического пространства $\mathbb{M}$ называется пара $(\mathbb{U}, f)$, где $\mathbb{U}$ – компактное топологическое пространство, а f – гомеоморфизм $\mathbb{M}$ на всюду плотное подмножество в $\mathbb{U}$. □

Пример 2.4.11. Отображение евклидовой плоскости на сферу, $f : \mathbb{R}^2 \to \mathbb{S}^2$, которое задается стереографической проекцией (см. раздел 25.1) является компактификацией. При этом все бесконечно удаленные точки евклидовой плоскости отображаются на одну точку сферы – северный полюс. □

Пример 2.4.12. Добавим к евклидову пространству $\mathbb{R}^n$ бесконечно удаленную точку ∞ и объявим ее окрестностями множества вида $(\mathbb{R}^n \setminus \mathbb{B}) \cup \infty$ для всех ограниченных замкнутых подмножеств $\mathbb{B} \subset \mathbb{R}^n$. Тогда объединение $\mathbb{R}^n \cup \infty$ будет компактным пространством, гомеоморфным сфере $\mathbb{S}^n$. В этом примере пара $\Big((\mathbb{R}^n \cup \infty), f\Big)$, где отображение $f :\ \mathbb{R}^n \to \mathbb{R}^n \cup \infty$ имеет вид

$$f :\quad \mathbb{R}^n \ni x \mapsto (x, \infty) \in \mathbb{R}^n \cup \infty,$$

является компактификацией евклидова пространства. □

Пусть задано семейство $\mathcal{F}$ отображений топологического пространства $\mathbb{M}_1$ в топологическое пространство $\mathbb{M}_2$, т.е. элементами семейства $f \in \mathcal{F}$ являются отображения (2.54). Тогда возникает естественный вопрос, можно ли на $\mathcal{F}$ задать такую топологию, что отображение топологического произведения

$$\mathcal{F} \times \mathbb{M}_1 \ni (f, x) \mapsto f(x) \in \mathbb{M}_2$$

является непрерывным. Ответ положительный: такие топологии существуют. Ниже мы рассмотрим две из них. ???

Отображения хаусдорфовых топологических пространств. Поскольку многообразия, по определению, являются хаусдорфовыми топологическими пространствами, то отображения хаусдорфовых пространств будут играть в дальнейшем важную роль.

Доказательство следующих двух теорем и следствий содержится, например, в [**?**].

Теорема 2.4.4. *Если компактное пространство $\mathbb{M}_1$ непрерывно отображается на пространство $\mathbb{M}_2$, то $\mathbb{M}_2$ компактно.*

Следствие. Если $f :\ \mathbb{M}_1 \to \mathbb{M}_2$ – непрерывное отображение компактного пространства $\mathbb{M}_1$ в хаусдорфово пространство $\mathbb{M}_2$, то $\overline{f(\mathbb{U})} = f(\overline{\mathbb{U}})$ для каждого подмножества $\mathbb{U} \subset \mathbb{M}_1$.

Следствие. Каждое непрерывное отображение компактного пространства в хаусдорфово пространство замкнуто.

Теорема 2.4.5. *Каждое непрерывное взаимно однозначное отображение компактного пространства на хаусдорфово пространство является гомеоморфизмом.*

Приведем несколько утверждений, доказательство которых можно найти, например, в [**?**].

Теорема 2.4.6. *Компактное подмножество* $\mathbb{U}$ *хаусдорфова топологического пространства* $\mathbb{M}$ *хаусдорфово и замкнуто.*

Теорема 2.4.7. *Пусть* $\mathbb{M}_1$ *и* $\mathbb{M}_2$ *– топологические пространства. Тогда* $\mathbb{M}_1$ *и* $\mathbb{M}_2$ *хаусдорфовы в том и только в том случае, если произведение* $\mathbb{M}_1 \times \mathbb{M}_2$ *хаусдорфово.*

Хотя подпространства и произведения хаусдорфовых пространств хаусдорфовы, факторпространство хаусдорфова пространства, вообще говоря, нехаусдорфово.

Пример 2.4.13. Рассмотрим вещественную прямую $\mathbb{R}$ и открытый интервал $(0,1)$. Введем отношение эквивалентности $\sim$ на $\mathbb{R}$, при котором $x \sim x'$ тогда и только тогда, когда $x = x'$ или $x, x' \in (0,1)$. Грубо говоря, фактор пространство $X/\sim$ – это вещественная прямая $\mathbb{R}$ со стянутым в точку интервалом $(0,1)$. Если снабдить $\mathbb{R}/\sim$ фактортопологией относительно естественной проекции $\pi : \mathbb{R} \ni x \mapsto \tilde{x} \in \mathbb{R}/\sim$, то прообраз точки $\tilde{x}_0$, где $x_0 \in (0,1)$ есть интервал $(0,1)$, который открыт в $\mathbb{R}$. Следовательно, точка $\tilde{x}_0$ открыта в $\mathbb{R}/\sim$. Таким образом фактор пространство $\mathbb{M}/\sim$ представляет собой объединение двух лучей $(-\infty, 0]$, $[1, \infty)$ и точки $\tilde{x}_0$, которая открыта. Любые окрестности точек 0 и 1 содержат точку $\tilde{x}_0$ и, потому, пересекаются. Следовательно, факторпространство $\mathbb{R}/\sim$ нехаусдорфово. Подчеркнем, что в рассматриваемом примере фактор пространство $\mathbb{R}/\sim$ снабжено фактор топологией (топологией идентификации), которая отличается от топологии, индуцируемой вложением в $\mathbb{R}^2$. □

В рассмотренном примере факторпространство $\mathbb{R}/\sim$ не является многообразием, хотя проекция π непрерывна. Для того, чтобы обеспечить хаусдорфовость факторпространства $\mathbb{N}$ хаусдорфова пространства $\mathbb{M}$ по некоторому отношению эквивалентности необходимо наложить дополнительные условия. Достаточное условие дает следующая

Теорема 2.4.8. *Пусть* $\mathbb{N}$ *– факторпространство топологического пространства* $\mathbb{M}$*, определенное при помощи сюръективного отображения* $f : \mathbb{M} \to \mathbb{N}$*. Если пространство* $\mathbb{M}$ *компактно и хаусдорфово, а отображение* f *замкнуто, то* $\mathbb{N}$ *компактно и хаусдорфово.*

Следствие. Если $\mathbb{M}$ – компактное хаусдорфово пространство, на котором действует конечная группа преобразований $\mathbb{G}$, то $\mathbb{M}/\mathbb{G}$ – компактное хаусдорфово пространство. □

Пример 2.4.14. Проективное пространство $\mathbb{RP}^n$ получается из сферы $\mathbb{S}^n$ отождествлением диаметрально противоположных точек, $\mathbb{RP}^n = \mathbb{S}^n/\mathbb{Z}_2$. Это есть компактное хаусдорфово пространство. □

Чтобы сформулировать другое следствие теоремы 2.4.8, рассмотрим пространство $\mathbb{M}$ вместе с его подмножеством $\mathbb{U} \subset \mathbb{M}$. Обозначим через $\mathbb{M}/\mathbb{U}$ фактор пространство $\mathbb{M}/\sim$, где $\sim$ – отношение эквивалентности на $\mathbb{M}$, при котором $x \sim x'$ тогда и только тогда, когда $x = x'$ или $x, x' \in \mathbb{U}$.

Следствие. Если $\mathbb{M}$ – компактное хаусдорфово пространство и $\mathbb{U}$ – его замкнутое подмножество, то $\mathbb{M}/\mathbb{U}$ – компактное хаусдорфово пространство. □

Замечание. В примере 2.4.13 подмножество $(0,1) \subset \mathbb{R}$ открыто, и фактор пространство $\mathbb{R}/\sim$ не является хаусдорфовым. □

2.5 Преобразования координат

Важным примером отображений являются преобразования координат, которые рассмотрены в настоящем разделе. Евклидово пространство $\mathbb{R}^n$ было определено в разделе 2.3.2 как топологическое произведение n прямых. При этом каждая точка $x \in \mathbb{R}^n$ задается набором n вещественных чисел:

$$x = (x^1, \ldots, x^n) = \{x^\alpha\}, \qquad \alpha = 1, \ldots, n,$$

– декартовых координат, которые покрывают все евклидово пространство. Рассмотрим n дифференцируемых функций (преобразование координат)

$$x^{\alpha'} = x^{\alpha'}(x) \quad \in \mathcal{C}^1(\mathbb{U}), \qquad \alpha' = 1, \ldots, n, \tag{2.55}$$

от декартовых координат x^α, заданных в некоторой области $\mathbb{U} \subset \mathbb{R}^n$. В общем случае эти функции отображают область $\mathbb{U} \subset \mathbb{R}^n$ на некоторое множество точек $\mathbb{U}' \subset \mathbb{R}^n$ – образ отображения. В наших обозначениях запись $x^{\alpha'}$ эквивалентна записи x'^α. При этом штрихованные индексы пробегают те же значения, что и нештрихованные $\alpha, \alpha' = 1, \ldots, n$. Такая запись удобнее, т.к. позволяет во многих случаях опускать букву x. В сокращенной записи мы опускаем индексы:

$$\mathbb{U} \ni x \mapsto x'(x) \in \mathbb{U}'. \tag{2.56}$$

Определение. Матрица, составленная из частных производных,

$$J_\alpha{}^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \partial_\alpha x^{\alpha'} \tag{2.57}$$

называется *матрицей Якоби*. Здесь индекс α считается первым, а индекс α' – вторым. Определитель матрицы Якоби называется *якобианом*,

$$J = \frac{\partial(x^{\alpha'})}{\partial(x^\alpha)} := \mathsf{det}\left(\frac{\partial x^{\alpha'}}{\partial x^\alpha}\right). \quad \square \tag{2.58}$$

Из курса математического анализа известно следующее утверждение.

Теорема 2.5.1. *Пусть функции (2.55) задают дифференцируемое отображение класса $\mathcal{C}^k$ открытого множества $\mathbb{U} \subset \mathbb{R}^n$ в $\mathbb{R}^n$. Если якобиан этого отображения не обращается в нуль в точке $x_0 \in \mathbb{U}$, то существуют такие окрестности $\mathbb{U}_0 \subset \mathbb{U}$ и $\mathbb{U}'_0 \subset \mathbb{R}^n$ точек x_0 и $x'_0 = x'_0(x_0)$, что функции $x^{\alpha'}(x)$ задают взаимно однозначное отображение окрестности $\mathbb{U}_0$ на $\mathbb{U}'_0$ (биекцию), и обратное ему отображение дифференцируемо и того же класса гладкости $\mathcal{C}^k$ в области $\mathbb{U}'_0$. В частности, возможно $k = \infty$.*

Следствие. Пусть $x'(x)$ – дифференцируемое отображение открытого множества $\mathbb{U} \subset \mathbb{R}^n$ в $\mathbb{R}^n$. Если якобиан этого отображения отличен от нуля на всем $\mathbb{U}$, то образ этого множества $\mathbb{U}'$ также является открытым множеством. В частности, окрестность точки x отображается в окрестность точки x'. $\square$

Пусть задана последовательность биекций:

$$\mathbb{U} \xrightarrow{x'(x)} \mathbb{U}' \xrightarrow{x''(x')} \mathbb{U}'' \to \ldots$$

Мы говорим, что точка $x \in \mathbb{U} \subset \mathbb{R}^n$ евклидова пространства задается n вещественными числами

$$x = \{x^\alpha\} = \{x^{\alpha'}\} = \{x^{\alpha''}\} = \dots,$$

которые являются координатами одной и той же точки в различных системах координат. Теперь мы можем расширить понятие евклидова пространства, допустив произвольные замены координат. Тогда мы считаем, что евклидово пространство $\mathbb{R}^n$ покрывается некоторыми областями (картами)

$$\mathbb{R}^n = \bigcup_i \mathbb{U}_i,$$

на каждой из которых выбрана своя система координат. Рассмотрим две пересекающиеся карты $\mathbb{U}_i$ и $\mathbb{U}_j$ в $\mathbb{R}^n$. В частном случае одна или обе области могут совпадать со всем евклидовым пространством. Обозначим координаты точек, принадлежащих $\mathbb{U}_i$ и $\mathbb{U}_j$, соответственно через x^α и $x^{\alpha'}$. Тогда в области пересечения карт $\mathbb{U}_i \cap \mathbb{U}_j$ координаты точек будут связаны некоторым преобразованием вида (2.55). По этой причине функции (2.55) называются *функциями перехода* к новой системе координат или *функциями склейки*. Заметим, что их одновременно можно рассматривать и как функции, осуществляющие склейку двух областей и как замену координат на пересечении $\mathbb{U}_i \cap \mathbb{U}_j$.

Замечание. В дифференциальной геометрии такие объекты как векторные поля, метрика, связность и другие геометрические структуры вводятся инвариантным образом, независимо от выбранной системы координат. В приложениях выбор той или иной системы координат диктуется симметрией задачи, и эти координаты, как правило, не покрывают все евклидово пространство. Например, сферические и цилиндрические координаты определены всюду в $\mathbb{R}^3$ за исключением оси z. □

Функции (2.55) осуществляют замену или *преобразование координат*. Это преобразование в зависимости от обстоятельств можно рассматривать как *пассивное* или *активное*. При пассивном рассмотрении считается, что одной и той же точке области соответствуют различные координаты в различных системах отсчета. Именно так мы и рассматривали преобразование координат до сих пор. В втором случае считается, что заданная точка евклидова пространства с координатами x^α в результате некоторого преобразования или деформирования области занимает новое положение с координатами $x^{\alpha'}$ относительно старой координатной системы. Такая точка зрения принята в теории групп преобразований.

Пример 2.5.1. Вращение евклидова пространства относительно начала в данной декартовой системе координат обычно рассматривают как активное. С математической точки зрения оба подхода равноправны, т.к. определяются одним и тем же набором функций (2.55), а деление преобразований на активные и пассивные зависит от физической интерпретации и традиций. □

Обычно от функций перехода (2.55) требуется, чтобы они и их обратные $x^\alpha = x^\alpha(x')$ были достаточно гладкими. В частности, непрерывность функций перехода означает гомеоморфность отображения (2.55). В дальнейшем, если не оговорено противное, мы рассматриваем преобразования координат класса $\mathcal{C}^\infty$. В этом случае гомеоморфизм (2.56) называется *диффеоморфизмом*.

Отличие от нуля якобиана (2.58) означает, что функции (2.55), определяющие преобразования координат, функционально независимы.

Определение. Набор функций $u^i(x^\alpha)$, $i = 1, \ldots, \mathrm{N}$, в некоторой области $\mathbb{U}$ называется *функционально зависимым*, если на любом компакте из $\mathbb{U}$ существует функция от N переменных $F(u^i)$, определенная в области значений u^i, которая непрерывна вместе со всеми частными производными первого порядка, не равна тождественно нулю ни в какой подобласти и для которой выполнено соотношение $F\Big(u^i(x^\alpha)\Big) = 0$. □

Дифференцируя уравнение $F\Big(u^i(x^\alpha)\Big) = 0$ по x^α, получим равенство

$$\frac{\partial F}{\partial u^i}\frac{\partial u^i}{\partial x^\alpha} = 0.$$

Поскольку среди функций $\partial F/\partial u^i$ по крайней мере одна отлична от нуля, то для функциональной зависимости необходимо, чтобы ранг матрицы $\partial u^i/\partial x^\alpha$ был меньше N. Это условие является и достаточным. В частности, при $\mathrm{N} > n$ функции всегда будут функционально зависимы. Тем самым можно дать эквивалентное

Определение. N функций, для которых ранг матрицы $\partial u^i/\partial x^\alpha$ равен N в каждой точке области $\mathbb{U} \subset \mathbb{R}^n$ называются *функционально независимыми* в этой области. □

Таким образом, n достаточно гладких функционально независимых функций (2.55) задают преобразование координат в некоторой области $\mathbb{U} \subset \mathbb{R}^n$.

Преобразование координат (2.55) называется *невырожденным* в некоторой точке или области, если якобиан преобразования отличен от нуля в этой точке или области, соответственно. Если преобразование координат невырождено в некоторой области $\mathbb{U}$, то согласно теореме 2.5.1 в достаточно малой окрестности произвольной точки $x \in \mathbb{U}$ это преобразование является взаимно однозначным. Для невырожденных преобразований координат локально существует обратное преобразование

$$x^{\alpha'} \mapsto x^\alpha = x^\alpha(x'),$$

при этом якобиан обратного преобразования равен

$$J^{-1} = \mathsf{det}\left(\frac{\partial x^\alpha}{\partial x^{\alpha'}}\right).$$

Если якобиан преобразования обращается в нуль в некоторой точке, то в этой точке преобразование координат является вырожденным, и обратного преобразования не существует.

Совершим два преобразования координат $x \mapsto x' \mapsto x''$. Тогда координаты точки x можно выразить, как в системе координат x', так и в x''. При этом, по-определению, справедливы равенства $x^\alpha(x'') = x^\alpha\Big(x'(x'')\Big)$. Отсюда следует, что каждая координатная функция x^α при преобразовании координат $x' \to x''$ ведет себя, как скалярное поле (см. раздел 3.5). Верно также и обратное утверждение. Если в некоторой области евклидова пространства $\mathbb{U} \subset \mathbb{R}^n$ задано n функций (скалярных полей) $f^a(x)$, $a = 1, \ldots, n$, таких, что $\mathsf{det}\,(\partial f^a/\partial x^\alpha) \neq 0$, то их всегда можно выбрать в качестве новой системы координат на $\mathbb{U}$.

Обозначим множество невырожденных преобразований координат в $\mathbb{R}^n$ через $\mathbb{G}(\mathbb{R}^n)$. При этом будем рассматривать преобразования координат как активные, т.е. отображающие области в одном фиксированном евклидовом пространстве. Для заданного преобразования координат $f = x'(x)$ будем писать $f \in \mathbb{G}(\mathbb{R}^n)$. Эти преобразования согласованы с топологией в $\mathbb{R}^n$.

Определение. Множество преобразований $\mathbb{G}(\mathbb{R}^n)$ топологического пространства $\mathbb{R}^n$ называется *псевдогруппой преобразований*, если выполнены следующие условия:

1) каждое преобразование $f \in \mathbb{G}(\mathbb{R}^n)$ является гомеоморфизмом открытого множества $\mathbb{U}_i$ (область определения f) на другое открытое множество $f(\mathbb{U}_i) = \mathbb{U}_j$ (область значений f);
2) если $f \in \mathbb{G}(\mathbb{R}^n)$, то сужение f на произвольное открытое подмножество области определения f также принадлежит $\mathbb{G}(\mathbb{R}^n)$;
3) пусть множество $\mathbb{U}$ есть объединение произвольного числа открытых множеств, $\mathbb{U} = \bigcup_i \mathbb{U}_i$, тогда гомеоморфизм f множества $\mathbb{U}$ принадлежит $\mathbb{G}(\mathbb{R}^n)$, если сужение f на $\mathbb{U}_i$ принадлежит $\mathbb{G}(\mathbb{R}^n)$ для каждого i;
4) тождественное преобразование каждого открытого множества $\mathbb{U} \subset \mathbb{R}^n$ принадлежит $\mathbb{G}(\mathbb{R}^n)$;
5) для каждого $f \in \mathbb{G}(\mathbb{R}^n)$ существует обратное преобразование $f^{-1} \in \mathbb{G}(\mathbb{R}^n)$;
6) если f_1 есть гомеоморфизм $\mathbb{U}_1 \to \mathbb{V}_1$, а f_2 – гомеоморфизм $\mathbb{U}_2 \to \mathbb{V}_2$, и пересечение $\mathbb{V}_1 \cap \mathbb{U}_2$ не пусто, то гомеоморфизм $f_2 f_1$ множества $f_1^{-1}(\mathbb{V}_1 \cap \mathbb{U}_2)$ на множество $f_2(\mathbb{V}_1 \cap \mathbb{U}_2)$ также принадлежит $\mathbb{G}(\mathbb{R}^n)$. □

Замечание. Первые три свойства являются топологическими, а остальные – групповыми, если под композицией преобразований понимать их последовательное выполнение в тех точках, где они определены. В общем случае преобразования координат группы не образуют, т.к. при композиции двух преобразований необходимо следить за их областью определений и значений. Например, если область значений преобразования f_1 не имеет общих точек с областью определения f_2, то композиция $f_2 \circ f_1$ не определена. Следовательно, преобразования координат группы не образуют, так как закон композиции в группе должен быть определен для всех элементов группы. Тем не менее часто употребляется термин группа преобразований координат. Для того, чтобы преобразования координат действительно образовали группу необходимо предположить, что области определений и значений $\mathbb{U} \subset \mathbb{R}^n$ для всех $f \in \mathbb{G}(\mathbb{R}^n)$ совпадают и отказаться от условия 2). Эта группа бесконечномерна, т.к. параметризуется n функциями от n переменных. □

Отображение f областей евклидова пространства называется *диффеоморфизмом*. Если все диффеоморфизмы $f \in \mathbb{G}(\mathbb{R}^n)$ одного класса гладкости $\mathcal{C}^k$, то будем писать $f \in \mathsf{diff}^{\,k}(\mathbb{R}^n)$. Для гладкой псевдогруппы преобразований координат индекс обычно опускают $\mathsf{diff}\,(\mathbb{R}^n) := \mathsf{diff}^{\,\infty}(\mathbb{R}^n)$. Ее часто называют просто *группой диффеоморфизмов* евклидова пространства $\mathbb{R}^n$. Очевидно, что $\mathsf{diff}^{\,l}(\mathbb{R}^n)$ – подпсевдогруппа $\mathsf{diff}^{\,k}(\mathbb{R}^n)$ при $l > k$. Множество преобразований координат с положительным определителем $\mathsf{diff}^{\,k}_0(\mathbb{R}^n)$ также является подпсевдогруппой: $\mathsf{diff}^{\,k}_0(\mathbb{R}^n) \subset \mathsf{diff}^{\,k}(\mathbb{R}^n)$. Все невырожденные преобразования координат можно разделить на два класса: с положительным, $\mathsf{diff}^{\,k}_0(\mathbb{R}^n)$, и отрицательным, $\mathsf{diff}^{\,k}(\mathbb{R}^n) \setminus \mathsf{diff}^{\,k}_0(\mathbb{R}^n)$, якобианом. Говорят, что координатный базис касательного пространства (см. раздел 3.6) имеет одинаковую *ориентацию* в двух системах координат, если якобиан соответствующего преобразования положителен. Если якобиан преобразования отрицателен, то при переходе между системами координат базис касательного пространства, по-определению, меняет ориентацию.

2.6 Группа двумерных вращений $\mathbb{O}(2)$

Рассмотрим евклидову плоскость $\mathbb{R}^2$ с декартовыми координатами $\{x, y\} = \{x^\alpha\}$, $\alpha = 1, 2$, и евклидовой метрикой $g_{\alpha\beta} = \delta_{\alpha\beta}$.

Определение. Невырожденные линейные однородные преобразования координат,

$$x^\alpha \mapsto x'^\alpha = x^\beta S_\beta{}^\alpha, \qquad \det S_\beta{}^\alpha \neq 0,$$

относительно которых евклидова метрика остается инвариантной,

$$\delta_{\alpha\beta} = S_\alpha{}^\gamma S_\beta{}^\epsilon \delta_{\gamma\epsilon}, \tag{2.59}$$

образуют группу, которая называется *группой двумерных вращений* и обозначается $\mathbb{O}(2)$. Квадратные матрицы, удовлетворяющие условию (2.59), которое можно переписать в матричном виде

$$SS^T = S^T S = 1,$$

где S^T обозначает транспонированную матрицу, называются *ортогональными.* □

Группа $\mathbb{O}(2)$ является группой ортогональных 2×2 матриц с обычным правилом умножения матриц. Вычисляя определитель левой и правой части уравнения (2.59) получим, что определитель матриц S по модулю равен единице,

$$\det S = \pm 1.$$

Предложение 2.6.1. *Любое решение уравнения (2.59) принадлежит одному из двух классов:*

$$S_+(\omega) := \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix}, \qquad \det S_{(+)} = 1, \tag{2.60}$$

$$S_-(\omega) := \begin{pmatrix} \cos\omega & -\sin\omega \\ -\sin\omega & -\cos\omega \end{pmatrix}, \qquad \det S_{(-)} = -1, \tag{2.61}$$

при некотором угле вращения $\omega \in [0, 2\pi]$. *При этом точки* $\omega = 0$ *и* $\omega = 2\pi$ *отождествляются. Следовательно, параметр* ω *лежит на окружности,* $\omega \in \mathbb{S}^1$.

Доказательство. Уравнение (2.59) – это система трех квадратных уравнений:

$$\begin{aligned} &(S_\alpha{}^1)^2 + (S_\alpha{}^2)^2 = 1, \quad \alpha = 1, 2, \\ &S_1{}^1 S_2{}^1 + S_1{}^2 S_2{}^2 = 0. \end{aligned}$$

Перебор всех возможностей показывает, что с точностью до переопределения угла поворота формулы (2.60), (2.61) дают общее решение. □

Для краткости обозначим множество матриц вида (2.60), (2.61) теми же, но ажурными буквами:

$$\begin{aligned} \mathbb{S}_+ &:= \{S_+(\omega) : \ \ \omega \in \mathbb{S}^1\}, \\ \mathbb{S}_- &:= \{S_-(\omega) : \ \ \omega \in \mathbb{S}^1\}. \end{aligned}$$

Тогда $\mathbb{O}(2) = \mathbb{S}_+ \cup \mathbb{S}_-$, и справедлива следующая таблица умножения:

$$\mathbb{S}_+\mathbb{S}_+ = \mathbb{S}_+, \quad \mathbb{S}_+\mathbb{S}_- = \mathbb{S}_-, \quad \mathbb{S}_-\mathbb{S}_+ = \mathbb{S}_-, \quad \mathbb{S}_-\mathbb{S}_- = \mathbb{S}_+.$$

Обратные матрицы для $\mathbb{S}_+$ имеют вид

$$S_+^{-1}(\omega) = S_+(-\omega) \tag{2.62}$$

и соответствует вращению на тот же угол, но в противоположном направлении. При этом единичная матрица $\mathbb{1} = S_+(0) \in \mathbb{S}_+$.

Таким образом, группа двумерных вращений является однопараметрической компактной группой Ли, состоящей из двух компонент. Матрицы $\mathbb{S}_+$ с единичным определителем образуют нормальную подгруппу группы вращений, которая называется группой *собственных* вращений и обозначается $\mathbb{SO}(2) = \mathbb{S}_+$. Эти матрицы образуют связную компоненту единицы. В этом нетрудно убедиться, заменив параметр ω на $t\omega$ в (2.60). Тогда мы получим непрерывное семейство ортогональных матриц, связывающее заданную матрицу при $t = 1$ с единичной матрицей при $t = 0$.

Отметим, что матрица $-\mathbb{1}$ также принадлежат связной компоненте единицы, $-\mathbb{1} = S_+(\pi/2)$.

Матрицы несобственных вращений S_- группы не образуют, т.к. не содержат, например, единицы. Их можно представить в виде

$$S_- = S_+ P, \tag{2.63}$$

где

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_-(0), \qquad \det P = -1, \tag{2.64}$$

– *оператор четности*, который отражает ось ординат,

$$P: \quad (x, y) \;\mapsto\; (x, -y).$$

При этом меняется ориентация декартовых координат.

Таким образом, матрицы с положительным и отрицательным определителем образуют две связные компоненты группы вращений $\mathbb{O}(2)$. Связная компонента единицы образует группу собственных вращений $\mathbb{SO}(2)$, а оператор четности определяет диффеоморфизм между компонентами.

Легко проверить, что группа собственных вращений $\mathbb{SO}(2)$ является абелевой. В то же время полная группа вращений $\mathbb{O}(2)$ неабелева. В этом просто убедиться, вычислив коммутатор

$$\Big[S_-(\omega_1), S_-(\omega_2)\Big] = 2 \begin{pmatrix} 0 & \sin(\omega_1 - \omega_2) \\ -\sin(\omega_1 - \omega_2) & 0 \end{pmatrix},$$

который не лежит ни в $\mathbb{S}_+$, ни в $\mathbb{S}_-$.

Элементы группы $\mathbb{SO}(2)$ можно представить в виде

$$S_{+\alpha}{}^{\beta} = (e^{\omega\varepsilon})_{\alpha}{}^{\beta} = \delta_{\alpha}^{\beta} \cos\omega + \varepsilon_{\alpha}{}^{\beta} \sin\omega, \tag{2.65}$$

где экспонента определена с помощью ряда Тейлора, и $\varepsilon_{\alpha\beta} = \varepsilon_{\alpha}{}^{\beta}$ – полностью антисимметричный тензор второго ранга (см. приложение 28.4).

Подгруппа $\mathbb{SO}(2) \subset \mathbb{O}(2)$ является нормальной подгруппой. Следовательно, фактор пространство $\mathbb{O}(2)/\mathbb{SO}(2)$ является группой, состоящей из двух элементов

$$\mathbb{Z}_2 \simeq \frac{\mathbb{O}(2)}{\mathbb{SO}(2)},$$

где $\mathbb{Z}_2 := \{1, -1\}$ – группа из двух элементов по умножению.

Матрицы (2.60), (2.61) можно рассматривать как представление группы вращений $\mathbb{O}(2)$ в двумерном векторном пространстве. Оно неприводимо и называется *фундаментальным* или *векторным* представлением.

Бесконечно малое вращение на угол $\delta\omega \ll 1$ в линейном приближении имеет вид

$$S_{+\alpha}{}^{\beta} \approx \delta_{\alpha}^{\beta} + \delta\omega L_{\alpha}{}^{\beta}, \tag{2.66}$$

где $L_{\alpha}{}^{\beta} = \varepsilon_{\alpha}{}^{\beta}$ – генератор (вектор базиса алгебры Ли) группы $\mathbb{SO}(2)$ в фундаментальном представлении.

Вращения евклидовой плоскости показаны на рис. 2.6,*a*. Орбитами действия и

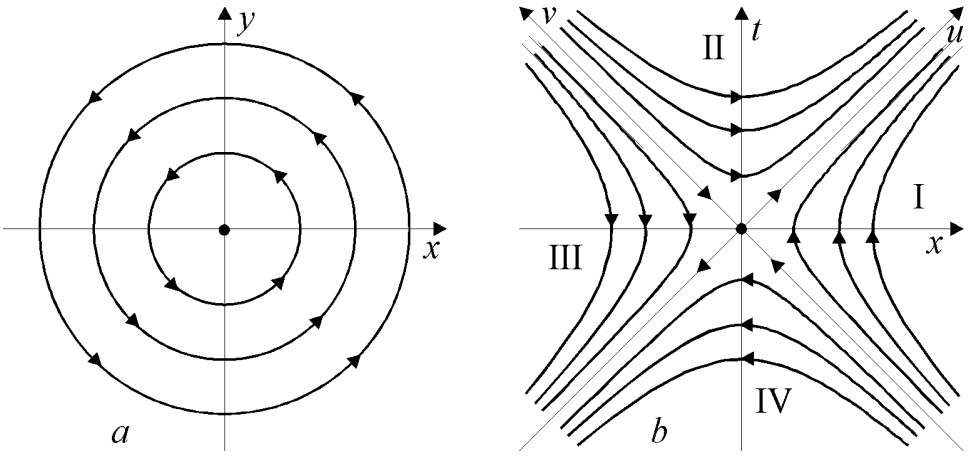


Рис. 2.6: Траектории Киллинга = орбиты действия групп $\mathbb{SO}(2)$ и $\mathbb{SO}_{\uparrow}(1,1)$ на евклидовой плоскости *(a)* и плоскости Минковского *(b)*.

связной компоненты единицы группы $\mathbb{SO}(2)$, и полной группы $\mathbb{O}(2)$ являются концентрические окружности,

$$x^2 + y^2 = r^2, \qquad r \in \mathbb{R}_+, \tag{2.67}$$

с центром в начале координат. Само начало координат является неподвижной точкой относительно действия группы вращений.

Определим *полярные координаты* на евклидовой плоскости

$$x = r\cos\varphi, \qquad y = r\sin\varphi, \tag{2.68}$$

где радиус точки r и полярный угол φ определены в интервалах: $r \in (0,\infty)$, $\varphi \in [0, 2\pi)$. Обратное преобразование координат имеет вид

$$r = \sqrt{x^2+y^2}, \qquad \varphi = \operatorname{arctg}\left(\frac{y}{x}\right). \tag{2.69}$$

Замечание. Эти координаты особенно удобны в задачах, обладающих вращательной симметрией, поскольку при вращении меняется только полярный угол, а радиус точки остается прежним. Например, для собственных вращений

$$S_+(\omega): \quad (r,\varphi) \;\mapsto\; (r,\varphi+\omega). \quad \square \tag{2.70}$$

Евклидов интервал в полярных координатах имеет нетривиальный вид

$$ds^2 = dr^2 + r^2 d\varphi^2. \tag{2.71}$$

Недостатком полярных координат является то, что они вырождены в начале координат, поскольку для этой точки полярный угол (2.69) не определен. Соответственно, якобиан преобразования (2.68),

$$J = \frac{\partial(x,y)}{\partial(r,\varphi)} = r,$$

равен нулю в начале координат. Во всех остальных точках преобразование координат является взаимно однозначным и сохраняет ориентацию, поскольку якобиан положителен.

Если исключить (выколоть) сингулярную точку, которой является начало координат, то евклидова плоскость в полярных координатах представляет собой прямое произведение окружности $\mathbb{S}^1$ на полубесконечный интервал $\mathbb{R}_+ = (0, \infty)$,

$$\mathbb{R}^2 \setminus \{0\} = \mathbb{S}^1 \times \mathbb{R}_+.$$

Рассмотрим два ненулевых вектора в начале координат, $X_1 = (x_1, y_1) = (r_1, \varphi_1)$ и $X_2 = (x_2, y_2) = (r_2, \varphi_2)$. Их длины равны $\|X_1\| = r_1$ и $\|X_2\| = r_2$. Тогда *угол между векторами* φ_{12}, по-определению, равен

$$\mathsf{cos}\,\varphi_{12} := \frac{(X_1, X_2)}{\|X_1\|\|X_2\|}, \tag{2.72}$$

где круглые скобки обозначают скалярное произведение векторов $(X_1, X_2) = x_1x_2 + y_1y_2$. Это соотношение определяет угол между векторами с точностью до знака, т.к. cos является четной функцией. Чтобы устранить возникшую неоднозначность, будем считать угол φ_{12} между векторами X_1 и X_2 положительным, если вращение от X_1 к X_2 происходит против часовой стрелки. По-определению, вращение от оси абсцисс к оси ординат по меньшему углу называется *вращением против часовой стрелки*. При вращении по часовой стрелке угол считается отрицательным. Тогда

$$\varphi_{12} = \varphi_2 - \varphi_1. \tag{2.73}$$

Отсюда следует, что углы между векторами складываются. То есть, если углы между векторами X_1, X_2 и X_2, X_3 равны, соответственно, φ_{12} и φ_{23}, то угол между векторами X_1 и X_3 равен сумме углов

$$\varphi_{13} = \varphi_{12} + \varphi_{23}. \tag{2.74}$$

Это правило сложения углов эквивалентно абелевости группы собственных вращений $\mathbb{SO}(2)$.

Рассмотрим бесконечно малые вращения евклидовой плоскости на угол $|\delta\omega| \ll 1$. Разлагая матрицу (2.60) в ряд, в линейном приближении получим следующие приращения координат:

$$\begin{aligned} \delta x &= -y\delta\omega, \\ \delta y &= x\delta\omega. \end{aligned} \tag{2.75}$$

Забегая вперед, скажем, что, если евклидова плоскость рассматривается, как многообразие, то генератором вращений является векторное поле Киллинга (см. главу 16)

$$K = -y\partial_x + x\partial_y.$$

Изменение формы функции $f(x, y)$ (см. раздел 3.13) в линейном приближении можно выразить через вектор Киллинга:

$$\delta f(x, y) := -\delta x^\alpha \partial_\alpha f = -\delta\omega K f|_{(x,y)},$$

где векторное поле действует на функцию как дифференцирование. В полярных координатах, генератор вращений (векторное поле Киллинга) есть дифференцирование по полярному углу

$$K = \partial_\varphi.$$

При этом интегральные кривые для генератора вращений (траектории Киллинга) совпадают с орбитами действия группы вращений и являются концентрическими окружностями (2.67) с центром в начале координат, как показано на рис. 2.6, *a*.

В комплексных координатах на евклидовой плоскости,

$$z := x + iy = r\,\mathsf{e}^{i\varphi}, \qquad \bar{z} := x - iy = r\,\mathsf{e}^{-i\varphi}, \tag{2.76}$$

где черта обозначает комплексное сопряжение, собственные и несобственные вращения на угол ω выглядят соответственно

$$\begin{aligned} S_+(\omega):\quad & z \mapsto z' = z\,\mathsf{e}^{i\omega}, \qquad & S_-(\omega):\quad & z \mapsto z' = \bar{z}\,\mathsf{e}^{-i\omega}, \\ & \bar{z} \mapsto z' = \bar{z}\,\mathsf{e}^{-i\omega}, & & \bar{z} \mapsto z' = z\,\mathsf{e}^{i\omega}. \end{aligned} \tag{2.77}$$

Отсюда следует, что связная компонента единицы $\mathbb{SO}(2)$ изоморфна группе $\mathbb{U}(1)$ комплексных чисел, равных по модулю единице, с обычным правилом умножения.

Приведем для справки ряд формул. Интервал на евклидовой плоскости в комплексных координатах (2.76) имеет вид

$$ds^2 = dz d\bar{z}.$$

Операторы дифференцирования по декартовым и комплексным координатам связаны между собой соотношениями:

$$\begin{aligned} \partial_z &= \frac{1}{2}(\partial_x - i\partial_y), \qquad & \partial_{\bar{z}} &= \frac{1}{2}(\partial_x + i\partial_y), \\ \partial_x &= \partial_z + \partial_{\bar{z}}, & \partial_y &= i(\partial_z - \partial_{\bar{z}}). \end{aligned} \tag{2.78}$$

Отсюда следует выражение для лапласиана

$$\triangle = \partial^2_{xx} + \partial^2_{yy} = 4\partial^2_{z\bar{z}}. \tag{2.79}$$

Из формул преобразования комплексных координат (2.78) сразу следует, что оператор Лапласа инвариантен относительно полной группы вращений $\mathbb{O}(2)$.

2.7 Группа двумерных преобразований Лоренца $\mathbb{O}(1,1)$

Определение. *Плоскостью Минковского* $\mathbb{R}^{1,1}$ называется евклидова плоскость с декартовыми координатами $\{x^\alpha\} = \{t, x\}$, $\alpha = 0, 1$, в которых задана двумерная метрика Лоренца

$$\eta = \{\eta_{\alpha\beta}\} := \mathsf{diag}\,(+-) := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.80}$$

В пространстве Минковского декартова система координат, в которой задана метрика (2.80), называется *инерциальной* системой отсчета. Метрике Лоренца соответствует интервал

$$ds^2 = dx^\alpha dx^\beta \eta_{\alpha\beta} = dt^2 - dx^2 = du dv, \tag{2.81}$$

где введены *координаты светового конуса* (см. рис. 2.6, *b*):

$$u = t + x, \quad v = t - x, \tag{2.82}$$

которые являются аналогом комплексных координат (2.76) на евклидовой плоскости. Координатные оси t и x называют соответственно *временем* и *пространством*. □

Якобиан преобразования координат (2.82) равен

$$J = \frac{\partial(u,v)}{\partial(x,t)} = 2.$$

Это преобразование сохраняет ориентацию, если координаты упорядочены следующим образом: $x, t \mapsto u, v$. В координатах светового конуса линии, параллельные осям координат, $u = \mathsf{const}$ и $v = \mathsf{const}$, имеют нулевую длину, т.е. светоподобны.

Замечание. На плоскости Минковского $\mathbb{R}^{1,1}$ время и пространство играют равноправную роль, поскольку пространство одномерно, и переобозначение $t \leftrightarrow x$ приводит только к изменению общего знака метрики. □

Подчеркнем, что на плоскости Минковского $\mathbb{R}^{1,1}$ задано две метрики: евклидова метрика и метрика Минковского (2.80). Первая необходима для задания топологии и определения дифференцируемых в обычном смысле функций на $\mathbb{R}^{1,1}$.

Предложение 2.7.1. *На плоскости Минковского любая изотропная кривая* $\{t(s), x(s)\}$, *т.е. линия, касательный вектор к которой имеет нулевую длину:*

$$\left(\frac{dt}{ds}\right)^2 - \left(\frac{dx}{ds}\right)^2 = 0, \tag{2.83}$$

является прямой линией вида $t = \pm x + \mathsf{const}$.

Доказательство. Из уравнения (2.83) следует, что $dt = \pm dx$. Общим решением данного уравнения являются прямые линии $t = \pm x + \mathsf{const}$. □

Определение. Уравнение

$$(t - t_0)^2 - (x - x_0)^2 = 0$$

определяет на плоскости Минковского $\mathbb{R}^{1,1}$ две перпендикулярные прямые, пересекающиеся в точке (t_0, x_0). Они называются *световым конусом* в данной точке, рис. 2.7. Световой конус состоит из двух связных компонент: светового конуса *прошлого*, $t \le t_0$, и *будущего*, $t \ge t_0$. Световые конусы прошлого и будущего имеют общую вершину в точке (t_0, x_0). □

Любая времениподобная кривая $\{x^\alpha(s)\}$,

$$\dot{x}^\alpha \dot{x}^\beta \eta_{\alpha\beta} > 0, \qquad \dot{x}^\alpha := \frac{dx^\alpha}{ds},$$

проходящая через точку (t_0, x_0), будет целиком лежать внутри светового конуса (заштрихованная область на рисунке).

Определение. Невырожденные линейные однородные преобразования координат

$$x^\alpha \mapsto x'^\alpha = x^\beta S_\beta{}^\alpha, \qquad \det S_\beta{}^\alpha \neq 0, \tag{2.84}$$

относительно которых метрика Минковского (2.80) остается инвариантной,

$$\eta_{\alpha\beta} = S_\alpha{}^\gamma S_\beta{}^\epsilon \eta_{\gamma\epsilon}, \tag{2.85}$$

образуют группу, которая называется *группой двумерных преобразований Лоренца* и обозначается $\mathbb{O}(1,1)$. □

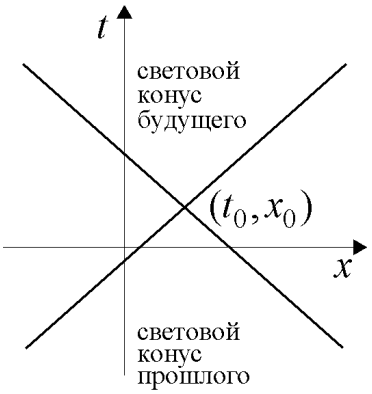


Рис. 2.7: Световой конус на плоскости Минковского.

Из уравнения (2.85) следует, что определитель матриц S, как и в случае вращений евклидовой плоскости, по модулю равен единице, $\det S = \pm 1$.

Предложение 2.7.2. *Любое решение уравнения (2.85) принадлежит одному из четырех классов:*

$$S_{+\uparrow}(\omega) := \begin{pmatrix} \operatorname{ch}\omega & \operatorname{sh}\omega \\ \operatorname{sh}\omega & \operatorname{ch}\omega \end{pmatrix}, \qquad \det S_{+\uparrow} = 1, \tag{2.86}$$

$$S_{-\uparrow}(\omega) := \begin{pmatrix} \operatorname{ch}\omega & -\operatorname{sh}\omega \\ \operatorname{sh}\omega & -\operatorname{ch}\omega \end{pmatrix}, \qquad \det S_{-\uparrow} = -1, \tag{2.87}$$

$$S_{+\downarrow}(\omega) := \begin{pmatrix} -\operatorname{ch}\omega & -\operatorname{sh}\omega \\ -\operatorname{sh}\omega & -\operatorname{ch}\omega \end{pmatrix}, \qquad \det S_{+\downarrow} = 1, \tag{2.88}$$

$$S_{-\downarrow}(\omega) := \begin{pmatrix} -\operatorname{ch}\omega & \operatorname{sh}\omega \\ -\operatorname{sh}\omega & \operatorname{ch}\omega \end{pmatrix}, \qquad \det S_{-\downarrow} = -1, \tag{2.89}$$

при некотором $\omega \in \mathbb{R}$. Параметр ω называется гиперболическим *углом вращения.*

Доказательство. Уравнение (2.80) – это система трех квадратных уравнений:

$$\begin{aligned} (S_1{}^1)^2 - (S_1{}^2)^2 &= 1, \\ (S_2{}^1)^2 - (S_2{}^2)^2 &= -1, \\ S_1{}^1 S_2{}^1 - S_1{}^2 S_2{}^2 &= 0. \end{aligned}$$

Перебор всех возможных решений показывает, что с точностью до переопределения угла поворота формулы (2.86)–(2.89) дают общее решение. □

Индекс $\pm$ у матрицы преобразований Лоренца соответствует знаку определителя, а смысл стрелки будет ясен из дальнейшего рассмотрения.

Из доказанного утверждения следует, что двумерная группа Лоренца $\mathbb{O}(1,1)$ является однопараметрической некомпактной группой Ли и состоит из четырех несвязных компонент.

Заметим, что

$$S^T_{-\uparrow}(\omega) = S_{-\uparrow}(-\omega) \quad \text{и} \quad S^T_{-\downarrow}(\omega) = S_{-\downarrow}(-\omega).$$

Как и в случае группы вращений евклидовой плоскости $\mathbb{O}(2)$, обозначим множества матриц (2.86)–(2.89) теми же, но ажурными буквами. Например,

$$\mathbb{S}_{+\uparrow} := \{S_{+\uparrow}(\omega) : \quad \omega \in \mathbb{R}\}.$$

Тогда выполнена следующая таблица умножения

	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$
$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$
$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{-\downarrow}$
$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{+\uparrow}$	$\mathbb{S}_{-\uparrow}$
$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{+\downarrow}$	$\mathbb{S}_{-\downarrow}$	$\mathbb{S}_{-\uparrow}$	$\mathbb{S}_{+\uparrow}$

(2.90)

Обратные матрицы из $\mathbb{S}_{+\uparrow}$ имеют вид (2.62), т.е. получаются заменой $\omega \mapsto -\omega$.

Единичная матрица $\mathbb{1} = S_{+\uparrow}(0)$ и матрица $-\mathbb{1} = S_{+\downarrow}(0)$ принадлежат разным компонентам группы.

Определение. Введем оператор *обращения времени* и *пространственного отражения (четности)*

$$\begin{aligned} T &:= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = S_{-\downarrow}(0), \qquad \det T = -1, \\ P &:= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_{-\uparrow}(0), \qquad \det P = -1. \end{aligned} \tag{2.91}$$

Действие этих операторов мы записываем в виде $x^\alpha \mapsto x^\beta T_\beta{}^\alpha$ и $x^\alpha \mapsto x^\beta P_\beta{}^\alpha$. Они отражают соответственно временну́ю и пространственную координаты:

$$\begin{aligned} T: &\quad (t,x) \;\mapsto\; (-t,x), \\ P: &\quad (t,x) \;\mapsto\; (t,-x). \end{aligned}$$

Введем также обозначение для их композиции

$$R := PT = TP = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = S_{+\uparrow}(0), \qquad \det R = 1, \tag{2.92}$$

которая соответствует полному отражению координат $R: \ (t,x) \mapsto (-t,-x)$. □

Тогда матрицы преобразований Лоренца (2.87)–(2.89) можно представить в виде

$$\begin{aligned} S_{-\uparrow} &= S_{+\uparrow} P, \\ S_{-\downarrow} &= S_{+\uparrow} T, \\ S_{+\downarrow} &= S_{+\uparrow} R. \end{aligned} \tag{2.93}$$

Преобразования (2.86) и (2.87) не меняют направления времени и называются *ортохронными*, что отмечено направленной вверх стрелкой. Эти преобразования образуют компоненты $S_{+\uparrow}$ и $S_{-\uparrow}$, для которых ${S_0}^0 > 0$. Матрицы из компонент $S_{+\downarrow}$ и $S_{-\downarrow}$ не являются ортохронными, и для них ${S_0}^0 < 0$.

Из таблицы умножения (2.90) следует, что только одна из четырех связных компонент $\mathbb{S}_{+\uparrow}$ (связная компонента единицы) образует группу. Эта группа обозначается

$$\mathbb{SO}_\uparrow(1,1) \approx \mathbb{S}_{+\uparrow}, \qquad \det S = 1, \quad {S_0}^0 > 0, \tag{2.94}$$

и называется *собственной ортохронной* группой Лоренца. Кроме этого группа Лоренца $\mathbb{O}(1,1)$ имеет еще три нетривиальные подгруппы, состоящие из двух компонент, которые включают отражение времени T, преобразование четности P и полного отражения R. Введем для этих подгрупп следующие обозначения:

$$\begin{aligned} \mathbb{O}_{\downarrow}(1,1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\downarrow}, & \det S &= -1, \quad S_0{}^0 < 0, \\ \mathbb{O}_{\uparrow}(1,1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\uparrow}, & \det S &= -1, \quad S_0{}^0 > 0, \\ \mathbb{SO}(1,1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{+\downarrow}, & \det S &= \quad 1, \quad S_0{}^0 < 0. \end{aligned} \tag{2.95}$$

Матрицы подгрупп $\mathbb{SO}(1,1)$ и $\mathbb{SO}_{\uparrow}(1,1)$ имеют положительный определитель. Легко проверить, что они являются абелевыми. Подгруппы $\mathbb{O}_{\uparrow}(1,1)$ и $\mathbb{O}_{\downarrow}(1,1)$, а также полная группа Лоренца $\mathbb{O}(1,1)$ являются неабелевыми.

Замечание. Знание подгрупп полной группы Лоренца $\mathbb{O}(1,1)$ важно для описания представлений, в том числе спинорных. □

Определение. Если при пространственных отражениях скаляр, вектор и вообще произвольный тензор меняет знак, то их принято называть *псевдоскаляр*, *псевдовектор* и *псевдотензор*. □

Все подгруппы в полной группе Лоренца $\mathbb{O}(1,1)$ являются нормальными. Поэтому определена факторгруппа

$$\mathbb{K}_4 = \frac{\mathbb{O}(1,1)}{\mathbb{SO}_{\uparrow}(1,1)},$$

состоящая из четырех элементов. Из выражений (2.93) следует, что в качестве представителей смежных классов можно выбрать четыре матрицы $\mathbb{1}, P, T, R$. Эти матрицы, как легко проверить, имеют следующую таблицу умножения

	$\mathbb{1}$	P	T	R
$\mathbb{1}$	$\mathbb{1}$	P	T	R
P	P	$\mathbb{1}$	R	T
T	T	R	$\mathbb{1}$	P
R	R	T	P	$\mathbb{1}$

, (2.96)

которая с точностью до обозначений совпадает с (2.90).

Определение. Группа состоящая из четырех элементов, удовлетворяющих таблице умножения (2.96), называется *4-группой Клейна*. □

Замечание. Группа Клейна не является циклической. □

Элементы группы $\mathbb{SO}_{\uparrow}(1,1)$ можно представить в виде

$$S_{+\uparrow\alpha}{}^{\beta} = (e^{-\omega\varepsilon})_{\alpha}{}^{\beta} = \delta_{\alpha}^{\beta}\,\mathsf{ch}\,\omega - \varepsilon_{\alpha}{}^{\beta}\,\mathsf{sh}\,\omega, \tag{2.97}$$

где $\varepsilon_{\alpha\beta} = \varepsilon_{\alpha}{}^{\gamma}\eta_{\gamma\beta}$, $\epsilon_{01} = 1$ – полностью антисимметричный тензор второго ранга в пространстве Минковского (см. приложение 28.4).

В координатах светового конуса (2.82) преобразования из группы Лоренца имеют вид:

$$\begin{aligned} S_{+\uparrow}: &\quad u \mapsto u' = ue^{\omega}, & S_{+\downarrow}: &\quad u \mapsto u' = -ue^{\omega}, \\ &\quad v \mapsto v' = ve^{-\omega}, & &\quad v \mapsto v' = -ve^{-\omega}, \\ S_{-\uparrow}: &\quad u \mapsto u' = ve^{-\omega}, & S_{-\downarrow}: &\quad u \mapsto u' = -ve^{-\omega}, \\ &\quad v \mapsto v' = ue^{\omega}, & &\quad v \mapsto v' = -ue^{\omega}. \end{aligned} \tag{2.98}$$

Матрицы (2.86)–(2.89) можно рассматривать как представление группы Лоренца $\mathbb{O}(1,1)$ в двумерном векторном пространстве. Оно неприводимо и называется *фундаментальным* или *векторным* представлением. Относительно группы собственных ортохронных преобразований $\mathbb{SO}_\uparrow(1,1)$ и группы собственных преобразований $\mathbb{SO}(1,1)$ это представление является приводимым и распадается на два неприводимых одномерных представления (2.98).

Замечание. Все неприводимые представления собственных групп Лоренца $\mathbb{SO}_\uparrow(1,1)$ и $\mathbb{SO}$ являются одномерными, поскольку группы абелевы. Это является причиной того, что выбор конусных координат часто приводит к существенному упрощению вычислений. □

Бесконечно малые собственные ортохронные лоренцевы вращения на угол $\delta\omega \ll 1$ имеют вид

$$S_{+\uparrow}{}^{\beta}_{\alpha} = \delta_\alpha{}^\beta + \delta\omega L_\alpha{}^\beta, \tag{2.99}$$

где $L_\alpha{}^\beta = -\varepsilon_\alpha{}^\beta$ – генераторы группы Лоренца в фундаментальном представлении.

Гиперболические вращения плоскости Минковского при действии собственной ортохронной группы Лоренца $\mathbb{SO}_\uparrow(1,1)$ показаны на рис. 2.6,*b*. При этом плоскость разделяется на четыре квадранта I–IV. Собственные ортохронные преобразования Лоренца преобразуют точки внутри каждого квадранта. Орбитами этих точек являются гиперболы

$$t^2 - x^2 = \pm r^2, \qquad r \in \mathbb{R}_+. \tag{2.100}$$

Здесь знак плюс соответствует II и IV, а минус – I и III квадрантам. Кроме того, орбитами группы $\mathbb{SO}_\uparrow(1,1)$ являются четыре луча $t = \pm x$, выходящих из начала координат $(0,0)$, которое является неподвижной точкой. В каждом из квадрантов можно ввести гиперболическую полярную систему координат:

$$\begin{aligned} \text{I}: \quad & t = r\,\mathsf{sh}\,\varphi, \\ & x = r\,\mathsf{ch}\,\varphi, \\ \text{II}: \quad & t = r\,\mathsf{ch}\,\varphi, \\ & x = r\,\mathsf{sh}\,\varphi, \end{aligned} \qquad \begin{aligned} \text{III}: \quad & t = -r\,\mathsf{sh}\,\varphi, \\ & x = -r\,\mathsf{ch}\,\varphi, \\ \text{IV}: \quad & t = -r\,\mathsf{ch}\,\varphi, \\ & x = -r\,\mathsf{sh}\,\varphi, \end{aligned} \tag{2.101}$$

где $r \in (0,\infty)$, $\varphi \in (-\infty,\infty)$ независимо от квадранта. При этом "началом" координат, $r = 0$, являются две перпендикулярные прямые

$$r = 0 \;\leftrightarrow\; t = \pm x,$$

т.е. световой конус в начале координат. На рис. 2.6,*b* стрелками показано направление возрастания угла φ внутри квадрантов и возрастание гиперболического угла вращения ω на световом конусе.

Обратные преобразования координат, например, во II квадранте имеют вид

$$r = \sqrt{t^2 - x^2}, \qquad \varphi = \mathsf{arcth}\left(\frac{x}{t}\right).$$

Знаки для преобразования координат в различных квадрантах (2.101) подобраны таким образом, чтобы при лоренцевом преобразовании (2.86) полярный угол менялся по закону

$$\varphi \;\mapsto\; \varphi' = \varphi + \omega, \tag{2.102}$$

при этом радиус точки остается прежним. Лоренцев интервал в полярных координатах отличается знаком в разных квадрантах:

$$\begin{aligned} \text{I,III}: &\quad ds^2 = -dr^2 + r^2 d\varphi^2, \\ \text{II,IV}: &\quad ds^2 = dr^2 - r^2 d\varphi^2. \end{aligned} \tag{2.103}$$

Отсюда следует, что радиус точки играет роль пространственной координаты в I и III квадранте и времениподобной координаты во II и IV квадранте. Угол, наоборот, времениподобен в I и III квадранте и пространственноподобен во II и IV квадранте. Якобиан преобразования (2.101) равен

$$\text{I,III}: \quad J = \frac{\partial(x,t)}{\partial(r,\varphi)} = r, \qquad \text{II,IV}: \quad J = \frac{\partial(x,t)}{\partial(r,\varphi)} = -r$$

и вырожден на линиях $r = 0$. Переход к полярным координатам сохраняет ориентацию базиса в I и III квадрантах и меняет ориентацию во II и IV квадрантах.

С топологической точки зрения внутренность каждого квадранта является прямым произведением полубесконечного интервала и прямой

$$\mathsf{int}\,(\mathrm{I}) = \mathbb{R}_+ \times \mathbb{R}.$$

Нетрудно выразить координаты светового конуса через полярные координаты. Например, во втором квадранте

$$u = re^{\varphi}, \qquad v = re^{-\varphi},$$

что является аналогом полярной записи комплексных чисел (2.76).

При пространственном отражении, обращении времени и полном отражении квадранты отображаются друг на друга следующим образом:

P:	$\mathrm{I} \leftrightarrow \mathrm{III}$,	$\mathrm{II} \to \mathrm{II}$,	$\mathrm{IV} \to \mathrm{IV}$,
T:	$\mathrm{I} \to \mathrm{I}$,	$\mathrm{II} \leftrightarrow \mathrm{IV}$,	$\mathrm{III} \to \mathrm{III}$,
R:	$\mathrm{I} \leftrightarrow \mathrm{III}$,	$\mathrm{II} \leftrightarrow \mathrm{IV}$,	

Определим угол φ_{12} между векторами X_1 и X_2 из одного квадранта следующим образом:

$$\mathsf{ch}\,\varphi_{12} := \frac{|(X_1, X_2)|}{|X_1||X_2|}. \tag{2.104}$$

Этот угол определен с точностью до знака. Будем считать угол φ_{12} между векторами $X_1 = \{r_1, \varphi_1\}$ и $X_2 = \{r_2, \varphi_2\}$ положительным, если поворот от X_1 к X_2 происходит в сторону увеличения полярного угла. Тогда

$$\varphi_{12} = \varphi_2 - \varphi_1. \tag{2.105}$$

Отсюда следует, что углы между векторами складываются (2.74), что соответствует абелевости собственной ортохронной группы Лоренца $\mathbb{SO}_{\uparrow}(1,1)$. Подчеркнем, что понятие угла вводится только для векторов из одного квадранта.

Бесконечно малые гиперболические вращения плоскости Минковского $\mathbb{R}^{1,1}$ на угол $\delta\omega \ll 1$ можно получить, разлагая матрицу (2.86) в ряд Тейлора. В линейном приближении получим следующие приращения координат:

$$\delta t = x\delta\omega, \quad \delta x = t\delta\omega. \tag{2.106}$$

Отсюда следует, что генератором лоренцевых преобразований является векторное поле Киллинга

$$K = x\partial_t + t\partial_x.$$

В полярных координатах, т.е. внутри каждого квадранта, генератор вращений есть дифференцирование по полярному углу:

$$K = \partial_\varphi.$$

При этом интегральные кривые для генератора вращений (траектории Киллинга) совпадают с орбитами действия группы и представляют собой гиперболы (2.100), как показано на рис. 2.6,*b*.

В физике собственные ортохронные преобразования Лоренца (2.97) обычно записывают для размерных координат $x^0 = ct$, $x^1 = x$, где c – скорость света. Они имеют вид

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \qquad x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \tag{2.107}$$

где V – скорость движения штрихованной системы координат вдоль оси x. При этом угол гиперболического поворота ω выражается через скорость относительного движения V с помощью соотношения

$$\mathsf{th}\,\omega = -\frac{V}{c}. \tag{2.108}$$

Поскольку $-1 < \mathsf{th} < 1$, то преобразования Лоренца определены для скоростей, меньших скорости света $V < c$. Поэтому из постулата ковариантности законов Природы относительно преобразований Лоренца вытекает, что скорость света является предельной скоростью для частиц и волн. Это предположение является постулатом специальной теории относительности и находится в прекрасном согласии с известными экспериментальными данными.

Преобразования Лоренца (2.107) имеют следующую физическую интерпретацию. Дополним двумерное пространство Минковского до четырехмерного, введя дополнительные декартовы координаты y, z. Предположим, что дополнительные координаты не меняются

$$y' = y, \qquad z' = z.$$

Тогда координаты t, x, y, z и t', x', y', z' являются координатами одного и того же события в двух инерциальных системах отсчета с параллельными осями координат (см. рис.2.8). При этом штрихованная система координат движется равномерно и прямолинейно со скоростью V относительно нештрихованной системы вдоль оси x. Системы координат выбраны таким образом, что их начала совпадают:

$$t = 0,\ x = 0 \quad \Leftrightarrow \quad t' = 0,\ x' = 0.$$

Преобразования Лоренца (2.107) при $c \to \infty$ переходят в преобразования Галилея

$$x' = x - vt, \qquad t' = t. \tag{2.109}$$

Эти преобразования хорошо знакомы из курса классической механики, где время является абсолютным и не зависит от выбора системы координат.

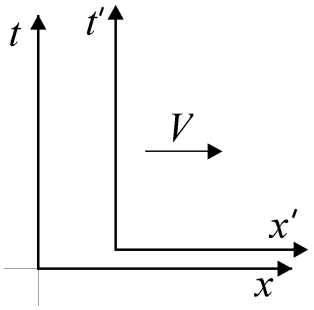


Рис. 2.8: Физическая интерпретация преобразований Лоренца. Штрихованная инерциальная система отсчета движется равномерно и прямолинейно относительно нештрихованной со скоростью V.

Вернемся к двумерному пространству Минковского. Поскольку в приложениях часто используются конусные координаты (2.82), в заключение настоящего раздела приведем для справок некоторые формулы:

$$t = \frac{1}{2}(u+v), \qquad x = \frac{1}{2}(u-v), \tag{2.110}$$

$$\begin{aligned} \partial_u &= \frac{1}{2}(\partial_t + \partial_x), \qquad & \partial_v &= \frac{1}{2}(\partial_t - \partial_x), \\ \partial_t &= \partial_u + \partial_v, & \partial_x &= \partial_u - \partial_v. \end{aligned} \tag{2.111}$$

Оператор Даламбера на плоскости Минковского имеет вид

$$\Box = \partial_{tt}^2 - \partial_{xx}^2 = 4\partial_{uv}^2. \tag{2.112}$$

Он инвариантен относительно действия группы Лоренца $\mathbb{O}(1,1)$.

2.8 Трехмерное евклидово пространство

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами $\{x^i\}=\{x,y,z\}$, $i=1,2,3$.

Замечание. В настоящем разделе для обозначения индексов используются буквы из середины латинского алфавита $i,j,k,\ldots$. Такое изменение обозначений связано с тем, что в дальнейшем координаты в пространстве Минковского $\mathbb{R}^{1,3}$ часто будут обозначаться через $\{x^a\}=\{x^0,x^i\}$, где явно выделено время $x^0=t$, $a=0,1,2,3$ и $i=1,2,3$. В обозначениях мы используем следующее правило. Множество чисел $\{1,2,3\}$ является подмножеством чисел $\{0,1,2,3\}$, так же как и буквы $\{i,j,\ldots\}$ являются подмножеством всего алфавита $a,b,\ldots$. □

Дифференциально геометрическая евклидова метрика в декартовой системе координат задается единичной матрицей

$$ds^2 = \delta_{ij}dx^i dx^j = dx^2 + dy^2 + dz^2. \tag{2.113}$$

Определение. Группа линейных неоднородных преобразований декартовых координат, оставляющая метрику (2.113) инвариантной, называется *неоднородной группой вращений* и обозначается $\mathbb{IO}(3)$. В общем случае преобразование из неоднородной группы вращений имеет вид

$$x^i \mapsto x'^i = x^j S_j{}^i + a^i,$$

где $S_j{}^i \in \mathbb{O}(3)$ – вещественная ортогональная матрица, определяемая условием

$$S^T S = S S^T = 1, \tag{2.114}$$

и $a^i \in \mathbb{R}^3$ – произвольный вектор. Матрицы $\mathbb{O}(3)$ образуют группу, которая называется *группой трехмерных вращений*. Абелева подгруппа в $\mathbb{IO}(3)$, параметризуемая вектором a, называется *группой трансляций*. □

Поскольку группа вращений сохраняет метрику (2.113), то при поворотах и трансляциях евклидова пространства сохраняются длины и углы между векторами. Конечно, трансляции можно было бы ввести и ранее в двумерном случае.

Можно показать, что любое преобразование декартовых координат, оставляющее метрику (2.113) инвариантной, является линейным. Поэтому группа $\mathbb{IO}(3)$ является максимальной группой преобразований симметрии евклидова пространства $\mathbb{R}^3$ с заданной метрикой. Группа неоднородных вращений, как будет показано в дальнейшем, является группой Ли.

Группа $\mathbb{IO}(3)$ представляет собой полупрямое произведение группы вращений $\mathbb{O}(3)$ вокруг начала координат на абелеву группу трансляций. Группа трансляций действует в $\mathbb{R}^3$ свободно и транзитивно. Как многообразие она диффеоморфна евклидову пространству $\mathbb{R}^3$. Группа вращений $\mathbb{O}(3)$ действует в $\mathbb{R}^3$ эффективно. При этом начало координат является неподвижной точкой. Многообразие группы вращений будет описано немного позже.

Из условия ортогональности (2.114) следует, что определитель ортогональной матрицы равен по модулю единице

$$\mathsf{det}\, S = \pm 1.$$

Группа трехмерных вращений $\mathbb{O}(3)$ является неабелевой и состоит из двух компонент. Связная компонента единицы обозначается $\mathbb{SO}(3)$ и называется группой *собственных* вращений. К ней относятся ортогональные матрицы с единичным определителем. Вторая связная компонента состоит из ортогональных матриц с отрицательным определителем и сама по себе группы не образует.

Определение. *Оператором пространственных отражений* называется матрица

$$P := \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{или} \quad P_j{}^i = -\delta^i_j, \tag{2.115}$$

при действии которой все координаты меняют знак $x^i \mapsto x^j P_j{}^i = -x^i$. □

Произвольную ортогональную матрицу с отрицательным определителем можно однозначно представить в виде произведения собственно ортогональной матрицы и оператора пространственных отражений. Это значит, что множество ортогональных матриц с отрицательным определителем представляет собой смежный класс матрицы отражений по подгруппе $\mathbb{SO}(3)$.

Связная компонента единицы $\mathbb{SO}(3)$ является нормальной подгруппой в группе вращений $\mathbb{O}(3)$. Поэтому определена фактор группа

$$\mathbb{Z}_2 = \frac{\mathbb{O}(3)}{\mathbb{SO}(3)},$$

состоящая из двух элементов $\mathbb{Z}_2 = \{\mathbb{1}, P\}$. Отображение

$$\mathbb{O}(3) \ni S \mapsto \det S \in \mathbb{Z}_2$$

есть гомоморфизм группы вращений на мультипликативную группу $\mathbb{Z}_2$. Как многообразие полная группа вращений гомеоморфна прямому произведению: $\mathbb{O}(3) \approx \mathbb{Z}_2 \times \mathbb{SO}(3)$.

Группа $\mathbb{SO}(3)$ проста (т.е. не содержит инвариантных абелевых подгрупп), и, значит, нетривиальные гомоморфизмы на другие группы Ли невозможны.

При вращениях произвольная точка $x \in \mathbb{R}^3$, не совпадающая с началом координат, пробегает все точки сферы $\mathbb{S}^2$ с центром в начале координат и содержащей точку $x \in \mathbb{S}^2$. При этом сфера является орбитой точки $x \in \mathbb{R}^3$ при действии групп $\mathbb{SO}(3)$ и $\mathbb{O}(3)$. Отметим, что две точки сферы $x_1, x_2 \in \mathbb{S}^2$ не определяют вращение однозначно (см. рис. 2.9). Действительно, пусть точки x_1 и x_2 лежат в плоскости x, z и симметричны относительно оси z. Тогда точку x_2 можно получить из точки x_1 вращением вокруг оси z на угол $\pm\pi$. Этого же можно добиться вращением вокруг оси y на меньший угол, причем траекторией точки x_1 будет часть большой окружности, соединяющей точки x_1 и x_2. В общем случае осью вращения может служить любой единичный вектор, лежащий в плоскости, перпендикулярной отрезку, соединяющему точки x_1 и x_2 и проходящей через начало координат. В рассматриваемом случае ось вращения должна лежать в плоскости y, z. Если на сфере задать перемещение не двух, а всех точек одновременно, тогда вращение будет определено однозначно.

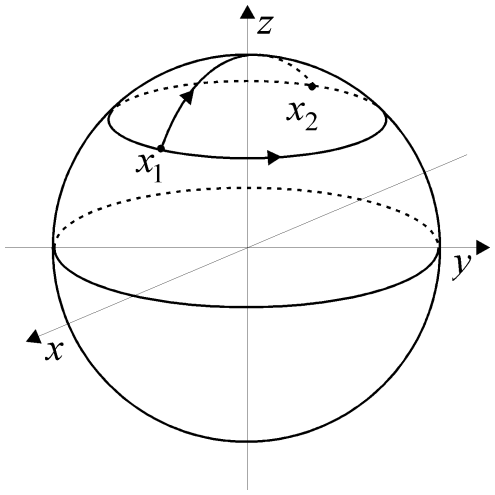


Рис. 2.9: Отображение двух точек на сфере $x_1 \mapsto x_2$ не определяет трехмерное вращение однозначно.

Генераторы группы вращений $\mathbb{SO}(3)$ могут быть представлены в евклидовом пространстве в виде дифференциальных операторов (векторных полей Киллинга)

$$L_{ij} = x_i\partial_j - x_j\partial_i. \tag{2.116}$$

Прямое вычисление коммутатора приводит к следующей алгебре Ли

$$[L_{ij}, L_{kl}] = -\delta_{ik}L_{jl} + \delta_{il}L_{jk} + \delta_{jk}L_{il} - \delta_{jl}L_{ik}. \tag{2.117}$$

Отсюда следует, что группа трехмерных вращений $\mathbb{SO}(3)$ является неабелевой группой Ли. Эта группа трехмерна, т.к. ввиду антисимметрии по индексам, существуют

только три независимых вектора: L_{12}, L_{13} и L_{23}. Алгебру Ли (2.117) можно переписать в более компактном виде для эквивалентного набора генераторов

$$J_i = \frac{1}{2}\varepsilon_{ijk}L^{jk}, \qquad L_{ij} = \varepsilon_{ijk}J^k, \tag{2.118}$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга, $\varepsilon_{123} = 1$, (см. Приложение 28.4), а подъем и опускание индексов производится с помощью евклидовой метрики. Тогда алгебра Ли группы $\mathbb{SO}(3)$ принимает вид

$$[J_i, J_j] = -\varepsilon_{ijk}J^k. \tag{2.119}$$

Алгебры Ли групп Ли $\mathbb{SO}(3)$ и $\mathbb{O}(3)$, естественно, совпадают.

При проведении расчетов с группой $\mathbb{O}(3)$ удобно использовать явную параметризацию элементов группы элементами ее алгебры. Элемент алгебры $\mathfrak{so}(3)$ может быть представлен в виде произвольной антисимметричной 3×3-матрицы

$$(\omega\varepsilon)_i{}^j = (\omega^k\varepsilon_k)_i{}^j = \omega^k\varepsilon_{ki}{}^j \quad \in \mathfrak{so}(3), \tag{2.120}$$

где $\boldsymbol{\omega} = \{\omega^k\} \in \mathbb{R}^3$ – *вектор вращения*. Здесь первый индекс k нумерует базис алгебры, а индексы i, j рассматриваются, как матричные. Для сравнения с (2.119) заметим, что полностью антисимметричный тензор третьего ранга можно рассматривать в качестве базиса алгебры $\mathfrak{so}(3)$:

$$(J_k)_i{}^j := \varepsilon_{ki}{}^j.$$

Элемент алгебры параметризуется трехмерным вектором $\boldsymbol{\omega} \in \mathbb{R}^3$, и, значит, группа $\mathbb{O}(3)$ является трехмерной группой Ли.

Предложение 2.8.1. *Каждому элементу алгебры Ли (вектору вращения)* $\boldsymbol{\omega} = \{\omega^i\} \in \mathfrak{so}(3)$ *соответствует элемент группы Ли* $S = \{S_i{}^j\} \in \mathbb{SO}(3)$ *(связной компоненты единицы):*

$$S_i{}^j(\boldsymbol{\omega}) = (\,\mathsf{e}^{(\omega\varepsilon)})_i{}^j = \delta_i^j\cos\omega + \frac{(\omega\varepsilon)_i{}^j}{\omega}\sin\omega + \frac{\omega_i\omega^j}{\omega^2}(1 - \cos\omega) \;\in \mathbb{SO}(3), \tag{2.121}$$

где $\omega = \sqrt{\omega^i\omega_i}$ *– модуль вектора вращения* $\boldsymbol{\omega}$.

Доказательство. Последнее равенство в (2.121) доказывается путем разложения обеих частей равенства в ряды, которые равномерно сходятся для всех $\omega < \infty$. С помощью прямых вычислений нетрудно убедиться, что $S_i{}^j$ является действительно ортогональной матрицей. □

Можно также доказать обратное утверждение, что любая ортогональная матрица имеет вид (2.123) для некоторого вектора $\boldsymbol{\omega}$.

Замечание. В отличие от элемента алгебры, элемент группы имеет как симметричную, так и антисимметричную части. □

Элемент группы $\mathbb{SO}(3)$ инвариантен относительно сдвига вектора вращения $\omega^i \mapsto \omega^i + 2\pi\omega^i/\omega$:

$$S_i{}^j\left(\omega^k + 2\pi\frac{\omega^k}{\omega}\right) = S_i{}^j(\omega^k),$$

и это единственная инвариантность. При сдвиге вектора $\boldsymbol{\omega}$ меняется только его длина $\omega \mapsto \omega + 2\pi$, а направление остается неизменным. В инвариантности матрицы вращений нетрудно убедиться, если заметить, что отношение ω^i/ω, определяющее направление вектора $\boldsymbol{\omega}$, не меняется при произвольном сдвиге:

$$\frac{\omega^i}{\omega} \mapsto \frac{\omega^i + c\omega^i/\omega}{\omega + c} = \frac{\omega^i}{\omega}, \qquad c \in \mathbb{R}.$$

Тем самым элемент группы вращений параметризуется точками евклидова пространства $\boldsymbol{\omega} \in \mathbb{R}^3$ с единственным отношением эквивалентности

$$\omega^i \sim \omega^i + 2\pi\frac{\omega^i}{\omega}. \tag{2.122}$$

При этом неопределенность в нуле раскрывается по радиальным направлениям $\omega^i = \epsilon k^i$, $\epsilon \to 0$. То есть начало координат отождествляется со всеми сферами радиуса $2\pi m$, $m = 1, 2, \ldots$

Вектор вращения $\boldsymbol{\omega}$ параметризует группу $\mathbb{SO}(3)$ следующим образом. Направление вектора $\boldsymbol{\omega}$ определяет ось вращения, а модуль вектора ω равен углу поворота. Таким образом, каждый элемент группы $\mathbb{SO}(3)$ отождествляется с точкой трехмерного шара $\mathbb{B}^3_\pi \subset \mathbb{R}^3$ радиуса π с центром в начале координат. При этом различным внутренним точкам шара соответствуют различные вращения, а диаметрально противоположные точки граничной сферы $\mathbb{S}^2_\pi$ необходимо отождествить, т.к. повороты вокруг фиксированной оси на углы π и $-\pi$ приводят к одному и тому же результату.

Таким образом, мы построили экспоненциальное отображение (2.121) алгебры Ли в группу Ли:

$$\mathfrak{so}(3) \ni \omega \mapsto S \in \mathbb{SO}(3).$$

Это отображение, конечно, не является взаимно однозначным. Оно также не является накрытием. Действительно, прообразом каждого нетривиального вращения вокруг некоторой оси является счетное число точек, лежащих на этой оси на расстоянии 2π друг от друга. В то же время прообразом единицы группы $e \in \mathbb{SO}(3)$ является начало координат в $\mathbb{R}^3 \approx \mathfrak{go}(3)$ и все сферы $\mathbb{S}^2_{2\pi m}$ радиусов $2\pi m$, $m = 1, 2, \ldots$.

Полная группа трехмерных вращений $\mathbb{O}(3)$ состоит из двух связных компонент: ортогональных матриц S_+ с положительным и S_- отрицательным определителем. Элементы полной группы вращений $\mathbb{O}(3)$ параметризуются элементом алгебры (2.120) следующим образом:

$$S_{\pm i}{}^j = \pm S_i{}^j \ \in \mathbb{O}(3). \tag{2.123}$$

То есть матрицы S_- получаются из матриц связной компоненты единицы умножением на оператор пространственного отражения $P := -\mathbb{1}$. При этом каждому элементу алгебры $\boldsymbol{\omega} \in \mathfrak{so}(3)$ соответствуют два элемента группы $S_\pm(\omega) \in \mathbb{O}(3)$ – по одному из каждой компоненты.

Обратные матрицы собственных вращений имеют вид

$$S^{-1}{}_i{}^j(\boldsymbol{\omega}) = S_i{}^j(-\boldsymbol{\omega}) = \delta^j_i \cos\omega - \frac{(\omega\varepsilon)_i{}^j}{\omega}\sin\omega + \frac{\omega_i\omega^j}{\omega^2}(1 - \cos\omega) \ \in \mathbb{SO}(3), \tag{2.124}$$

т.е. соответствуют противоположному вектору вращения $-\boldsymbol{\omega} \in \mathbb{R}^3$. Другими словами, обратный элемент группы $\mathbb{SO}(3)$ представляет собой поворот евклидова пространства вокруг той же оси, но на противоположный угол.

Свернем матрицы вращений $S_\pm$ с вектором $\{\omega_j\}$

$$S_{\pm i}{}^j\omega_j = \pm\omega_i.$$

Это значит, что вектор $\boldsymbol{\omega}$ является собственным вектором матриц вращений с собственными значениями ± 1. Другими словами, вращения S_+ и S_- оставляют, соответственно, ось вращения без изменения и меняют ее направление на противоположное.

Допустим, что элемент алгебры $\mathfrak{so}(3)$ (вектор поворота) дифференцируемо зависит от точки некоторого многообразия $\mathbb{M}$, т.е.

$$\mathbb{M} \ni x \mapsto \boldsymbol{\omega}(x) \in \mathfrak{so}(3).$$

Каждому вектору поворота по формуле (2.121) ставится в соответствие матрица вращений $S \in \mathbb{SO}(3)$. Введем обозначение

$$l_{\alpha i}{}^{j} = (\partial_\alpha S_\pm^{-1} S_\pm)_i{}^j = (\partial_\alpha S^{-1} S)_i{}^j, \tag{2.125}$$

где $\partial_\alpha := \partial/\partial x^\alpha$ – частная производная в локальной системе координат. Последнее равенство следует из (2.123). С помощью прямых вычислений можно убедиться в справедливости следующей формулы

$$l_{\alpha i}{}^{j} = -\frac{(\partial_\alpha \omega\varepsilon)_i{}^j}{\omega}\,\mathsf{sin}\,\omega - \frac{\partial_\alpha \omega (\omega\varepsilon)_i{}^j}{\omega}\left(1 - \frac{\mathsf{sin}\,\omega}{\omega}\right) + \\ + \frac{\partial_\alpha \omega_i \omega^j - \omega_i \partial_\alpha \omega^j}{\omega^2}(1 - \mathsf{cos}\,\omega) \quad \in \mathfrak{so}(3). \tag{2.126}$$

Эта матрица антисимметрична по своим индексам, и, значит, является элементом алгебры $\mathfrak{so}(3)$. Она представляет собой тривиальную $\mathbb{SO}(3)$-связность, для которой тензор кривизны тождественно равен нулю (чистая калибровка (14.64)). Если в качестве многообразия $\mathbb{M}$ выбрать многообразие самой группы вращений $\mathbb{SO}(3)$, то 1-форма $dx^\alpha l_{\alpha i}{}^j$ будет являться канонической 1-формой на $\mathbb{SO}(3)$ (см. раздел 9.2).

Построим гомоморфизм группы двумерных унитарных матриц $\mathbb{SU}(2)$ на группу собственных трехмерных вращений $\mathbb{SO}(3)$. Алгебра Ли $\mathfrak{su}(2)$ состоит из антиэрмитовых матриц $A = -A^\dagger$ с нулевым следом $\mathsf{tr}\,A = 0$. Элемент алгебры $\mathfrak{su}(2)$ имеет вид

$$r^k \frac{i}{2}(\sigma_k)_{\text{A}}{}^{\text{B}} \ \in \mathfrak{su}(2), \qquad \boldsymbol{r} = \{r^k\} \in \mathbb{R}^3, \tag{2.127}$$

где σ_k – матрицы Паули (см. приложение 28.2), и $\text{A}, \text{B} = 1, 2$ – матричные индексы. Множитель $i/2$ выбран с тем, чтобы коммутаторы векторов базиса $\frac{i}{2}\sigma_k$ имели вид (2.119) как и для алгебры $\mathfrak{so}(3)$. Это доказывает, что алгебры $\mathfrak{su}(2)$ и $\mathfrak{so}(3)$ изоморфны. Задать этот изоморфизм можно в явном виде. Для этого заметим, что произвольную антиэрмитову матрицу можно взаимно однозначно представить виде (2.127):

$$A = \frac{i}{2}\begin{pmatrix} r^3 & r^1 - ir^2 \\ r^1 + ir^2 & -r^3 \end{pmatrix}. \tag{2.128}$$

Предложение 2.8.2. *Отображение*

$$\varphi: \quad \mathfrak{su}(2) \ni A \mapsto \varphi(A) := (\omega\epsilon) \in \mathfrak{so}(3), \tag{2.129}$$

где $\boldsymbol{\omega} = \boldsymbol{r}$ из представления (2.128) и элемент алгебры $(\omega\epsilon)$ определен формулой (2.120), является изоморфизмом алгебр $\mathfrak{su}(2) \leftrightarrow \mathfrak{so}(3)$.

Доказательство. Взаимная однозначность и линейность отображения φ очевидны. Сохранение коммутатора при отображении,

$$\varphi\Big([A, B]\Big) = \Big[\varphi(A), \varphi(B)\Big], \qquad A, B \in \mathfrak{su}(2),$$

проверяется прямой проверкой. Это также следует из того, что коммутаторы базисных векторов в алгебрах совпадают. □

Очевидно, что отображение φ является гладким. Заметим также, что

$$\det A = \frac{1}{4}\Big((\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2\Big).$$

Пусть $U \in \mathbb{SU}(2)$ – произвольная унитарная матрица, $UU^\dagger = U^\dagger U = 1$, с единичным определителем, $\det U = 1$. Построим новую матрицу

$$A' = UAU^\dagger.$$

Эта матрица также антиэрмитова, $A' = -A'^\dagger$, и ее след равен нулю, $\mathrm{tr}\, A' = 0$. Поэтому матрицу A' также можно представить в виде

$$A' = \begin{pmatrix} r'^3 & r'^1 - ir'^2 \\ r'^1 + ir'^2 & -r'^3 \end{pmatrix}.$$

Поскольку $\det A' = \det A$, то

$$(\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2 = (\omega'^1)^2 + (\omega'^2)^2 + (\omega'^3)^2$$

Следовательно, вектор вращения $\boldsymbol{\omega}'$ получаются из вектора $\boldsymbol{\omega}$ с помощью некоторой ортогональной матрицы. Поскольку группа $\mathbb{SU}(2)$ связна, то из непрерывности следует, что каждой унитарной матрице $U \in \mathbb{SU}(2)$ однозначно ставится в соответствие некоторая ортогональная матрица $S \in \mathbb{SO}(3)$ из связной компоненты единицы. Полученное отображение гомоморфно, т.к. групповые операции, как нетрудно проверить, согласованы. Ядро гомоморфизма состоит из двух элементов $\{\mathbb{1}, -\mathbb{1}\}$, где $\mathbb{1}$ – двумерная единичная матрица. Это значит, что каждой ортогональной матрице $S \in \mathbb{SO}(3)$ соответствуют две унитарные матрицы U и $-U$ группы $\mathbb{SU}(2)$. Тем самым отображение $\mathbb{SU}(2) \to \mathbb{SO}(3)$ является двулистным накрытием, и это накрытие универсально, т.к. многообразие группы $\mathbb{SU}(2)$ связно и односвязно (см. раздел 12). Таким образом, справедлива

Теорема 2.8.1. *Существует гомоморфизм $\mathbb{SU}(2) \to \mathbb{SO}(3)$, при котором группа $\mathbb{SU}(2)$ является двулистным универсальным накрытием группы $\mathbb{SO}(3)$.*

Произвольную унитарную матрицу с единичным определителем $U \in \mathbb{SU}(2)$ можно параметризовать двумя комплексными числами $a, b \in \mathbb{C}$:

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \tag{2.130}$$

которые удовлетворяют условию $|a|^2 + |b|^2 = 1$. Это значит, что элементы группы $\mathbb{SU}(2)$ находятся во взаимно однозначном соответствии с точками трехмерной сферы $\mathbb{S}^3$, вложенной в четырехмерное евклидово пространство $\mathbb{R}^4$, т.е.

$$\mathbb{SU}(2) \approx \mathbb{S}^3.$$

Другими словами, трехмерную сферу можно оснастить групповой структурой. При гомоморфизме $\mathbb{SU}(2) \to \mathbb{SO}(3)$ два элемента U и $-U$ отображаются в один элемент группы вращений, и этим элементам соответствуют две противоположные точки на сфере $\mathbb{S}^3$. Это значит, что с топологической точки зрения группа $\mathbb{SO}(3)$ представляет собой сферу $\mathbb{S}^3$ с отождествленными диаметрально противоположными точками, т.е. проективную плоскость

$$\mathbb{SO}(3) \approx \mathbb{RP}^3 = \frac{\mathbb{S}^3}{\mathbb{Z}_2}.$$

Важно отметить, что из каждой пары элементов U и $-U$ невозможно выбрать по одному элементу группы $\mathbb{SU}(2)$ так, чтобы эти элементы образовали бы группу сами по себе. Это связано с тем, что при вращении на угол $2\pi t$, где $t \in [0,1]$, при $t = 1$ мы имеем тождественное преобразование в группе $\mathbb{SO}(3)$, а матрицы $U \in \mathbb{SU}(2)$ меняют свой знак $U \to -U$. Другими словами, когда параметр t пробегает единичный отрезок мы имеем замкнутый путь в $\mathbb{SO}(3)$ и отрезок в группе $\mathbb{SU}(2)$. Это доказывает, в частности, что группа $\mathbb{SO}(3)$ не является подгруппой в $\mathbb{SU}(2)$.

Гомоморфизм $\mathbb{SU}(2) \to \mathbb{SO}(3)$ имеет следующую геометрическую интерпретацию. Каждое вращение трехмерного евклидова пространства $\mathbb{R}^3$ порождает вращение сферы $\mathbb{S}^2$ с центром в начале координат. Верно и обратное утверждение: каждое вращение сферы однозначно определяет вращение евклидова пространства. При стереографической проекции все точки сферы за исключением северного полюса проектируются на евклидову плоскость $\mathbb{R}^2$, на которой можно ввести комплексные координаты $z \in \mathbb{C} \approx \mathbb{R}^2$ (детали конструкции содержатся в [?]).

Теорема 2.8.2. *Каждому собственному вращению $S \in \mathbb{SO}(3)$ соответствуют два дробно линейных преобразования комплексных координат*

$$z' = \pm\frac{az + b}{-\bar{b}\bar{z} + \bar{a}}, \qquad |a|^2 + |b|^2 = 1.$$

Обратно. Каждой унитарной матрице (2.130) соответствует некоторое вращение.

Доказательство. См., например, [?]. □

Замечание. Матрицу (2.130) можно рассматривать, как представление кватернионов (28.45) с единичной нормой. □

Отображение $U \mapsto S$ унитарной матрицы (2.130) в ортогональную матрицу $S \in \mathbb{SO}(3)$ можно записать в явном виде [?]

$$S = \begin{pmatrix} \frac{1}{2}(a^2 - b^2 + \bar{a}^2 - \bar{b}^2) & \frac{i}{2}(-a^2 - b^2 + \bar{a}^2 + \bar{b}^2) & -ab - \bar{a}\bar{b} \\ \frac{i}{2}(a^2 - b^2 - \bar{a}^2 + \bar{b}^2) & \frac{1}{2}(a^2 + b^2 + \bar{a}^2 + \bar{b}^2) & -i(ab - \bar{a}\bar{b}) \\ a\bar{b} + \bar{a}\bar{b} & i(-a\bar{b} + \bar{a}b) & a\bar{a} - b\bar{b} \end{pmatrix}. \tag{2.131}$$

Выше было доказано, что группа $\mathbb{SU}(2)$ дважды накрывает группу вращений $\mathbb{SO}(3)$. Покажем, как это выглядит при использовании явной параметризации унитарной группы $\mathbb{SU}(2)$.

Предложение 2.8.3. *Пусть элемент алгебры $\mathfrak{su}(2)$ имеет вид (2.127) с $\boldsymbol{r} = \boldsymbol{\omega}$. Тогда ему соответствует элемент группы $\mathbb{SU}(2)$*

$$U_{\mathrm{A}}{}^{\mathrm{B}}(\boldsymbol{\omega}) = \left(\mathsf{e}^{(i\omega^k \sigma_k/2)}\right)_{\mathrm{A}}{}^{\mathrm{B}} = \delta_{\mathrm{A}}^{\mathrm{B}} \cos\frac{\omega}{2} + i\frac{\omega^k \sigma_{k\mathrm{A}}{}^{\mathrm{B}}}{\omega} \sin\frac{\omega}{2} \quad \in \mathbb{SU}(2), \tag{2.132}$$

где $\omega = \sqrt{\omega^i \omega_i}$ – модуль вектора вращения $\boldsymbol{\omega}$.

Доказательство. Прямая проверка путем разложения в ряды. □

Верно также и обратное утверждение: любая унитарная матрица с единичным определителем имеет вид (2.132) для некоторого вектора $\boldsymbol{\omega} \in \mathbb{R}^3$. Из вида элемента группы следует равенство

$$U_{\text{A}}{}^{\text{B}}\left(\omega^i + 4\pi\frac{\omega^i}{\omega}\right) = U_{\text{A}}{}^{\text{B}}(\omega^i).$$

То есть для группы $\mathbb{SU}(2)$ в пространстве параметров $\boldsymbol{\omega} \in \mathbb{R}^3$ имеется отношение эквивалентности

$$\omega^i \sim \omega^i + 4\pi\frac{\omega^i}{\omega}. \tag{2.133}$$

При этом начало координат отождествляется со всеми сферами радиуса $4\pi m$, $m = 1, 2, \ldots$. По сравнению с отношением эквивалентности (2.122) для группы $\mathbb{SO}(3)$ сдвиг происходит на вектор вдвое большей длины. Тем самым групповое многообразие параметризуется внутренними точками шара $\mathbb{B}^3_{2\pi} \subset \mathbb{R}^3$ радиуса 2π. При этом все точки граничной сферы $\mathbb{S}^3_{2\pi}$ необходимо отождествить, т.к. второе слагаемое в (2.132) пропадает. Это приводит к второму отношению эквивалентности в пространстве параметров

$$\omega^i_1|_{\omega_1=2\pi} \sim \omega^i_2|_{\omega_2=2\pi}. \tag{2.134}$$

Дополнительное соотношение эквивалентности связано с отсутствием третьего слагаемого в (2.132) по сравнению с (2.121). В этом также есть отличие от группы вращений, для которой отождествляются только диаметрально противоположные точки граничной сферы, что входит в отношение эквивалентности (2.122). Можно показать, что в пространстве параметров для группы $\mathbb{SU}(2)$ других отношений эквивалентности кроме (2.133) и (2.134) не существует.

Выше было показано, что групповое многообразие унитарной группы $\mathbb{SU}(2)$ представляет собой трехмерную сферу $\mathbb{S}^3_{2\pi}$. Если начало координат в приведенной параметризации соответствует южному полюсу сферы $\mathbb{S}^3_{2\pi}$, то граничная сфера $\omega = 2\pi$ – северному. Это является следствием отношения эквивалентности (2.134).

Из гомоморфизма $\mathbb{SU}(2) \to \mathbb{SO}(3)$ следует гомоморфизм представлений этих групп. При этом каждой матрице представления группы вращений $\mathbb{SO}(3)$ соответствуют две матрицы представления группы $\mathbb{SU}(2)$, которые отличаются знаком, если представление точное. В физической литературе принято говорить, что точное представление группы $\mathbb{SU}(2)$ является двузначным представлением группы вращений $\mathbb{SO}(3)$. Оно называется *спинорным представлением* группы $\mathbb{SO}(3)$. (См. также пример 10.4.2) Это представление двузначно, т.к. одному вращению $S \in \mathbb{SO}(3)$ соответствует два элемента унитарной группы $\pm U \in \mathbb{SU}(2)$. И эту неоднозначность нельзя устранить. Действительно, если вращению S поставить в соответствие только одну унитарную матрицу U, то после тождественного преобразования (поворота на угол $\omega = 2\pi$), матрица U изменит знак. Это следует из явного вида унитарного преобразования (2.132), т.к. аргумент тригонометрических функций равен половине угла поворота.

Поскольку группа вращений компактна, то все ее неприводимые представления унитарны и конечномерны. При этом любое другое ее представление разлагается в прямую сумму неприводимых представлений. Подробное описание всех представлений группы вращений содержится в [**?**, **?**].

В заключение настоящего раздела скажем несколько слов о группе вращений $\mathbb{O}(n)$ n-мерного евклидова пространства $\mathbb{R}^n$. Под группой $\mathbb{O}(1) \approx \mathbb{Z}_2$ удобно понимать группу, состоящую из двух элементов $\{1, -1\}$, которые соответствуют тождественному преобразованию и отражению вещественной прямой относительно начала

координат. При $n \geq 2$ группа $\mathbb{O}(n)$ сохраняет евклидову метрику и состоит из двух несвязных компонент при любом $n \geq 2$: матриц с положительным и отрицательным определителем. Матрица пространственных отражений $-\mathbb{1}$ при четных n имеет положительный определитель, принадлежит связной компоненте единицы $\mathbb{SO}(n)$ и не меняет ориентации декартовых осей. Поэтому преобразование четности в этом случае целесообразно определить, как отражение только одной, например, первой координаты

$$P = \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix},$$

где $\mathbb{1}$ – единичная матрица размера $(n-1) \times (n-1)$.

2.9 Пространство Минковского

В настоящем разделе мы рассмотрим пространство-время Минковского произвольного числа измерений и его общие свойства. Затем остановимся на специфических свойствах пространств Минковского трех и четырех измерений, которые наиболее часто встречаются в физических приложениях.

Определение. *Пространством* или *пространством-временем Минковского* $\mathbb{R}^{1,n-1}$ размерности n называется евклидово пространство $\mathbb{R}^n$, на котором в декартовых координатах задана *лоренцева метрика*

$$\eta = \{\eta_{ab}\} = \mathsf{diag}\,(1\underbrace{-1,\ldots,-1}_{n-1}) \tag{2.135}$$

Декартовы системы координат в пространстве Минковского называются *инерциальными*. □

Замечание. Выбор общего знака в метрике (2.135) является условным, и часто используется метрика противоположного знака, когда время входит в интервал со знаком минус, а пространственные координаты – со знаком плюс. □

Таким образом в пространстве Минковского задано две метрики: евклидова и лоренцева. Евклидова метрика задается естественным образом на прямом произведении прямых и определяет топологию $\mathbb{R}^{1,n-1}$. Лоренцева метрика (2.135) не является положительно определенной, и ее нельзя использовать для определения расстояния в топологическом смысле (см. раздел 2.3.1). Поэтому не следует рассматривать пространство Минковского просто как прямое произведение прямых, на котором задана только метрика (2.135), поскольку евклидова метрика необходима для задания топологии. На самом деле это всегда подразумевается, т.к. непрерывность функций в пространстве Минковского понимается относительно естественной топологии евклидова пространства. Инерциальные координаты, в которых метрика имеет вид (2.135), обозначим $\{x^0, x^1, \ldots x^{n-1}\} = \{x^a\}$, $a = 0, 1, \ldots, n-1$. Здесь первая координата $t := x^0$ называется временем. Остальные координаты x^i, $i = 1, \ldots, x^{n-1}$, называются пространственноподобными и параметризуют пространственные сечения $x^0 = \mathsf{const}$. Очевидно, что пространство Минковского есть прямое произведение

$$\mathbb{R}^{1,n-1} = \mathbb{R} \times \mathbb{R}^{n-1},$$

где первый сомножитель соответствует времени, а второй – $(n-1)$-мерному евклидову пространству с отрицательно определенной метрикой.

Замечание. Здесь и в дальнейшем пространственные индексы будут нумероваться латинскими буквами из середины алфавита $\{i, j, k, \dots\}$. Мнемоническое правило следующее. Эти индексы образуют подмножество всего алфавита $\{a, b, c, \dots\}$, так же как множество чисел $\{1, 2, \dots, n-1\}$ является подмножеством чисел $\{0, 1, 2, \dots, n-1\}$. □

Замечание. Между евклидовым пространством и пространством Минковского существует связь. Если время x^0 умножить на мнимую единицу $x^0 \mapsto ix^0$ и изменить общий знак метрики, то получим евклидову метрику. К тому же результату можно прийти, если на мнимую единицу умножить все пространственные координаты $x^i \mapsto ix^i$. Оба преобразования называются *комплексным поворотом*. При комплексном повороте оператор Даламбера переходит в оператор Лапласа, что приводит к качественному отличию решений уравнений, содержащих эти операторы. В приложениях важную роль играет связь между этими решениями. □

Определение. В каждой точке пространства Минковского $y \in \mathbb{R}^{1,n-1}$ уравнение

$$\eta_{ab}(x^a - y^a)(x^b - y^b) = 0,$$

задает *световой конус* с вершиной в точке y (см. рис.2.10,a). Конус представляет собой объединение двух связных компонент: светового конуса *прошлого* $x^0 \le y^0$ и *будущего* $x^0 \ge y^0$ с общей вершиной в точке y. □

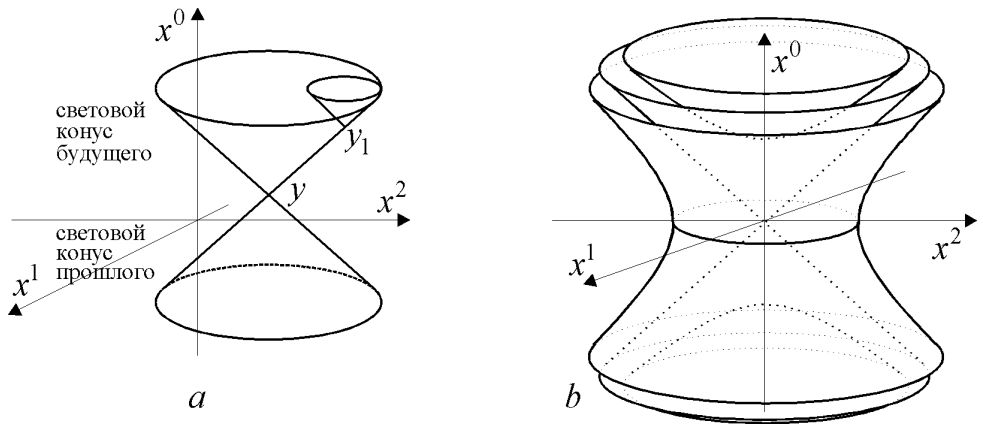


Рис. 2.10: Световой конус будущего и прошлого в трехмерном пространстве Минковского

Световые конусы прошлого и будущего являются $(n-1)$-мерными подмногообразиями в пространстве Минковского $\mathbb{R}^{1,n-1}$. В целом световой конус подмногообразием не является из-за общей вершины. Очевидно, что образующие световых конусов имеют нулевую длину.

Определение. Дифференцируемая кривая $\{x^a(s)\}$ в пространстве Минковского называется *времениподобной*, *пространственноподобной* или *светоподобной (нулевой)*, если касательный вектор, соответственно, удовлетворяет условиям:

$$\dot{x}^a \dot{x}^b \eta_{ab} > 0, \qquad \dot{x}^a \dot{x}^b \eta_{ab} < 0, \qquad \dot{x}^a \dot{x}^b \eta_{ab} = 0, \tag{2.136}$$

для всех значений параметра $s \in \mathbb{R}$. Если в некоторой точке y, через которую проходит кривая, выполнены два условия $\dot{x}^a \dot{x}^b \eta_{ab} > 0$ и $\dot{x}^0 > 0$, то мы говорим, что касательный вектор к кривой в точке y времениподобен и направлен в будущее. □

Предложение 2.9.1. *Любая дифференцируемая времениподобная кривая в пространстве Минковского $\mathbb{R}^{1,n-1}$, проходящая через точку y, целиком лежит внутри светового конуса с вершиной в этой точке.*

Доказательство. Если кривая дифференцируема, времениподобна и проходит через точку y, то из дифференцируемости следует, что она будет лежать внутри светового конуса в некоторой окрестности точки y. Кроме того она не может касаться или пересекать световой конус, т.к. в такой точке она не была бы времениподобна. □

Для полноты картины обобщим пространство Минковского на метрику произвольной сигнатуры.

Определение. Псевдоевклидовым пространством $\mathbb{R}^{p,q}$, $p+q=n$, называется евклидово пространство $\mathbb{R}^n$ с метрикой

$$\eta = \{\eta_{ab}\} = \mathsf{diag}\,(\underbrace{+\ldots+}_{p}\underbrace{-\ldots-}_{q}), \qquad a,b=1,\ldots,n, \tag{2.137}$$

заданной в декартовой системе координат. □

Многие свойства пространства Минковского естественным образом переносятся на псевдоевклидово пространство. При этом группа Лоренца $\mathbb{O}(1,n-1)$ заменяется на группу псевдовращений $\mathbb{O}(p,q)$, оставляющую метрику (2.137) инвариантной. Группы псевдовращений $\mathbb{O}(p,q)$ при $1<p<n-1$ устроены сложнее группы Лоренца и в настоящей монографии рассматриваться не будут.

2.9.1 Группа Пуанкаре

Определение. Линейные неоднородные преобразования декартовых координат пространства Минковского $\mathbb{R}^{1,n-1}$, оставляющие инвариантной квадратичную форму

$$ds^2 = \eta_{ab}dx^a dx^b \tag{2.138}$$

образуют группу Ли, которая называется *группой Пуанкаре* и обозначается $\mathbb{IO}(1,n-1)$. Эта группа состоит из лоренцевых вращений с матрицами $S=\{S_b{}^a\}$ и сдвигов на вектор $a=\{a^a\}$:

$$x^a \mapsto x'^a = x^b S_b{}^a + a^a, \tag{2.139}$$

где постоянная матрица S является решением уравнения

$$\eta_{ab} = S_a{}^c S_b{}^d \eta_{cd}, \tag{2.140}$$

и $a\in\mathbb{R}^{1,n-1}$ – постоянный вектор. Множество матриц S образует группу Ли, которая называется *группой Лоренца* и обозначается $\mathbb{O}(1,n-1)$. Абелева подгруппа, параметризуемая вектором a, называется *группой трансляций.* □

Замечание. В отличие от группы аффинных преобразований (2.49) матрицы S являются элементами группы Лоренца $\mathbb{O}(1,n-1)$, а не общей линейной группы $\mathbb{GL}(n,\mathbb{R})$. □

Предложение 2.9.2. *Любое (возможно, нелинейное) преобразование декартовых координат пространства Минковского, оставляющее квадратичную форму (2.138) инвариантной, является аффинным, т.е. имеет вид (2.139) с некоторой невырожденной матрицей S.*

Доказательство. См., например, [?]. □

Это утверждение доказывает, что преобразования из группы Пуанкаре исчерпывают все возможные преобразования, сохраняющие метрику Лоренца (2.135). Другими словами, преобразования (2.140) являются движениями пространства Минковского общего вида.

Два последовательных преобразования S_1, a_1 и S_2, a_2, которые отображают $x \mapsto x' \mapsto x''$ имеют вид

$$x''^a = x'^b S_{2b}{}^a + a_2^a = x^c (S_{1c}{}^b S_{2b}{}^a) + (a_1^b S_{2b}{}^a + a_2^a). \tag{2.141}$$

Группа Лоренца является подгруппой группы Пуанкаре и имеет размерность

$$\dim \mathbb{O}(1, n-1) = \frac{n(n-1)}{2}.$$

Она неабелева и некомпактна. Сдвиги образуют нормальную абелеву подгруппу группы Пуанкаре размерности n. При этом параметр сдвига a преобразуется как вектор относительно преобразований Лоренца. В целом группа Пуанкаре представляет собой полупрямое произведение группы Лоренца на подгруппу сдвигов и имеет размерность

$$\dim \mathbb{IO}(1, n-1) = \frac{n(n+1)}{2}.$$

Из определяющего уравнения (2.140) следует, что $\det S = \pm 1$. Кроме этого, 00 компоненту данного уравнения можно переписать в виде

$$(S_0{}^0)^2 = 1 + \sum_{i=1}^{n-1} (S_0{}^i)^2.$$

Заметим также, что из условия инвариантности обратной метрики Лоренца:

$$\eta^{ab} = \eta^{cd} S_c{}^a S_d{}^b,$$

которое эквивалентно (2.140), вытекает равенство

$$(S_0{}^0)^2 = 1 + \sum_{i=1}^{n-1} (S_i{}^0)^2, \tag{2.142}$$

где суммирование проходит по нижнему индексу. Отсюда вытекает, что для любой матрицы лоренцевых вращений либо $S_0{}^0 \geq 1$, либо $S_0{}^0 \leq -1$. Ясно, что никакую матрицу преобразований Лоренца с определителем $\det S = 1$ нельзя непрерывно деформировать в матрицу с определителем $\det S = -1$, т.к. множество из двух элементов $\{1, -1\}$ не является связным. Аналогично, матрицу с $S_0{}^0 \geq 1$ нельзя непрерывно деформировать в матрицу с $S_0{}^0 \leq -1$. Отсюда вытекает, что группа Лоренца состоит не менее, чем из четырех связных компонент.

Теорема 2.9.1. *Группа Лоренца $\mathbb{O}(1, n-1)$ при любом $n \geq 2$ состоит из четырех связных компонент.*

Доказательство. При $n = 2$ теорема была доказана в разделе 2.7. Доказательство для произвольного n содержится, например, в [?]. □

Определение. Преобразования Лоренца с положительным определителем называются *собственными*. Преобразования Лоренца, для которых ${S_0}^0 \geq 1$ называются *ортохронными*. □

Предложение 2.9.3. *Преобразование Лоренца является ортохронным тогда и только тогда, когда оно всякий времениподобный вектор $X = \{X^0, X^1, \ldots, X^{n-1}\}$ в начале координат, направленный в будущее, $X^0 > 0$, переводит во времениподобный вектор, который также направлен в будущее.*

Доказательство. Пусть $S:\ X^a \mapsto X'^a = X^b {S_b}^a$ – преобразование Лоренца. При этом времениподобный вектор переходит во времениподобный, т.к. квадрат вектора сохраняется при любом преобразовании Лоренца. Для произвольного времениподобного вектора выполнено неравенство

$$(X^0)^2 - (X^1)^2 - \ldots - (X^{n-1})^2 > 0.$$

Применим неравенство Коши–Буняковского:

$$\begin{aligned}(X^1 {S_1}^0 + \ldots + X^{n-1} {S_{n-1}}^0)^2 \leq \Big(({S_1}^0)^2 + \ldots + ({S_{n-1}}^0)^2\Big)\Big((X^1)^2 + \ldots + (X^{n-1})^2\Big) \leq \\ \leq \Big(({S_0}^0)^2 - 1\Big)(X^0)^2 < ({S_0}^0)^2 (X^0)^2,\end{aligned}$$

где мы воспользовались равенством (2.142). Отсюда вытекает, что нулевая компонента преобразованного вектора

$$X'^0 := X^0 {S_0}^0 + X^1 {S_1}^0 + \ldots + X^{n-1} {S_{n-1}}^0$$

имеет тот же знак, что и ${S_0}^0$, если $X^0 > 0$. Отсюда же следует, что если $X^0 > 0$ и X'^0, то ${S_0}^0 > 0$. □

Определение. Введем оператор *обращения времени* T и *пространственного отражения (четности)* P:

$$\begin{aligned} T:&\quad (x^0, x^i) \mapsto (-x^0, x^i), && \det T = -1, \\ P:&\quad (x^0, x^1, x^2, \ldots, x^{n-1}) \mapsto (x^0, -x^1, x^2, \ldots, x^{n-1}), && \det P = -1. \end{aligned} \tag{2.143}$$

Введем также обозначение для их композиции

$$R := PT = TP:\quad (x^0, x^1, x^2, \ldots, x^{n-1}) \mapsto (-x^0, -x^1, x^2, \ldots, x^{n-1}), \qquad \det R = 1, \tag{2.144}$$

которая отражает только первые две координаты. □

Замечание. Пространственные отражения при четных n можно было бы определить, как отражение всех пространственных координат, т.к. в этом случае $\det P = -1$, и ориентация осей координат меняется. Тогда оператор R соответствует полному отражению всех координат. При нечетных n отражение всех пространственных координат имеет положительных определитель и не подходит, т.к. принадлежит связной компоненте единицы группы. Отражение одной (любой) из координатных осей меняет ориентацию пространства Минковского и может быть использовано в качестве оператора четности. Такое определение подходит для пространства Минковского произвольного числа измерений. □

Так же, как и в двумерном случае, полная группа Лоренца $\mathbb{O}(1,n-1)$ при $n \geq 3$ состоит из четырех связных компонент, которые получаются из связной компоненты единицы $\mathbb{S}_{+\uparrow}(1,n-1)$ обращением времени, пространственным отражением и их композицией:

$$\begin{aligned} \mathbb{S}_{-\downarrow} &= \mathbb{S}_{+\uparrow}(1,n-1)T, \\ \mathbb{S}_{-\uparrow} &= \mathbb{S}_{+\uparrow}(1,n-1)P, \\ \mathbb{S}_{+\downarrow} &= \mathbb{S}_{+\uparrow}(1,n-1)R. \end{aligned} \tag{2.145}$$

Единица полной группы Лоренца $\mathbb{O}(1,n-1)$ содержится в компоненте $\mathbb{S}_{+\uparrow}$. Эта компонента является связной группой Ли, называется *собственной ортохронной группой Лоренца* и обозначается $\mathbb{SO}_{\uparrow}(1,n-1)$. Кроме того, полная группа Лоренца $\mathbb{O}(1,n-1)$ содержит еще три подгруппы, состоящие из двух компонент каждая:

$$\begin{aligned} \mathbb{O}_{\downarrow}(1,n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\downarrow}, && \det S = -1, \quad S_0{}^0 < 0, \\ \mathbb{O}_{\uparrow}(1,n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{-\uparrow}, && \det S = -1, \quad S_0{}^0 > 0, \\ \mathbb{SO}(1,n-1) &\approx \mathbb{S}_{+\uparrow} \cup \mathbb{S}_{+\downarrow}, && \det S = 1, \quad S_0{}^0 < 0. \end{aligned} \tag{2.146}$$

Все четыре подгруппы являются нормальными подгруппами в $\mathbb{O}(1,n-1)$.

Очевидно, что отображения (2.145) взаимно однозначны и гладки. Поэтому как многообразия все четыре связные компоненты групп $\mathbb{O}(1,n-1)$ диффеоморфны между собой:

$$\mathbb{S}_{+\uparrow} \approx \mathbb{S}_{-\downarrow} \approx \mathbb{S}_{-\uparrow} \approx \mathbb{S}_{+\downarrow}.$$

Компоненты (2.145) являются смежными классами в группе $\mathbb{O}(1,n-1)$ по нормальной подгруппе $\mathbb{SO}_{\uparrow}(1,n-1)$. В качестве их представителей в фактор группе $\mathbb{O}(1,n-1)/\mathbb{SO}_{\uparrow}(1,n-1)$ можно выбрать матрицы T, P, R. Фактор группа $\mathbb{O}(1,n-1)/\mathbb{SO}_{\uparrow}(1,n-1)$ – это 4-группа Клейна $\mathbb{K}_4 = \{\mathbb{1}, T, P, R\}$, рассмотренная в разделе 2.7.

Рассмотрим автоморфизмы связной компоненты единицы группы Лоренца произвольной размерности. Пусть $S \in \mathbb{SO}_{\uparrow}(1,n-1)$, тогда отображение

$$S \mapsto S_0 S S_0^{-1},$$

где $S_0 \in \mathbb{SO}_{\uparrow}(1,n-1)$ – произвольный фиксированный элемент из связной компоненты единицы, задает внутренний автоморфизм группы собственных преобразований Лоренца. Отображения

$$S \mapsto TST^{-1} \quad \text{и} \quad S \mapsto PSP^{-1}, \tag{2.147}$$

определяемые обращением времени и пространственным отражением, также задают автоморфизм собственных преобразований Лоренца. Однако этот автоморфизм будет внешним, т.к. преобразования T и P не принадлежат $\mathbb{SO}_{\uparrow}(1,n-1)$. Можно доказать, что внешние автоморфизмы (2.147) совпадают между собой с точностью до внутреннего автоморфизма[2]. Чтобы найти явный вид внешнего автоморфизма, перепишем определение лоренцевых вращений (2.140), не делая различия между верхними и нижними индексами,

$$\eta = S\eta S^{\mathrm{T}} \quad \text{или} \quad (S^{\mathrm{T}})^{-1} = \eta^{-1} S \eta.$$

[2]Если в пространстве Минковского четной размерности оператор пространственного отражения определен, как отражение всех пространственных координат, то внешние автоморфизмы (2.147) просто совпадают.

Поскольку $\eta^{-1} = -T$, то отсюда следует формула для внешнего автоморфизма

$$TST^{-1} = (S^{\mathrm{T}})^{-1}. \tag{2.148}$$

Конечно, внешние автоморфизмы связной компоненты единицы $\mathbb{SO}_{\uparrow}(1, n-1)$ являются внутренними для полной группы Лоренца $\mathbb{O}(1, n-1)$.

Замечание. Автоморфизм (2.148) играет большую роль при рассмотрении спинорных представлений полной группы Лоренца $\mathbb{O}(1, n-1)$, когда требуется определить оператор обращения времени и четности. □

Орбитами точек пространства Минковского (см. рис.**??**,*b*), лежащих вне светового конуса, относительно действия собственной ортохронной группы Лоренца $\mathbb{SO}_{\uparrow}(1, n-1)$ являются однополостные гиперболоиды, определяемые уравнением

$$(x^0)^2 - (x^1)^2 - \ldots - (x^{n-1})^2 = -r^2, \qquad r > 0.$$

Для точек, лежащих внутри конусов будущего и прошлого, орбитами являются, соответственно, верхняя и нижняя полы двуполостного гиперболоида

$$(x^0)^2 - (x^1)^2 - \ldots - (x^{n-1})^2 = r^2, \qquad r > 0.$$

Сами конусы будущего и прошлого представляют собой орбиты лежащих на них точек. Начало координат является неподвижной точкой относительно лоренцевых вращений. На каждой из орбит собственная группа Лоренца $\mathbb{SO}_{\uparrow}$ действует транзитивно, т.е. для любых двух точек, лежащих на одной орбите, найдется по крайней мере одно собственное ортохронное преобразование Лоренца, переводящее одну точку в другую. В целом действие группы $\mathbb{SO}_{\uparrow}(1, n-1)$ в пространстве Минковского $\mathbb{R}^{1,n-1}$ эффективно, но не свободно, т.к. начало координат является неподвижной точкой.

Замечание. Инвариантность законов Природы относительно действия группы Пуанкаре является фундаментальным требованием к современным моделям математической физики и составляет основное содержание специальной теории относительности. Эта группа не является полупростой, т.к. содержит нормальную абелеву подгруппу (сдвиги). Следовательно ее форма Киллинга–Картана вырождена, и это создает существенные трудности при построении физических моделей, т.к. на многообразии параметров группы Пуанкаре не существует двусторонне инвариантной невырожденной метрики. □

Бесконечно малые преобразования группы Пуанкаре в пространстве Минковского $\mathbb{R}^{1,n-1}$ в линейном приближении можно записать с помощью дифференциальных операторов $L_{ab} = -L_{ba}$ и P_a:

$$\delta x^a = \left(\frac{1}{2}\delta\omega^{bc}L_{bc} + \delta a^b P_b\right) x^a, \tag{2.149}$$

где $\delta\omega^{ab} = -\delta\omega^{ba} \ll 1$ и $\delta a^a \ll 1$ – параметры преобразований Лоренца и сдвигов, соответственно. Множитель $\frac{1}{2}$ перед $\delta\omega^{bc}L_{bc}$ связан с тем, что независимые параметры входят в эту сумму дважды. Например,

$$\delta\omega^{01}L_{01} + \delta\omega^{10}L_{10} = 2\delta\omega^{01}L_{01}.$$

Генераторы лоренцевых вращений и сдвигов представляются дифференциальными операторами (векторными полями Киллинга) на $\mathbb{R}^{1,n-1}$

$$L_{ab} = x_a\partial_b - x_b\partial_a, \tag{2.150}$$
$$P_a = \partial_a, \tag{2.151}$$

где подъем и опускание индексов осуществляется с помощью метрики Лоренца. При этом формула (2.149) переходит в равенство

$$\delta x^a = x^b \delta\omega_b{}^a + \delta a^a.$$

Дифференциальные операторы (2.150) и (2.151) действуют в алгебре функций $\mathcal{C}^1(\mathbb{R}^{1,n-1})$, заданных в пространстве Минковского $\mathbb{R}^{1,n-1}$. Например, изменение формы функции (см. раздел 3.13) при бесконечно малых преобразованиях в линейном приближении имеет вид

$$\delta f(x) := -\delta x^a \partial_a f(x) = -(x^b \delta\omega_b{}^a + \delta a^a)\partial_a f(x).$$

Матрица бесконечно малых вращений $S \in \mathbb{SO}_\uparrow(1, n-1)$ в линейном приближении имеет вид

$$S_b{}^a \approx \delta_b^a + \omega_b{}^a. \tag{2.152}$$

Отметим, что антисимметрия параметров лоренцевых вращений следует из уравнения (2.140) в линейном приближении. Каждый элемент антисимметричной матрицы ω^{ab} представляет собой угол бесконечно малого поворота в плоскости x^a, x^b. Лоренцевы вращения с параметрами ω^{0i}, $i = 1, \ldots, n-1$ называют *бустом*, чтобы отличать их от чисто пространственных вращений, соответствующих параметрам ω^{ij}.

Генераторы вращений и сдвигов удовлетворяют *алгебре Пуанкаре*

$$[L_{ab}, L_{cd}] = -\eta_{ac}L_{bd} + \eta_{ad}L_{bc} + \eta_{bc}L_{ad} - \eta_{bd}L_{ac}, \tag{2.153}$$
$$[L_{ab}, P_c] = -\eta_{ac}P_b + \eta_{bc}P_a, \tag{2.154}$$
$$[P_a, P_b] = 0, \tag{2.155}$$

что проверяется прямой проверкой.

При действии вращений на векторы в пространстве Минковского $\mathbb{R}^{1,n-1}$ генераторы лоренцевых вращений, но не сдвиги, можно представить также в матричном виде

$$S_b{}^a \approx \delta_b^a + \frac{1}{2}\delta\omega^{cd} L_{cd\,b}{}^a,$$

где

$$L_{cd\,b}{}^a = \eta_{db}\delta_c^a - \eta_{cb}\delta_d^a. \tag{2.156}$$

Это – фундаментальное или векторное представление образующих алгебры Ли группы Лоренца. Нетрудно проверить, что матричное представление генераторов вращений также удовлетворяет алгебре Ли (2.153), которую можно записать в виде

$$[L_{ab}, L_{cd}] = f_{ab\,cd}{}^{ef} L_{ef}, \tag{2.157}$$

где $f_{ab\,cd}{}^{ef}$ – структурные константы группы Лоренца для коммутационных соотношений (2.153). Трансляции не имеют матричного представления в пространстве Минковского. Группа Лоренца при $n > 2$ является неабелевой и простой. Ее форма Киллинга–Картана невырождена

$$\eta_{ab\,cd} = f_{ab\,ef}{}^{gh} f_{cd\,gh}{}^{ef} = 4(-\eta_{ac}\eta_{bd} + \eta_{ad}\eta_{bc}), \tag{2.158}$$

и поэтому может быть использована в качестве инвариантной метрики для построения инвариантов. Для матричного представления генераторов (2.156) справедливо равенство

$$\mathsf{tr}\,(L_{ab}L_{cd}) = \frac{1}{2}\eta_{ab\,cd}. \tag{2.159}$$

Замечание. При $n = 2$ собственная ортохронная группа Лоренца $\mathbb{SO}_{\uparrow}(1,1)$ является абелевой, ее структурные константы равны нулю и форма Киллинга–Картана вырождена. Заметим, что правая часть (2.158) отлична от нуля и в этом случае. Формула для следа (2.159) справедлива также при $n = 2$, если под $\eta_{ab\,cd}$ понимать правую часть (2.158), а не форму Киллинга–Картана. □

Группа Пуанкаре в n-мерном пространстве Минковского связана с группой Лоренца в $(n+1)$-мерном пространстве следующим образом. Пусть индекс $\text{A} = \{0, 1, \ldots, n-1, n\}$ нумерует координаты $(n+1)$-мерного пространства Минковского $\mathbb{R}^{1,n}$ с метрикой

$$\{\eta_{\text{AB}}\} = \mathsf{diag}\,(1, \underbrace{-1, \ldots, -1}_{n-1}, -k^2), \qquad k > 0. \tag{2.160}$$

Обозначим генераторы лоренцевых вращений, затрагивающих n-тую координату, через $\widetilde{P}_a = L_{an}$. Тогда алгебра Лоренца в $(n+1)$-мерном пространстве примет вид

$$\begin{aligned} [L_{ab}, L_{cd}] &= -\eta_{ac}L_{bd} + \eta_{ad}L_{bc} + \eta_{bc}L_{ad} - \eta_{bd}L_{ac}, \\ [L_{ab}, \widetilde{P}_c] &= -\eta_{ac}\widetilde{P}_b + \eta_{bc}\widetilde{P}_a, \\ [\widetilde{P}_a, \widetilde{P}_b] &= -k^2 L_{ab}. \end{aligned} \tag{2.161}$$

Эта алгебра отличается от алгебры Пуанкаре тем, что "сдвиги" $\widetilde{P}_a$ уже не коммутируют. Алгебра Пуанкаре получается после формального предела $k \to 0$, который называют *контракцией*. В этом пределе $\widetilde{P} \to P$ и метрика Лоренца (2.160) в $n+1$ мерном пространстве Минковского $\mathbb{R}^{1,n}$ вырождается.

2.9.2 Группа Галилея

Модели математической физики, которые инвариантны относительно действия группы Пуанкаре называются *релятивистскими*. Постулат о том, что физические модели должны быть инвариантны относительно группы Пуанкаре, лежит в основе *специальной теории относительности*, которая была предложена А. Эйнштейном в 1905 году []. До создания специальной теории относительности ньютонова механика точечных частиц рассматривалась как основная фундаментальная модель. В механике Ньютона важнейшую роль играют преобразования Галилея, которые также образуют группу симметрии пространства-времени.

Рассмотрим пространство-время Галилея $\mathbb{R}^1 \times \mathbb{R}^{n-1} = \mathbb{R}^4$, где первый сомножитель соответствует времени, $t = x^0 \in \mathbb{R}$, а второй – пространству. Пространственные декартовы координаты занумеруем буквами из середины латинского алфавита x^i, $i = 1, \ldots, n-1$.

Определение. Зададим преобразования в $\mathbb{R}^1 \times \mathbb{R}^{n-1}$ следующими формулами:

$$\begin{aligned} t &\mapsto t' = t + a, \\ x &\mapsto x'^i = x^j S_j{}^i + v^i t + b^i, \end{aligned} \tag{2.162}$$

где ${S_j}^i \in \mathbb{O}(n-1)$ – матрица ортогональных вращений, действующая в пространстве. Преобразования (2.162) содержат дополнительные параметры: *скорость* $\{v^i\} \in \mathbb{R}^{n-1}$, сдвиги по времени $a \in \mathbb{R}$ и сдвиги пространства $\{b^i\} \in \mathbb{R}^{n-1}$. Эти преобразования образуют *группу Галилея* $\mathbb{G}(1, n-1)$. Декартовы системы координат, связанные преобразованием (2.162), как и в пространстве Минковского, называются *инерциальными*.

Подгруппа группы Галилея, для которой начало координат является неподвижной точкой, называется *однородной группой Галилея* и обозначается $\mathbb{HG}(1, n-1)$. Она состоит из пространственных вращений с матрицей ${S_j}^i$ и преобразований, которые параметризуются вектором скорости v^i. Последние преобразования называются *галилеевыми бустами*. □

Посчитаем размерность групп. Пространственные вращения параметризуются $(n-1)(n-2)/2$ параметрами. Галилеевы бусты – $(n-1)$ параметром. Кроме того, сдвиги времени и пространства задаются n параметрами. Таким образом, преобразования из полной группы Галилея задаются $n(n+1)/2$ параметрами, также как и преобразования из группы Пуанкаре. Размерность однородной группы Галилея равна размерности группы Лоренца,

$$\mathsf{dim}\, \mathbb{HG}(1, n-1) = \frac{n(n-1)}{2}.$$

Пусть в галилеевом пространстве-времени произошло два события в точках (t_1, x_1) и (t_2, x_2). Эти события разделены во времени $\triangle t := t_2 - t_1$ и пространстве $\triangle x^i := x_2^i - x_1^i$. Тогда преобразования Галилея – это такие линейные неоднородные преобразования пространства-времени, которые оставляют инвариантными временной интервал $\triangle t$ и расстояние между двумя одновременными событиями $\delta_{ij} \triangle x^i \triangle x^j$, при $t = \mathsf{const}$.

Генераторы полной группы Галилея можно представить в виде дифференциальных операторов (векторных полей):

$$\begin{aligned} L_{ij} &= x_i \partial_j - x_j \partial_i, \\ P_i &= \partial_i, \\ B_i &= t\partial_i, \\ P_0 &= \partial_0. \end{aligned}$$

Эти генераторы удовлетворяют алгебре Галилея:

$$\begin{aligned} [L_{ij}, L_{kl}] &= -\delta_{ik} L_{jl} + \delta_{il} L_{jk} + \delta_{jk} L_{il} - \delta_{jl} L_{ik}, \\ [L_{ij}, P_k] &= -\delta_{ik} P_j + \delta_{jk} P_i, \\ [L_{ij}, B_k] &= -\delta_{ik} B_j + \delta_{jk} B_i, \\ [B_i, P_0] &= -P_i, \end{aligned} \tag{2.163}$$

где выписаны только отличные от нуля коммутаторы. Отсюда следует, что при пространственных вращениях генераторы галилеевых бустов преобразуются, как ковекторы $\{0, x_i\} \in \mathbb{R}^{1,n-1}$.

В этом представлении группа Галилея действует в алгебре функций $\mathcal{C}^1(\mathbb{R}^{1,n-1})$. При этом временна́я координата может быть сдвинута только на постоянный вектор. Механика Ньютона инварианта относительно преобразований Галилея. Поэтому говорят, что в механике Ньютона время имеет абсолютное значение.

Однородная группа Галилея представляет собой полупрямое произведение группы пространственных вращений $\mathbb{O}(n-1)$ на группу галилеевых бустов. Группа галилеевых бустов образует абелеву нормальную подгруппу однородной группы Галилея, которая, следовательно, не является полупростой. Отсюда следует, что в галилеевом пространстве $\mathbb{R}\times\mathbb{R}^{n-1}$ не существует метрики, инвариантной относительно преобразований Галилея.

2.9.3 Группа конформных преобразований

Рассмотрим пространство Минковского $\mathbb{R}^{1,n-1}$ с декартовыми координатами x^a, $a = 0,1,\ldots,n-1$. В настоящем разделе подъем и опускание индексов производится с помощью метрики Лоренца (2.135). Рассмотрим преобразования декартовых координат $x \mapsto x'$ и ослабим требование инвариантности лоренцевой метрики (2.140), заменив его следующим условием

$$\eta_{ab}dx'^a dx'^b = \Omega^2(x)\eta_{ab}dx^a dx^b, \tag{2.164}$$

где $\Omega(x)$ – произвольная отличная от нуля функция (конформный множитель). Для определенности, будем считать $\Omega > 0$. При конформных преобразованиях длины векторов меняются, а углы между ними сохраняются. Нетрудно доказать, что преобразования пространства Минковского (возможно, нелинейные), удовлетворяющие условию (2.164), образуют группу Ли.

Определение. Преобразования декартовых координат пространства Минковского, при которых метрика Лоренца умножается на некоторый отличный от нуля множитель (2.164) образуют *конформную группу.* □

Замечание. В настоящем разделе мы рассматриваем только связные компоненты единиц групп для того, чтобы не обсуждать сложные вопросы глобального устройства групп. □

Предложение 2.9.4. *Конформная группа состоит из подгруппы преобразований из группы Пуанкаре (2.139), для которых* $\Omega = 1$, специальных конформных преобразований, *которые параметризуются постоянным вектором* $b = \{b^a\} \in \mathbb{R}^{1,n-1}$,

$$x'^a = \frac{x^a + b^a x^2}{1 + 2bx + b^2x^2}, \qquad b \neq 0, \quad \Omega = \frac{1}{1 + 2bx + b^2x^2}, \tag{2.165}$$

где $bx = b^a x_a$, $b^2 = b^a b_a$, $x^2 = x^a x_a$, *и* дилатаций

$$x'^a = \Omega x^a, \qquad \Omega = \mathsf{const} \neq 1. \quad \square \tag{2.166}$$

Доказательство. См., например, []. □

Дилатации называются также *гомотетией.*

Генераторы бесконечно малых специальных конформных преобразований и дилатаций можно представить в виде дифференциальных операторов (векторных полей) на $\mathbb{R}^{1,n-1}$

$$K_a = x^2\partial_a - 2x_a x^b\partial_b, \tag{2.167}$$

$$D = x^a\partial_a. \tag{2.168}$$

Эти векторные поля не являются полями Киллинга, т.к. при соответствующих преобразованиях метрика меняется. Прямые вычисления с учетом представления (2.150), (2.151) для генераторов группы Пуанкаре приводят к следующим коммутационным соотношениям

$$\begin{aligned}
[L_{ab}, K_c] &= \eta_{ac} K_b - \eta_{bc} K_a, \\
[P_a, K_b] &= -2\eta_{ab} D - 2L_{ab}, \\
[P_a, D] &= P_a, \\
[K_a, D] &= -K_a, \\
[L_{ab}, D] &= [K_a, K_b] = [D, D] = 0.
\end{aligned} \tag{2.169}$$

Эти соотношения вместе с коммутационными соотношениями (2.153)–(2.155) для группы Пуанкаре показывают, что генераторы L_{ab}, P_a, K_a и D образуют алгебру Ли, соответствующую конформной группе преобразований. Она имеет размерность

$$\frac{1}{2}(n+1)(n+2).$$

В отличие от группы Пуанкаре группа конформных преобразований проста и ее связная компонента единицы изоморфна связной компоненте единицы группы псевдовращений $\mathbb{SO}(2,n)$, действующей в пространстве $\mathbb{R}^{2,n}$, размерность которого на два превышает размерность исходного пространства Минковского $\mathbb{R}^{1,n-1}$. Изоморфизм алгебр Ли можно задать следующим образом

$$L_{\text{AB}} = \begin{pmatrix} \frac{1}{2}L_{ab} & \frac{1}{4}(P_a + K_a) & \frac{1}{4}(P_a - K_a) \\ -\frac{1}{4}(P_b + K_b) & 0 & \frac{1}{2}D \\ -\frac{1}{4}(P_b - K_b) & -\frac{1}{2}D & 0 \end{pmatrix}, \tag{2.170}$$

где индексы A, B пробегают значения $\text{A} = \{a, n, n+1\} = \{0, 1, \ldots, n+1\}$, и метрика в пространстве $\mathbb{R}^{2,n}$ имеет вид

$$\eta_{\text{AB}} = \begin{pmatrix} \eta_{ab} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Поскольку конформная группа проста, то на групповом многообразии существует двусторонне инвариантная метрика. Это существенное отличие от группы Пуанкаре.

Оператор Даламбера в n-мерном пространстве Минковского

$$\Box = \partial_0^2 - \sum_{i=1}^{n-1} \partial_i^2$$

инвариантен относительно группы Пуанкаре $\mathbb{IO}(1, n-1)$, причем сдвиги действуют на оператор дифференцирования тривиально $\partial_a \mapsto \partial_a$. При специальных конформных преобразованиях и дилатациях оператор Даламбера умножается на конформный множитель:

$$\Box = \Omega^2 \Box',$$

где $\Box' := \eta^{ab} \partial'_a \partial'_b$. Это означает, что, если некоторая функция удовлетворяет уравнению Даламбера $\Box f = 0$, то она будет удовлетворять уравнению Даламбера после произвольного конформного преобразования координат.

Замечание. Конформные преобразования были введены в физику в 1909 году Кунигамом и Бейтменом [?], которые показали, что уравнения Максвелла ковариантны относительны конформных преобразований. Позже Дирак показал конформную инвариантность уравнения для нейтрино (безмассового уравнения Дирака) [?]. В настоящее время принято считать, что конформная инвариантность играет большую роль в асимптотических режимах, где отсутствуют размерные параметры. Подробное изложение конформной квантовой теории поля можно найти в монографии [?]. □

2.9.4 Трехмерное пространство Минковского

Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$ с декартовыми координатами x^a, $a = 0, 1, 2$, и метрикой Лоренца $\{\eta_{ab}\} = \mathsf{diag}\,(+--)$. В трехмерном случае группа Лоренца $\mathbb{O}(1,2)$ имеет всего три независимых генератора: L_{01}, L_{02} и L_{12}. Первые два генератора соответствуют лоренцевым бустам, а третий – пространственным вращениям в плоскости x^1, x^2. Генераторы группы Лоренца удовлетворяют алгебре (2.153), которая в рассматриваемом случае существенно упрощается:

$$\begin{aligned}[L_{01}, L_{02}] &= L_{12},\\ [L_{01}, L_{12}] &= L_{02},\\ [L_{02}, L_{12}] &= -L_{01},\end{aligned}$$

а все остальные коммутаторы равны нулю.

Используя полностью антисимметричный тензор третьего ранга ε^{abc}, $\varepsilon^{012} = 1$, (см. приложение 28.4) удобно перейти к дуальному базису алгебры Ли $\mathfrak{so}(1,2)$

$$M^a = \frac{1}{2}\varepsilon^{abc}L_{bc}, \qquad L_{bc} = \varepsilon_{bca}M^a,$$

где подъем и опускание индексов осуществляется с помощью метрики Лоренца. При этом генератор $M^0 = L_{12}$ соответствует вращениям в пространственной плоскости, а генераторы $M^1 = -L_{02}$ и $M^2 = L_{01}$ – лоренцевым бустам соответственно в плоскостях x^0, x^2 и x^0, x^1. Нетрудно проверить, что новый базис удовлетворяет коммутационным соотношениям

$$[M^a, M^b] = \varepsilon^{abc}M_c.$$

В отличие от алгебры Ли группы трехмерных вращений (2.119), подъем и опускание индексов в рассматриваемом случае осуществляется не с помощью евклидовой метрики, а с помощью метрики Лоренца: $M_a = \eta_{ab}M^b$ ($M_0 = M^0$, $M_{1,2} = -M^{1,2}$).

Группа Лоренца $\mathbb{O}(1,2)$ неабелева и некомпактна.

Собственная ортохронная группа Лоренца $\mathbb{SO}_\uparrow(1,2)$ имеет двумерные неабелевы некомпактные подгруппы. В качестве такой подгруппы можно выбрать группу, генерируемую, например, базисными векторами M_1 и $K = M_0 + M_2$, которые соответствуют бусту и вращению в перпендикулярной нулевой плоскости. Эти векторы образуют базис двумерной неабелевой алгебры Ли $\mathfrak{g} \subset \mathfrak{so}(1,2)$ и удовлетворяют следующим коммутационным соотношениям

$$\begin{aligned}[M_1, M_1] = 0, \qquad [K, K] = 0,\\ [M_1, K] = K.\end{aligned}$$

Это алгебра генерирует простейшую двумерную неабелеву группу Ли $\mathbb{G}$. Ее групповое многообразие некомпактно и подробно изучено в разделе 9.7. С топологической

точки зрения эта подгруппа диффеоморфна евклидовой плоскости $\mathbb{R}^2$. Любое преобразование из группы $\mathbb{SO}_\uparrow(1,2)$ можно представить в виде композиции некоторого преобразования из данной подгруппы и некоторого пространственного вращения из подгруппы $\mathbb{SO}(2) \subset \mathbb{SO}_\uparrow(1,2)$. Поэтому в соответствии с общей теоремой **??** как многообразие трехмерная группа Лоренца диффеоморфна прямому произведению

$$\mathbb{SO}_\uparrow(1,2) \approx \mathbb{S}^1 \times \mathbb{R}^2,$$

так как $\mathbb{SO}(2) \approx \mathbb{S}^1$. Этот диффеоморфизм описан явно в примере 9.12.8.

Поскольку трехмерная группа Лоренца $\mathbb{SO}_\uparrow(1,2)$ является подгруппой всех групп Лоренца более высоких размерностей $\mathbb{SO}_\uparrow(1,n-1)$, $n>3$, то соответствующая двумерная подгруппа является подгруппой всех групп Лоренца $\mathbb{G} \subset \mathbb{SO}_\uparrow(1,n-1)$, $n \geq 3$.

Построим гомоморфизм группы вещественных 2×2 матриц с единичным определителем $\mathbb{SL}(2,\mathbb{R})$ на связную компоненту единицы трехмерной группы Лоренца $\mathbb{SO}_\uparrow(1,2)$. Для этого рассмотрим произвольную симметричную 2×2 матрицу A (поскольку матрица вещественна, то она также эрмитова), которую параметризуем следующим образом

$$A = \begin{pmatrix} t+x & y \\ y & t-x \end{pmatrix}, \qquad t,x,y \in \mathbb{R}. \tag{2.171}$$

Рассмотрим числа t,x,y как декартовы координаты в пространстве Минковского $\mathbb{R}^{1,2}$. Поскольку

$$\mathsf{det}\, A = t^2 - x^2 - y^2,$$

то определитель $\mathsf{det}\, A$ задает квадратичную форму Лоренца в $\mathbb{R}^{1,2}$.

Пусть $M \in \mathbb{SL}(2,\mathbb{R})$ – произвольная вещественная 2×2 матрица с единичным определителем. Тогда каждому преобразованию

$$A \mapsto A' = MAM^T \tag{2.172}$$

можно сопоставить некоторое преобразование Лоренца в пространстве Минковского. Действительно, при таком преобразовании матрица A' остается бесследовой, и ее можно представить в виде

$$A' = \begin{pmatrix} t'+x' & y' \\ y' & t'-x' \end{pmatrix}$$

с некоторыми новыми координатами t',x',y'. Поскольку при преобразовании (2.172) определитель не меняется

$$\mathsf{det}\, A' = \mathsf{det}\, A,$$

то каждому элементу $M \in \mathbb{SL}(2,\mathbb{R})$ однозначно ставится в соответствие элемент из группы Лоренца $\mathbb{SO}_\uparrow(1,2)$. При этом групповые операции, как нетрудно видеть, согласованы. Обратно, по заданному преобразованию Лоренца уравнение (2.172) определяет матрицу M, имеющую единичный определитель, с точностью до умножения на -1. Поскольку группа $\mathbb{SL}(2,\mathbb{R})$ связна, то это устанавливает изоморфизм групп

$$\mathbb{SO}_\uparrow(1,2) \simeq \frac{\mathbb{SL}(2,\mathbb{R})}{\mathbb{Z}_2}. \tag{2.173}$$

Отображение $\mathbb{SL}(2,\mathbb{R}) \to \mathbb{SO}_\uparrow(1,2)$ является двулистным накрытием. Группа $\mathbb{SL}(2,\mathbb{R})$ как многообразие диффеоморфна прямому произведению $\mathbb{S}^1\times\mathbb{R}^2$ (см. пример 9.12.8). Поэтому она является связной, но не односвязной. Следовательно, накрытие (2.173)

не является универсальным. Универсальная накрывающая для групп $\mathbb{SO}_\uparrow(1,2)$ и $\mathbb{SL}(2,\mathbb{R})$ построена в примере 9.12.8.

Построим явную параметризацию элементов собственной ортохронной группы Лоренца $\mathbb{SO}_\uparrow(1,2)$ элементами ее алгебры (экспоненциальное отображение). Для этого рассмотрим произвольный элемент алгебры $A \in \mathfrak{su}(2)$ (2.127)

$$A = \omega^k \frac{i}{2}\sigma_k = \frac{i}{2}\begin{pmatrix} \omega^3 & \omega^1 - i\omega^2 \\ \omega^1 + i\omega & -\omega^3 \end{pmatrix}$$

где σ_k, $k = 1,2,3$, – матрицы Паули. Положим

$$\omega^1 = -2ix, \quad \omega^2 = -2t, \quad \omega^3 = -2iy, \tag{2.174}$$

где t, x, y – некоторые вещественные числа, тогда матрица A примет вид

$$A = \begin{pmatrix} y & x - t \\ x + t & -y \end{pmatrix}, \quad t, x, y \in \mathbb{R}.$$

Здесь мы используем те же буквы, что и в параметризации (2.171), но они имеют другой смысл. Полученная матрица вещественна, симметрична и имеет нулевой след. То есть представляет собой элемент алгебры $\mathfrak{sl}(2,\mathbb{R})$ общего вида, т.к. параметризуется тремя параметрами.

Поскольку экспонента от произвольной комплексной $n \times n$ матрицы равномерно сходится в любой ограниченной области комплексного пространства $\mathbb{C}^{n^2}$, то экспоненциальной отображение (2.132) можно переписать для группы $\mathbb{SL}(2,\mathbb{R})$, просто произведя замену координат (2.174). Поскольку

$$\mathsf{sin}\,(ia) = i\,\mathsf{sh}\,a, \qquad \mathsf{cos}\,(ia) = \mathsf{ch}\,a, \qquad a \in \mathbb{R},$$

то преобразованные матрицы выглядят по-разному в различных точках пространства Минковского $(t,x,y) \in \mathbb{R}^{1,2}$. Внутри светового конуса $t^2 > x^2 + y^2$ экспоненциальное отображение имеет вид

$$M = \begin{pmatrix} \mathsf{cos}\,s + \frac{y}{s}\,\mathsf{sin}\,s & \frac{x-t}{s}\,\mathsf{sin}\,s \\ \frac{x+t}{s}\,\mathsf{sin}\,s & \mathsf{cos}\,s - \frac{y}{s}\,\mathsf{sin}\,s \end{pmatrix}, \tag{2.175}$$

где $s = \sqrt{t^2 - x^2 - y^2}$. На световом конусе $t^2 = x^2 + y^2$

$$M = \begin{pmatrix} 1 + y & x - t \\ x + y & 1 - y \end{pmatrix}. \tag{2.176}$$

Вне светового конуса

$$M = \begin{pmatrix} \mathsf{ch}\,r + \frac{y}{r}\,\mathsf{sh}\,r & \frac{x-t}{r}\,\mathsf{sh}\,r \\ \frac{x+t}{r}\,\mathsf{sh}\,r & \mathsf{ch}\,r - \frac{y}{r}\,\mathsf{sh}\,r \end{pmatrix}, \tag{2.177}$$

где $r = \sqrt{x^2 + y^2 - t^2}$. Матрицы (2.175) и (2.177) имеют одинаковый предел при $s \to 0$ и $r \to 0$, который совпадает с матрицей (2.176).

Экспоненциальное отображение не является однозначным, и многим элементам алгебры соответствует один и тот же элемент группы. То есть групповое многообразие $\mathbb{SO}_\uparrow(1,2)$ получается из $\mathbb{R}^3$ после отождествления некоторых точек. Например, единице группы соответствует начало координат и все двуполостные гиперболоиды

$$s = 2\pi m, \qquad m = \pm 1, \pm 2, \ldots .$$

Из вида матрицы (2.175) следует, что внутри светового конуса существует отношение эквивалентности

$$x^a \sim x^a + 2\pi\frac{x^a}{s}.$$

Как и в случае группы $\mathbb{SU}(2)$ существует второе отношение эквивалентности. А именно, необходимо отождествить также точки, лежащие на гиперболоиде $s = \pi$, т.к. всем этим точкам соответствует одна и та же матрица $-\mathbb{1}$.

2.9.5 Четырехмерное пространство Минковского

Обсудим ряд специфических свойств группы Лоренца в четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$, которые часто используются в физических приложениях.

Предложение 2.9.5. *В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ любая дифференцируемая изотропная кривая $\{x^a(s)\}$, определяемая уравнением*

$$\eta_{ab}\dot{x}^a\dot{x}^b = 0, \tag{2.178}$$

представима в виде

$$\begin{aligned} x^0 &= \int^s duR, & x^2 &= \int^s duR\sin\theta\sin\varphi, \\ x^1 &= \int^s duR\sin\theta\cos\varphi, & x^3 &= \int^s duR\cos\theta, \end{aligned} \tag{2.179}$$

где каждая из координат определена с точностью до постоянной, а $R(u)$, $\theta(u)$ и $\varphi(u)$ – произвольные непрерывные функции от u. Эти линии являются прямыми тогда и только тогда, когда функции R, θ, φ постоянны.

Доказательство. Явное решение алгебраического уравнения (2.178) относительно производных и последующее интегрирование. □

Отсюда следует, что, в отличии от плоскости Минковского, класс изотропных кривых состоит не только из прямых линий. Поскольку все экстремали (см. главу 17) в пространстве Минковского и только они являются прямыми, то это доказывает, что не всякая изотропная кривая является экстремалью. Исключение составляет только двумерное пространство Минковского, где все изотропные кривые являются прямыми и, следовательно, экстремалями.

Предложение 2.9.6. *Любая изотропная кривая в пространстве Минковского, проходящая через точку y, не может выйти за пределы светового конуса с вершиной в этой точке.*

Доказательство. Допустим, что в некоторой точке y_1, лежащей на световом конусе будущего и не совпадающей с y, изотропная кривая покидает световой конус (см. рис.2.10). Тогда в некоторой окрестности данной точки условие (2.178) будет нарушено. Это ясно из того, что световой конус будущего с вершиной в точке y_1 касается светового конуса будущего в точке y и целиком лежит внутри него. □

Четырехмерное пространство Минковского, $\mathbb{R}^{1,3} = \mathbb{R}^{1,1} \times \mathbb{R}^2$, можно представить как прямое произведение двумерной плоскости Минковского $\mathbb{R}^{1,1}$ и евклидовой плоскости $\mathbb{R}^2$ с координатами $(x^0,\ x^1)$ и $(x^2,\ x^3)$ соответственно. При таком разбиении

можно ввести координаты светового конуса (2.82) на $\mathbb{R}^{1,1}$ и комплексные координаты (2.76) на $\mathbb{R}^2$. Тогда лоренцев интервал примет вид

$$ds^2 = dudv - dzd\bar{z}. \tag{2.180}$$

Соответствующая метрика в координатах $\{x^a\} = \{u, v, z, \bar{z}\}$ имеет вид

$$\eta = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

В разделе **??** будет введена комплексная изотропная тетрада, которая соответствует введенным выше координатам в касательном пространстве к пространству-времени.

Коммутационные соотношения, определяющие алгебру Пуанкаре (2.153)–(2.155), в четырехмерном пространстве-времени можно переписать в другом, эквивалентном, виде, выделив явно пространственные и временные компоненты. Среди координат пространства Минковского, как и ранее, явно выделяем время $\{x^a\} = \{x^0, x^i\}$, $i = 1, 2, 3$. Группа Пуанкаре содержит подгруппу трехмерных вращений, алгебра которой определяется генераторами L_{ij}. В связи с этим вместо генераторов L_{ab} можно рассматривать 6 генераторов

$$J_i = \frac{1}{2}\varepsilon_i{}^{jk}L_{jk}, \qquad N_i = L_{0i},$$

где подъем и опускание латинских индексов $i, j = 1, 2, 3$ осуществляется с помощью евклидовой метрики δ_{ij}. Образующие алгебры Ли группы Лоренца J_i и N_i представляют собой генераторы пространственных вращений и бустов. Нетрудно проверить, что они удовлетворяют следующим коммутационным соотношениям

$$\begin{aligned} [J_i, J_j] &= -\varepsilon_{ij}{}^k J_k, \\ [N_i, N_j] &= \ \ \varepsilon_{ij}{}^k J_k, \\ [J_j, N_j] &= -\varepsilon_{ij}{}^k N_k. \end{aligned} \tag{2.181}$$

Последнее коммутационное соотношение показывает, что базисные векторы N_i преобразуются как компоненты ковектора при пространственных вращениях. Остальные коммутационные соотношения (2.153), (2.154) примут вид

$$\begin{aligned} [J_i, P_0] &= 0, \\ [J_i, P_j] &= -\varepsilon_{ijk}P^k, \\ [N_i, P_0] &= P_i, \\ [N_i, P_j] &= \delta_{ij}P_0. \end{aligned} \tag{2.182}$$

Построим гомоморфизм группы комплексных 2×2-матриц с единичным определителем $\mathbb{SL}(2, \mathbb{C})$ на связную компоненту единицы группы Лоренца $\mathbb{SO}_\uparrow(1, 3)$. Это построение аналогично доказательствам существования накрытий $\mathbb{SU}(2) \to \mathbb{SO}(3)$ и $\mathbb{SL}(2, \mathbb{R}) \to \mathbb{SO}_\uparrow(1, 2)$, проведенным в разделах 2.8 и 2.9.4. Для построения гомоморфизма заметим, что произвольную эрмитову матрицу $A = A^\dagger$ можно представить в виде

$$A = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} = t + x\sigma_1 + y\sigma_2 + z\sigma_3, \qquad t, x, y, z \in \mathbb{R}, \tag{2.183}$$

Рассмотрим числа t, x, y, z как координаты в пространстве-времени Минковского $\mathbb{R}^{1,3}$. Очевидно, что

$$\mathsf{det}\, A = t^2 - x^2 - y^2 - z^2.$$

Пусть $M \in \mathbb{SL}(2,\mathbb{C})$ – произвольная комплексная матрица с единичным определителем. Тогда каждому преобразованию

$$A \mapsto A' = MAM^\dagger, \tag{2.184}$$

можно однозначно поставить в соответствие некоторое собственное ортохронное преобразование Лоренца. Действительно, при таком преобразовании матрица A' остается эрмитовой, и ее можно записать в виде

$$A' = \begin{pmatrix} t'+z' & x'-iy' \\ x'+iy' & t'-z' \end{pmatrix}$$

с некоторыми новыми координатами t', x', y', z' в $\mathbb{R}^{1,3}$. Поскольку при этом определитель матрицы не меняется,

$$\mathsf{det}\, A' = \mathsf{det}\, A,$$

то каждому элементу $M \in \mathbb{SL}(2,\mathbb{C})$ однозначно ставится в соответствие элемент из группы Лоренца $\mathbb{SO}_\uparrow(1,3)$. Прямая проверка показывает, что отображение

$$\mathbb{SL}(2,\mathbb{C}) \to \mathbb{SO}_\uparrow(1,3) \tag{2.185}$$

является гомоморфизмом групп. Обратно, по заданному преобразованию Лоренца уравнение (2.184) определяет матрицу M с точностью до умножения на -1, причем отображение (2.185) сюръективно [**?**]. Поскольку группа $\mathbb{SL}(2,\mathbb{C})$ связна, то это устанавливает изоморфизм групп

$$\mathbb{SO}_\uparrow(1,3) \simeq \frac{\mathbb{SL}(2,\mathbb{C})}{\mathbb{Z}_2}. \tag{2.186}$$

При этом тождественному преобразованию Лоренца соответствуют две матрицы $\mathbb{1}$ и $-\mathbb{1}$ из группы $\mathbb{SL}(2,\mathbb{C})$. Можно доказать [], что группа $\mathbb{SL}(2,\mathbb{C})$ является односвязной. Это значит, что построенное отображение представляет собой двулистное универсальное накрытие (см. раздел 12). Таким образом доказана

Теорема 2.9.2. *Существует гомоморфизм групп (2.185), который представляет собой двулистное универсальное накрытие.*

Определение. Точное представление универсальной накрывающей $\mathbb{SL}(2,\mathbb{C})$ называется *спинорным* представлением собственной ортохронной группы Лоренца $\mathbb{SO}_\uparrow(1,3)$. При этом каждому вращению $S \in \mathbb{SO}_\uparrow(1,3)$ ставится в соответствие неупорядоченная пара элементов $\pm M \in \mathbb{SL}(2,\mathbb{C})$. □

Избавиться от неоднозначности спинорного представления, т.е. упорядочить пару элементов $\pm M$, нельзя, т.к. при тождественном преобразовании (пространственное вращение на угол 2π) матрица M изменит знак. Поэтому говорят, что оно является двузначным представлением собственной ортохронной группы Лоренца. Таким образом, спинорное представление связной компоненты единицы группы Лоренца $\mathbb{SO}_\uparrow(1,3)$ является комплексным двумерным. Спинорные представления полной группы Лоренца $\mathbb{O}(1,3)$ являются комплексными четырехмерными, и будут рассмотрены в разделе **??**.

Собственные вращения $\mathbb{SO}(3)$ в трехмерном евклидовом пространстве, определяемом уравнением $t = 0$ в пространстве Минковского, образуют подгруппу группы Лоренца $\mathbb{SO}_\uparrow(1,3)$. Пусть $\tilde{S} \in \mathbb{SO}(3)$. Тогда каждой матрице вращений $\tilde{S}$ соответствуют две матрицы $\pm U$ из группы специальных унитарных матриц $\mathbb{SU}(2)$, которая является подгруппой в $\mathbb{SL}(2,\mathbb{C})$. Это следует из параметризации (2.183). Заметим, что для преобразований (2.184) с унитарной матрицей $U \in \mathbb{SU}(2) \subset \mathbb{SL}(2,\mathbb{C})$ сохраняется не только определитель, но и след, $\mathsf{tr}\, A' = tr A$. Так как $\mathsf{tr}\, A = 2t$, то каждому преобразованию из подгруппы унитарных унимодулярных матриц ставится в соответствие такое преобразование Лоренца, которое не затрагивает время, т.е. некоторое вращение $\tilde{S} \in \mathbb{SO}(3)$.

Аналогичное рассуждение применимо к подгруппе вещественных матриц $\mathbb{SL}(2,\mathbb{R}) \subset \mathbb{SL}(2,\mathbb{C})$. Если $M \in \mathbb{SL}(2,\mathbb{R})$, то $M^T = M^\dagger$, и преобразование (2.184) переходит в $A' = MAM^T$, которое отображает вещественные симметричные матрицы на себя. Это соответствует таким преобразованиям пространства Минковского, которые не затрагивает координату y в параметризации (2.183). Следовательно, каждой вещественной унимодулярной матрице M из $\mathbb{SL}(2,\mathbb{R})$ ставится в соответствие некоторое собственное ортохронное преобразование Лоренца из $\mathbb{SO}_\uparrow(1,2)$, которое действует только на координаты t, x, z в пространстве Минковского.

Посмотрим, что представляют из себя группа Лоренца $\mathbb{SO}_\uparrow(1,3)$ и ее универсальная накрывающая $\mathbb{SL}(2,\mathbb{C})$ как многообразия.

Предложение 2.9.7. *Группа специальных комплексных* 2×2 *матриц* $\mathbb{SL}(2,\mathbb{C})$ *и собственная ортохронная группа Лоренца* $\mathbb{SO}_\uparrow(1,3)$ *как многообразия представляют собой прямые произведения:*

$$\mathbb{SL}(2,\mathbb{C}) \approx \mathbb{S}^3 \times \mathbb{R}^3. \tag{2.187}$$

$$\mathbb{SO}_\uparrow(1,3) \approx \mathbb{RP}^3 \times \mathbb{R}^3. \tag{2.188}$$

Доказательство. Согласно полярному разложению матриц (теорема 28.1.7) произвольная матрица $M \in \mathbb{SL}(2,\mathbb{C})$ однозначно представима в виде

$$M = RU, \tag{2.189}$$

где $R = \sqrt{MM^\dagger}$ – положительно определенная эрмитова матрица и U – унитарная матрица. Поскольку $\mathsf{det}\, M = 1$, то выполнены равенства $\mathsf{det}\, R = 1$ и $\mathsf{det}\, U = 1$. Следовательно, $U \in \mathbb{SU}(2)$. Множество эрмитовых 2×2 матриц с единичным определителем имеет вид (2.183) с единственным ограничением

$$\mathsf{det}\, R = t^2 - x^2 - y^2 - z^2 = 1.$$

Поскольку матрица R положительно определена, то нам необходимо выбрать верхнюю полу однополостного гиперболоида, соответствующую значениям $t > 0$. Она, как многообразие, диффеоморфна трехмерному евклидову пространству $\mathbb{R}^3$. Ранее было установлено, что многообразие группы $\mathbb{SU}(2)$ представляет собой сферу $\mathbb{S}^3$. Таким образом построен диффеоморфизм (2.187).

Собственная ортохронная группа Лоренца $\mathbb{SO}_\uparrow(1,3)$ возникает после отождествления $M \sim -M$. В полярном разложении (2.189) это означает отождествление унитарных матриц $U \sim -U$, т.к. матрица R положительно определена. Это отождествление соответствует накрытию трехмерного проективного пространства $\mathbb{RP}^3$ трехмерной сферой $\mathbb{S}^3$ (см. раздел 2.8). Следовательно, существует диффеоморфизм (2.188). $\square$

Существование диффеоморфизмов (2.187) и (2.188) согласуется с общей теоремой 9.11.13. Тем самым мы показали, что подгруппы $\mathbb{SU}(2)$ и $\mathbb{SO}(3)$ являются максимальными компактными подгруппами в $\mathbb{SL}(2,\mathbb{C})$ и $\mathbb{SO}_\uparrow(1,3)$ соответственно.

Определим оператор пространственного отражения (четности) в пространстве Минковского как отражение всех пространственных координат

$$P: \quad (x^0,x^1,x^2,x^3)\mapsto(x^0,-x^1,-x^2,-x^3).$$

Тогда внешние автоморфизмы (2.147), вызванные обращением времени и преобразованием четности, совпадают, т.к. $P=-T$. При инверсии пространственных координат матрица преобразований Лоренца меняется по правилу (2.148):

$$P: \quad S\mapsto PSP^{-1}=(S^T)^{-1}.$$

Пусть пространственному вращению $\tilde S\in\mathbb{SO}(3)$ соответствуют две комплексные матрицы $\pm\tilde A\in\mathbb{SU}(2)$. Тогда внешнему автоморфизму (2.148) соответствует автоморфизм в группе $\mathbb{SL}(2,\mathbb{C})$

$$P: \quad \pm A \;\mapsto\; \pm(A^T)^{-1}.$$

Последний будет внутренним, т.к. группа $\mathbb{SL}(2,\mathbb{C})$ связна. Это значит, что представление группы $\mathbb{SL}(2,\mathbb{C})$ является спинорным представлением не только для связной компоненты единицы $\mathbb{SO}_\uparrow(1,3)$, но и группы $\mathbb{O}_\uparrow(1,3)$, включающей преобразование четности.

Группа Лоренца является некомпактной, ее представления могут быть неунитарны, а среди ее неприводимых представлений есть бесконечномерные. Известно, что все конечномерные неприводимые представления, за исключением тривиального единичного, группы Лоренца неунитарны [?]. Это существенно отличает ее от группы вращений в евклидовом пространств и представляет существенные трудности при построении физических моделей, где требуется положительная определенность энергии и скалярного произведения в гильбертовом пространстве соответствующей квантовой теории. Из изоморфизма (2.186) следует, что точное представление группы $\mathbb{SL}(2,\mathbb{C})$ является двузначным или спинорным представлением группы $\mathbb{SO}_\uparrow(1,3)$. При этом представление собственной ортохронной группы Лоренца $\mathbb{SO}_\uparrow(1,3)$ однозначно или двузначно одновременно с порожденным им представлением группы вращений $\mathbb{SO}(3)$.

Построим параметризацию группы $\mathbb{SL}(2,\mathbb{C})$ элементами ее алгебры. Алгебра Ли $\mathfrak{sl}(2,\mathbb{C})$ состоит из комплексных 2×2-матриц с нулевым следом (6 вещественных параметров). Алгебру Ли $\mathfrak{sl}(2,\mathbb{C})$ можно рассматривать как трехмерное векторное пространство над полем комплексных чисел. В качестве базиса выберем матрицы Паули σ_i (см. приложение 28.2). Тогда произвольный элемент алгебры Ли $\mathfrak{sl}(2,\mathbb{C})$ параметризуется вектором $\boldsymbol{z}$ в трехмерном комплексном пространстве

$$\boldsymbol{z}=z^k\sigma_k,\qquad \boldsymbol{z}=\{z^k\}\in\mathbb{C}^3,\quad k=1,2,3.$$

Предложение 2.9.8. *Каждому элементу алгебры Ли* $\boldsymbol{z}=\{z^k\}\in\mathfrak{sl}(2,\mathbb{C})\approx\mathbb{C}^3$ *соответствует элемент группы Ли* $S=\{S_{\text{A}}{}^{\text{B}}\}\in\mathbb{SL}(2,\mathbb{C})$

$$\begin{aligned}S_{\text{A}}{}^{\text{B}}(\boldsymbol{z})=\left(\mathrm{e}^{iz^k\sigma_k/2}\right)_{\text{A}}{}^{\text{B}}=\delta^{\text{B}}_{\text{A}}\cos\frac{\sqrt{\boldsymbol{z}^2}}{2}+i\frac{z^k\sigma_{k\text{A}}{}^{\text{B}}}{\sqrt{\boldsymbol{z}^2}}\sin\frac{\sqrt{\boldsymbol{z}^2}}{2}=\\ =\begin{pmatrix}\cos\frac{\sqrt{\boldsymbol{z}^2}}{2}+i\frac{z^3}{\sqrt{\boldsymbol{z}^2}}\sin\frac{\sqrt{\boldsymbol{z}^2}}{2} & i\frac{z^1-iz^2}{\sqrt{\boldsymbol{z}^2}}\sin\frac{\sqrt{\boldsymbol{z}^2}}{2}\\ i\frac{z^1+iz^2}{\sqrt{\boldsymbol{z}^2}}\sin\frac{\sqrt{\boldsymbol{z}^2}}{2} & \cos\frac{\sqrt{\boldsymbol{z}^2}}{2}-i\frac{z^3}{\sqrt{\boldsymbol{z}^2}}\sin\frac{\sqrt{\boldsymbol{z}^2}}{2}\end{pmatrix},\end{aligned}\tag{2.190}$$

где $\boldsymbol{z}^2:=(z^1)^2+(z^2)^2+(z^3)^2$ *– комплексной число.*

Доказательство. Формальное переписывание (2.132), т.к. все разложения равномерно сходятся в произвольном ограниченном шаре в $\mathbb{C}^3$. □

Многообразие группы $\mathbb{SL}(2,\mathbb{C})$ возникает после отождествления точек в $\mathbb{C}^3$, т.к. некоторым элементам алгебры соответствует один и тот же элемент группы. В рассматриваемом случае это отождествление выглядит громоздко, и мы не будем его проводить.

2.10 Специальная теория относительности

Специальная теория относительности – это учение о пространстве-времени, в котором строятся современные модели всех взаимодействия за исключением гравитационных. В двух словах суть специальной теории относительности состоит в утверждении, что пространство-время представляет собой пространство Минковского $\mathbb{R}^{1,3}$, в котором определена метрика Лоренца $\eta_{ab} = \mathsf{diag}\,(c^2 - - -)$, где c – скорость света. В настоящем разделе мы будем писать скорость света явно, не полагая ее равной единице, чтобы быть ближе к стандартному изложению специальной теории относительности. При этом координаты в системе СГС имеют следующие размерности:

$$[x^0] = \text{сек}, \qquad [x^\mu] = \text{см}, \qquad \mu = 1,2,3$$

Мы начнем с описания основных свойств нерелятивистских моделей, а затем обсудим некоторые свойства пространства Минковского, которые затем будут обобщены в теории тяготения Эйнштейна. Кроме того, рассмотренные примеры из специальной теории относительности помогают развить определенную интуицию, которая полезна при исследовании многообразий общего вида, на которых задана метрика лоренцевой сигнатуры.

2.10.1 Нерелятивистские модели

До создания специальной теории относительности модели математической физики были нерелятивистскими и строились в пространстве-времени, в котором время играет выделенную роль.

Определение. *Галилеевым* пространством-временем называется топологически тривиальное четырехмерное многообразие $\mathbb{M} = \mathbb{R} \times \mathbb{R}^3$, где первый сомножитель $\mathbb{R}$ соответствует *времени* t, а второй $\mathbb{R}^3$ – *пространству*, со следующей структурой. Мы предполагаем, что в пространстве $\mathbb{R}^3$ задана евклидова метрика. Пусть x^i, $i = 1,2,3$, – декартовы координаты в пространстве $\mathbb{R}^3$, в которых метрика имеет диагональный вид $\delta_{ij} = \mathsf{diag}\,(+++)$. Каждая точка пространства-времени $x = \{t, x^i\} \in \mathbb{M}$ называется *событием* и говорит о том, что нечто произошло в момент времени t в точке пространства $\{x^i\}$. События, соответствующие фиксированному значению t, называются *одновременными*. Пусть задано два произвольных события $x_1, x_2 \in \mathbb{M}$ с координатами $x_1 = \{t_1, x_1^i\}$ и $x_2 = \{t_2, x_2^i\}$. Тогда временной интервал между этими событиями определяется разностью

$$\triangle t := t_2 - t_1. \tag{2.191}$$

Если $t_2 > t_1$, то мы говорим, что событие x_2 произошло позже события x_1. В противном случае, $t_2 < t_1$, событие x_2 произошло раньше события x_1. Пространственное

расстояние между двумя одновременными событиями $x_1 = \{t, x_1^i\}$ и $x_2 = \{t, x_2^i\}$ равно евклидову расстоянию в $\mathbb{R}^3$:

$$\triangle x := \sqrt{\delta_{ij}(x_2^i - x_1^i)(x_2^j - x_1^j)}. \tag{2.192}$$

Система координат t, x^i называется *инерциальной* и обозначается O_0. Кроме того, мы предполагаем, что галилеево пространство-время снабжено естественной линейной и аффинной структурой, такой же как и евклидово пространство $\mathbb{R}^4$ в разделах 2.3.3 и 2.3.4. □

Одновременные события составляют пространственные сечения пространства-времени $\mathbb{M}$, соответствующие фиксированному значению времени t. Из определения следует, что понятия "раньше" и "позже" не зависят от выбора событий в каждом сечении, т.е. временной интервал определен для двух произвольных сечений t_1 и t_2. Другими словами, на оси времени $t \in R$ задана обычная топологическая евклидова метрика $l(t_1, t_2) = |t_2 - t_1|$, которая определяет расстояние между двумя пространственными сечениями. Единственное отличие – это наличие знака у временно́го интервала, который соответствует понятиям "раньше" и "позже".

Подчеркнем, что пространственное расстояние можно определить только для одновременных событий. Если две системы координат движутся друг относительно друга с ненулевой скоростью и события не являются одновременными, то, как легко видеть, пространственное расстояние (2.192) между этими событиями, измеренное в этих системах координат, будет различным.

Данное выше определение галилеева пространства-времени содержит одну инерциальную систему отсчета O_0. Расширим это понятие. Временно́й интервал между двумя произвольными событиями и пространственное расстояние между двумя одновременными событиями инвариантны относительно преобразований Галилея (2.162). Обратное утверждение в общем случае неверно.

Обсудим это. Сохранение временно́го интервала оставляет возможность только для сдвигов вдоль оси времени. Эти сдвиги не могут зависеть от точки пространства, т.к. мы требуем сохранение интервала (2.191) не для отдельных событий, а для сечений.

Пространственная метрика (2.192) инвариантна только относительно преобразований из неоднородной группы вращений $\mathbb{IO}(3)$. В общем случае преобразований пространства-времени $\mathbb{R} \times \mathbb{R}^3$ и матрица вращений, и вектор сдвига пространства могут зависеть от времени, причем произвольно. Однако, если ограничиться линейными преобразованиями, то остаются только преобразования из группы Галилея. Снабдим галилеево пространство-время обычной структурой линейного пространства из $\mathbb{R}^4$. Тогда два произвольных различных события определяют единственную прямую (2.25), проходящую через эти события. При преобразованиях Галилея линейная структура сохраняется и прямые переходят в прямые. Верно также обратное утверждение: любое преобразование пространства-времени, которое сохраняет временно́й интервал (2.191), пространственную метрику (2.192) и линейную структуру пространства-времени, называется преобразованием Галилея.

Определение. Любая система координат в галилеевом пространстве-времени $\mathbb{M} = \mathbb{R} \times \mathbb{R}^3$, связанная с системой координат O_0 преобразованием из группы Галилея, называется *инерциальной*. □

Точечные частицы движутся в пространстве по траекториям $\{x^i(t)\} \in \mathbb{R}^3$. Эти три функции задают мировую линию частицы в пространстве-времени $\{t, x^i(t)\} \in \mathbb{M}$.

Мировые линии частиц вводятся точно таким же образом и в специальной, и в общей теории относительности. Различие заключается лишь в том, что время в нерелятивистской механике время имеет абсолютный характер, и его естественно выбрать в качестве параметра вдоль мировых линий всех частиц. Такой выбор общепринят и нагляден с физической точки зрения.

Определение. Если в некоторой инерциальной системе координат мировая линия частицы представляет собой прямую линию вида $\{t, x^i = v^i t + x^i_0\} \in \mathbb{M}$, где v^i, x^i_0 – некоторые постоянные, то такая частица называется *свободной*. Точка $\{x^i_0\} \in \mathbb{R}^3$ является точкой пространства, в которой частица расположена в начальный момент времени $t = 0$, а векторное поле $\boldsymbol{v} = \{v^i\} \in \mathbb{T}(\mathbb{R}^3)$, определенное вдоль траектории частицы, называется *наблюдаемой скоростью*. □

Из определения инерциальной системы координат следует, что, если мировая линия частицы является прямой в какой либо инерциальной системе отсчета, то она будет прямой и в любой другой инерциальной системе.

Замечание. В физической литературе понятие инерциальной системы координат часто определяют следующим образом. Систему координат называют инерциальной, если свободная частица движется в ней равномерно и прямолинейно. В таком определении не сказано, что значит свободная частица. Поэтому мы пошли другим путем. Сначала было построено пространство-время и инерциальные системы отсчета. Это позволило дать определение свободной частицы. □

Замечание. В пространстве-времени $\mathbb{M}$ не существует метрики, инвариантной относительно преобразований Галилея, потому что эта группа не является полупростой. Поэтому говорить про экстремали не имеет смысла. В галилеевом пространстве-времени заданы естественные линейная и аффинная структуры, и параллельный перенос мы отождествляем с трансляциями. Это определяет связность на $\mathbb{M}$, которая в инерциальной системе отсчета имеет нулевые компоненты. Эта связность плоская, а геодезические являются прямыми линиями. □

Линейность функций $x^i(t)$ означает, что свободная частица движется в пространстве равномерно и прямолинейно. Это утверждение известно как *первый закон Ньютона*. Мировая линия свободной частицы в пространстве-времени в любой инерциальной системе координат представляет собой прямую линию. Мы говорим, что частица не является свободной, если ее мировая линия отличается от прямой. В этом случае траектория частицы имеет ненулевое ускорение $\boldsymbol{a} = d\boldsymbol{v}/dt$. Мы говорим, что на частицу действует сила $\boldsymbol{F}$, которая, по-определению, пропорциональна ускорению:

$$\boldsymbol{F} = m\boldsymbol{a}, \tag{2.193}$$

где коэффициент пропорциональности m называется *инертной массой*. Это соотношение называется *вторым законом Ньютона*.

Если имеется две взаимодействующих частицы, то между ними действует сила. Пусть $\boldsymbol{F}$ – сила, действующая на первую частицу со стороны второй. Тогда сила, действующая на вторую частицу со стороны первой, равна $-\boldsymbol{F}$. Это утверждение составляет *третий закон Ньютона*. Коротко говорят: “Сила действия равна силе противодействия”.

Для механики Галилея справедливо следующее правило сложения скоростей. Если частица движется относительно инерциальной системы координат с постоянной

скоростью $\boldsymbol{v} = \{v^i\}$, то ее скорость движения относительно другой инерциальной системы координат, движущейся со скоростью $\boldsymbol{V} = \{V^i\}$ относительно первой, равна разности $\boldsymbol{v} - \boldsymbol{V}$.

Модели математической физики, которые инвариантны или ковариантны относительно преобразований Галилея называются *нерелятивистскими*.

Пример 2.10.1. В теории тяготения Ньютона две точечные частицы с массами m_1 и m_2, находящиеся соответственно в точках $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$, испытывают гравитационное притяжение. При этом сила, действующая со стороны второй частицы на первую, равна

$$\boldsymbol{F}_{12} = G\frac{m_1 m_2}{\boldsymbol{r}^2}\frac{\boldsymbol{r}}{|\boldsymbol{r}|}, \tag{2.194}$$

где G – *гравитационная постоянная*, $\boldsymbol{r} = \boldsymbol{x}_2 - \boldsymbol{x}_1$ и $|\boldsymbol{r}|$ – расстояние между частицами. Выражение (2.194) называется законом *всемирного тяготения* или *законом тяготения Ньютона*. Теория тяготения Ньютона является нерелятивистской моделью, в которой время имеет абсолютный характер. Она инвариантна относительно преобразований Галилея, поскольку сила гравитационного взаимодействия двух тел в каждый момент времени зависит только от их масс и расстояния между ними. Массы частиц m_1 и m_2 называются *гравитационными массами*. В механике Ньютона инертная и гравитационная массы частиц считаются равными. В этом случае из закона всемирного тяготения и второго закона Ньютона (2.193) следует, что ускорение, которая испытывает частица массы m_1 в поле тяжести частицы m_2, не зависит от ее массы. Это следствие проверено в поле тяжести Земли и Солнца с высокой степенью точности. □

В классической механике взаимодействие частиц описывается энергией потенциального взаимодействия, которая является функцией от координат частиц. Тем самым изменение положения одной из взаимодействующих частиц отразится на других частицах в тот же момент времени. Это соответствует предположению о том, что взаимодействие распространяется с бесконечной скоростью. Опыт, однако, показывает, что мгновенных взаимодействий в природе не существует. Если сдвинуть одну частицу, то это отразится на других частицах только через некоторое время, которое определяется скоростью распространения взаимодействий.

Электромагнитные взаимодействия распространяются со скоростью света

$$c = 2,998 \cdot 10^{10} \text{ см/с} \approx 300000 \text{ км/с}.$$

В 1887 году Майкельсон и Морли [?] экспериментально установили с точностью до 5 км/с, что скорость света постоянна и не зависит от того, движется ли он параллельно траектории Земли или в перпендикулярном направлении. Точность этого результата сейчас доведена до 1 км/с [?].

Постоянство скорости света находится в явном противоречии с правилом сложения скоростей в механике Галилея и с предположением о мгновенности распространения взаимодействий. Это и привело к созданию специальной теории относительности.

2.10.2 Релятивистские модели

Специальная теория относительности изменила наше представление о структуре пространства-времени. В релятивистских моделях отсутствует понятие одновремен-

ных событий и пространственного расстояния между двумя одновременными событиями. Эти обстоятельства являются наиболее трудными в процессе понимания релятивистских моделей, т.к. наша интуиция основана на механике Ньютона.

Определение. *Пространством Минковского* $\mathbb{R}^{1,3}$ называется четырехмерное топологически тривиальное многообразие с заданной лоренцевой метрикой в инерциальной системе отсчета и естественными линейной и аффинной структурами евклидова пространства. □

В специальной теории относительности мы предполагаем, что все законы природы формулируются в пространстве Минковского. Как линейное пространство пространство-время $\mathbb{R}^{1,3}$ такое же, как и в нерелятивистских моделях. С топологической точки зрения это просто прямое произведение, $\mathbb{M} = \mathbb{R}^{1,3} = \mathbb{R} \times \mathbb{R}^3$. Точка пространства-времени, как и раньше, называется событием, и каждая точечная частица движется вдоль своей мировой линии в $\mathbb{M}$. Отличие заключается в том, что вместо понятий одновременности и пространственного расстояния между одновременными событиями, мы постулируем существование инвариантной метрики в пространстве Минковского $\mathbb{R}^{1,3}$. В декартовой системе координат x^a, $a = 0, 1, 2, 3$, лоренцева метрика, по-определению, имеет диагональный вид $\eta_{ab} = \mathsf{diag}\,(c^2, -1, -1, -1)$, где c – скорость света.

Как и в нерелятивистских моделях, в специальной теории относительности вводится понятие инерциальной системы координат. Однако ее определение другое.

Определение. Система координат в пространстве Минковского $\mathbb{R}^{1,3}$, в которой метрика имеет вид $\eta_{ab} = \mathsf{diag}\,(c^2 - - -)$, называется *инерциальной*. Если мировая линия частицы в инерциальной системе координат представляет собой прямую линию, то она называется *свободной*. □

Замечание. Напомним, что прямые линии в пространстве Минковского и только они являются экстремалями. Поэтому в определении свободной частицы прямую линию можно заменить на экстремаль. □

Замечание. Данное определение инерциальной системы координат по своей сути совпадает с определением декартовой системы координат, которое было дано в разделе 2.3.1 в случае евклидовой метрики. □

Инерциальные системы координат определены неоднозначно. Любые две инерциальные системы координат связаны между собой преобразованием из полной группы Пуанкаре, которая включает в себя преобразования Лоренца (собственные и несобственные), сдвиги, преобразование четности и обращение времени (см. раздел 2.9). Обратно, если некоторая система координат связана с инерциальной преобразованием из группы Пуанкаре, то она сама является инерциальной. Это следует из определений группы Пуанкаре и инерциальной системы координат.

Вторым постулатом специальной теории относительности является предположение о том, что все физические явления описываются некоторым набором полей на пространстве Минковского $\mathbb{R}^{1,3}$, которые преобразуются по какому то, возможно, приводимому представлению группы Пуанкаре.

В физической литературе принято формулировать

Принцип относительности. Все физические законы природы (уравнения движения, равновесия и т.д.) инвариантны или ковариантны относительно преобразований из группы Пуанкаре. □

Пример 2.10.2. Уравнение Клейна–Гордона–Фока для скалярного поля (23.3) в пространстве Минковского инвариантно относительно преобразований из группы Пуанкаре. Уравнение Дирака для спинорного поля и уравнения Максвелла для электромагнитного поля в пространстве Минковского ковариантны относительно преобразований из группы Пуанкаре. □

Замечание. Часто принцип относительности формулируют следующим образом: законы природы выглядят одинаково во всех инерциальных системах отсчета. В такой формулировке за кадром остается определение инерциальной системы отсчета и смысл термина “одинаково”. □

Приведенные рассуждения приводит к следующему итогу. В специальной теории относительности имеется три основных постулата:

1. Пространство-время, в котором происходят все окружающие нас явления, представляет собой четырехмерное пространство Минковского $\mathbb{R}^{1,3}$.

2. Модели математической физики строятся из соответствующего набора полей на пространстве Минковского, которые преобразуются по некоторому представлению группы Пуанкаре.

3. Все законы природы описываются некоторыми уравнениями, которые либо инвариантны, либо ковариантны относительно преобразований из группы Пуанкаре.

Мы всегда предполагаем, что точечные частицы, на которые не действуют никакие силы, кроме гравитационных, движутся в пространстве-времени вдоль экстремалей. В специальной теории относительности гравитационные эффекты не учитываются. Поэтому свободные частицы, на которые не действуют никакие силы, движутся в пространстве Минковского вдоль экстремалей. В пространстве Минковского экстремали совпадают с геодезическими и являются прямыми линиями. Поэтому свободные точечные частицы в специальной теории относительности в инерциальной системе координат, как и в нерелятивистских моделях, движутся равномерно и прямолинейно.

Мы говорим, что пространство-время специальной теории относительности однородно и изотропно. Математически это значит, что метрика инвариантна относительно сдвигов из группы Пуанкаре и вращений пространства Минковского, т.е. преобразований Лоренца. Другими словами, физические свойства пространства-времени не зависят от выбора начала отсчета декартовой системы координат и направления.

Понятие инерциальной системы координат является математической абстракцией. Поскольку во вселенной существуют материальные тела, то гравитационное поле не устранимо. Поэтому системы координат, которые используются в экспериментах, можно считать инерциальными только с определенной степенью точности.

Пример 2.10.3. Система координат, покоящаяся относительно Земли, с хорошей степенью точности является инерциальной при изучении электромагнитных явлений. □

Несмотря на то, что инерциальных систем отсчета, строго говоря, в природе не существует, это понятие чрезвычайно важно в современной теоретической физике. Модели математической физики (например, электродинамика), сформулированные

в пространстве Минковского, приводят к следствиям, которые находятся в прекрасном согласии с экспериментами. В настоящее время известно только об одном случае несогласия специальной теории относительности с опытом: нарушение четности в слабых взаимодействиях. Поэтому считается, что специальная теория относительности находится в прекрасном согласии с экспериментом, и все взаимодействия, кроме гравитационных, инвариантны относительно трансляций и преобразований Лоренца.

Наиболее успешные модели математической физики: модели электромагнитных, слабых, сильных (квантовая хромодинамика) взаимодействий, а также их объединения, не претендующие на описание гравитационных взаимодействий, построены в пространстве Минковского. С этой целью выбирается некоторое представление группы группы Пуанкаре и строится инвариантное действие в пространстве Минковского, которое приводит к ковариантным уравнениям движения. В последние десятилетия такой подход оказался самым распространенным и успешным.

Обсудим понятие одновременности и пространственного расстояния в пространстве Минковского. Зафиксируем два события $x_1 = \{t_1, x_1^i\}$ и $x_2 = \{t_2, x_2^i\}$ в $\mathbb{R}^{1,3}$. Если сказать, что эти события разделены временны́м интервалом $\triangle t := t_2 - t_1$, то такое определение будет некорректным, т.к. $\triangle t$ зависит от выбора инерциальной системы отсчета. Это следует прямо из выражения для преобразований координат (2.107) (лоренцевых бустов в плоскости x^0, x^1). Поэтому понятие одновременности в специальной теории относительности отсутствует. Как следствие, нельзя также корректно ввести понятие пространственного расстояния между одновременными событиями, т.к. одновременные события не определены. Вместо этого вводится понятие интервала между двумя событиями

$$\triangle s^2 = c^2(x_2^0 - x_1^0)^2 - (x_2^1 - x_1^1)^2 - (x_2^2 - x_1^2)^2 - (x_2^3 - x_1^3)^2. \tag{2.195}$$

Это определение корректно, т.к. не зависит от выбора инерциальной системы отсчета. Интервал для двух событий может быть положителен, равен нулю или отрицателен. Мы говорим, что два события причинно связаны, если интервал между ними положителен, $\triangle s^2 > 0$. В этом случае существует такая инерциальная система координат, в которой оба события происходят в одной точке. Если свет испущен в точке x_1 и получен в точке x_2, то для этих событий интервал равен нулю, $\triangle s^2 = 0$. Если интервал отрицателен, $\triangle s^2 < 0$, то события являются причинно не связанными. Для этих событий существует такая инерциальная система координат, в которой оба события происходят одновременно.

Рассмотрим две инерциальные системы координат O и O' с декартовыми координатами t, x, y, z и t', x', y', z' соответственно. Пусть начала систем координат совпадают и система координат O' движется со скоростью V относительно O вдоль оси x. Тогда события в этих системах координат связаны между собой преобразованием Лоренца (2.107):

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}, \quad x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}}, \quad y' = y, \quad z' = z. \tag{2.196}$$

Обратные преобразования координат от системы координат O' к O получаются простой заменой $V \to -V$:

$$t = \frac{t' + Vx'/c^2}{\sqrt{1 - V^2/c^2}}, \quad x = \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}, \quad y = y', \quad z = z'. \tag{2.197}$$

Напомним, что эти преобразования координат называются бустами.

Аналогичные формулы преобразования координат справедливы, если система координат O' движется относительно O либо вдоль оси y, либо оси z.

В приложениях иногда необходимо знать выражение для преобразований Лоренца в общем случае, когда система координат O' движется относительно O равномерно и прямолинейно, но в произвольном направлении. Эти формулы легко получить из (2.196) в виде, который инвариантен относительно пространственных вращений. Пусть $\boldsymbol{V} = \{V^i\}$ – скорость движения штрихованной системы координат. Введем векторное обозначение $\boldsymbol{x} = \{x^i\} = \{x, y, z\}$ для пространственных координат событий и аналогичные обозначения в штрихованной системе координат. Введем также обычное евклидово скалярное произведение пространственных векторов:

$$\boldsymbol{V}^2 = -V^i V_i, \qquad (\boldsymbol{x}, \boldsymbol{V}) = -\boldsymbol{x}^i V_i.$$

Знак минус в этих выражениях следует из того, что при опускании пространственных индексов знак компонент меняется: $V_i = V^j \eta_{ji} = -V^i$. Тогда радиус-вектор можно разложить на составляющие, $\boldsymbol{x} = \boldsymbol{x}_\parallel + \boldsymbol{x}_\perp$, которые параллельны и перпендикулярны вектору скорости:

$$\boldsymbol{x}_\parallel = \frac{(\boldsymbol{x}, \boldsymbol{V})\boldsymbol{V}}{\boldsymbol{V}^2}, \qquad \boldsymbol{x}_\perp = \boldsymbol{x} - \frac{(\boldsymbol{x}, \boldsymbol{V})\boldsymbol{V}}{\boldsymbol{V}^2}.$$

Аналогично

$$\boldsymbol{x}'_\parallel = \frac{(\boldsymbol{x}', \boldsymbol{V})\boldsymbol{V}}{\boldsymbol{V}^2}, \qquad \boldsymbol{x}'_\perp = \boldsymbol{x}' - \frac{(\boldsymbol{x}', \boldsymbol{V})\boldsymbol{V}}{\boldsymbol{V}^2}.$$

Из преобразований Лоренца (2.196) следует, что перпендикулярная составляющая радиус-вектора не меняется, а параллельная меняется так же, как координата x. Отсюда вытекает общее правило преобразования координат:

$$t' = \frac{t - (\boldsymbol{x}, \boldsymbol{V})/c^2}{\sqrt{1 - \boldsymbol{V}^2/c^2}}, \qquad \boldsymbol{x}'_\parallel = \frac{\boldsymbol{x}_\parallel - \boldsymbol{V}t}{\sqrt{1 - \boldsymbol{V}^2/c^2}}, \qquad \boldsymbol{x}'_\perp = \boldsymbol{x}_\perp. \tag{2.198}$$

Для полного радиуса-вектора получаем следующую формулу преобразования

$$\boldsymbol{x}' = \boldsymbol{x} + \frac{(\boldsymbol{x}, \boldsymbol{V})\boldsymbol{V}}{\boldsymbol{V}^2}\left(\frac{1}{\sqrt{1 - \boldsymbol{V}^2/c^2}} - 1\right) - \frac{\boldsymbol{V}t}{\sqrt{1 - \boldsymbol{V}^2/c^2}}$$

или, в компонентах,

$$x'^i = x^i - \frac{x^j V_j V^i}{\boldsymbol{V}^2}\left(\frac{1}{\sqrt{1 - \boldsymbol{V}^2/c^2}} - 1\right) - \frac{V^i t}{\sqrt{1 - \boldsymbol{V}^2/c^2}}. \tag{2.199}$$

В общем случае координаты события преобразуются по правилу

$$t' = tS_0{}^0 + x^i S_i{}^0, \qquad x'^i = tS_0{}^i + x^j S_j{}^i,$$

где $\{S_a{}^b\} = \{S_0{}^0, S_i{}^0, S_0{}^i, S_i{}^j\}$ – матрица преобразований Лоренца. Сравнение этого

выражения с формулами (2.198) и (2.199) дает довольно симметричные выражения:

$$
\begin{aligned}
S_0{}^0 &= \frac{1}{\sqrt{1-\boldsymbol{V}^2/c^2}},\\
S_i{}^0 &= \frac{V_i}{c^2\sqrt{1-\boldsymbol{V}^2/c^2}},\\
S_0{}^i &= -\frac{V^i}{\sqrt{1-\boldsymbol{V}^2/c^2}},\\
S_i{}^j &= \delta_i^j - \frac{V_iV^j}{\boldsymbol{V}^2}\left(\frac{1}{\sqrt{1-\boldsymbol{V}^2/c^2}}-1\right).
\end{aligned}
\tag{2.200}
$$

Нетрудно проверить инвариантность лоренцевой метрики η_{ab} относительно этих преобразований:

$$\eta_{ab} = S_a{}^c S_b{}^d \eta_{bd}.$$

Эти формулы для лоренцевых бустов получены в [**?**].

Полученные общие формулы для лоренцевых бустов (2.200) будут использованы в разделе 22.2 при АДМ параметризации репера.

В релятивистской механике скорость света является универсальной постоянной, входящей в метрику Лоренца, и не зависит от выбора инерциальной системы отсчета. Если скорость света стремится к бесконечности, $c \to \infty$, то метрика Лоренца вырождается, а лоренцевы бусты (2.196) стремятся к преобразованиям Галилея.

2.10.3 Замедление времени и лоренцево сокращение

Сначала выведем формулу для замедления времени в движущейся системе отсчета. Пусть часы покоятся в системе координат O' и, следовательно, движутся равномерно и прямолинейно в системе O. Время t', которое показывают часы в той системе координат, где они покоятся, называется собственным временем. Введем для него специальное обозначение $t_0 = t'$. Это время совпадает с длиной мировой линии часов, деленной на c, которая является прямой, параллельной оси времени. Поскольку для часов $x' = \mathsf{const}$, то из первой формулы преобразований Лоренца (2.197) следует, что для произвольных промежутков времени справедливо равенство

$$\triangle t_0 = \triangle t\sqrt{1-V^2/c^2} \quad\Leftrightarrow\quad \triangle t = \frac{\triangle t_0}{\sqrt{1-V^2/c^2}}.$$

Это соотношение показывает, что в движущейся системе координат время течет медленнее. Это свойство называется эффектом *замедления времени*.

Эффект замедления времени в движущейся системе координат породил множество “парадоксов”.

Парадокс близнецов. Допустим, что в инерциальной системе координат в некоторый момент времени родились два близнеца, событие A на рис.2.11. Первый близнец остался на месте, а второго посадили в ракету и отправили в космическое путешествие с большой скоростью. Через некоторое время ракета вернулась, и близнецы встретились снова, событие B. На рисунке мировые линии близнецов помечены цифрами 1 и 2 соответственно. С точки зрения первого близнеца ракета быстро двигалась и из-за замедления времени второй близнец должен оказаться более молодым, чем

он сам. С точки зрения второго близнеца ситуация прямо противоположная: он покоился относительно ракеты, а двигался первый близнец. Поэтому именно первый близнец должен оказаться моложе.

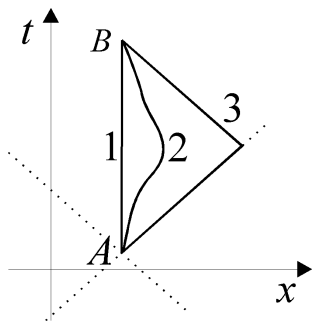


Рис. 2.11: Мировые линии близнецов 1 и 2. Мировая линия первого близнеца "длиннее" мировой линии второго.

Чтобы разъяснить парадокс, часто говорят, что рассуждения второго близнеца неверны, т.к. его система координат не может быть инерциальной (ракета должна замедлить свой полет и повернуть обратно). Однако это замечание ничего не объясняет. Действительно, если оба близнеца отправятся в космические путешествие в разные стороны, а потом вернутся и встретятся. Кто из них будет моложе ? Они оба находились в неинерциальных системах отсчета, и с этой точки зрения совершенно равноправны. Для того, чтобы разрешить парадокс, необходимо дать

Определение. Возрастом человека называется промежуток собственного времени, который прошел с момента рождения. Или, эквивалентно, длина мировой линии человека от момента рождения, деленная на скорость света c. □

Это определение корректно, так как инвариантно относительно произвольных преобразований координат. Поэтому возраст можно вычислить в любой, в том числе неинерциальной системе отсчета.

Теперь все становится на свои места. Поскольку на плоскости x, t задана метрика Лоренца, то длина любого наклонного отрезка между двумя моментами времени, меньше длины вертикального отрезка. Из рисунка ясно, что длина (собственное время) вертикального отрезка между точками A и B, который является мировой линией первого близнеца, максимальна. Поэтому длина любой (времениподобной) мировой линии, соединяющей точки A и B, меньше длины вертикального отрезка. Заметим, что если бы на плоскости x, t была задана евклидова метрика, то длина вертикального отрезка была бы наименьшей среди всех кривых, соединяющих точки A и B.

Заметим, что разница в возрасте может оказаться большой, если ракета движется достаточно быстро. В предельном случае, когда ракета удаляется от первого близнеца со скоростью света, затем отражается и возвращается обратно также со скоростью света, второй близнец за время своего путешествия вообще не состарится (траектория 3 на рисунке). □

Теперь предположим, что в движущейся системе координат O' на оси x' покоится линейка длины $l_0 = x'_2 - x'_1$. Собственная длина линейки l_0 измеряется при постоянном времени $t' = \mathsf{const}$ в сопутствующей системе координат O'. В системе координат O линейка равномерно движется и ее длина измеряется при постоянном времени t. Из второй формулы в (2.196) вытекает соотношение между длинами:

$$l_0 = \frac{l}{\sqrt{1 - V^2/c^2}} \quad \Leftrightarrow \quad l = l_0\sqrt{1 - V^2/c^2}.$$

Отсюда следует, что движущаяся линейка в покоящейся системе координат имеет длину l и выглядит короче. Это явление называется *лоренцевым сокращением.*

Из формул для лоренцевых преобразований (2.196) следует, что поперечные наблюдаемые размеры тел вдоль осей y и z не испытывают никаких сокращений. Отсюда следует, что объем движущегося тела W уменьшается так же, как и его длина:

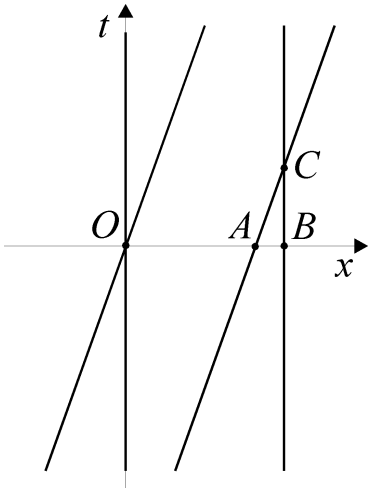
$$W = \triangle x \triangle y \triangle z = W_0\sqrt{1 - V^2/c^2},$$

где W_0 – собственный объем в системе координат, связанной с телом.

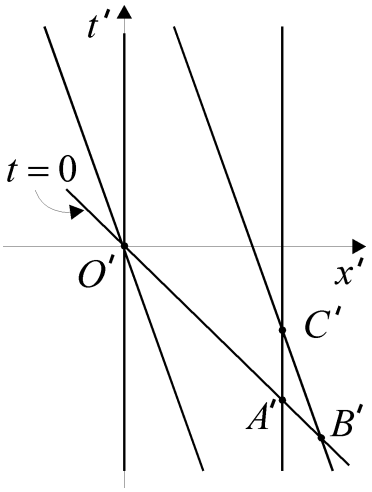
Сокращение длин при движении, так же, как и замедление времени, породило много "парадоксов".

Парадокс машины и гаража. Предположим, что машина и гараж имеют одинаковую собственную длину L. Водитель на большой скорости V въезжает в гараж, а дежурный по гаражу закрывает ворота в тот момент, когда задний бампер машины войдет в гараж. С точки зрения дежурного никаких проблем не возникает: машина испытывает лоренцево сокращение и, следовательно, без труда поместится в гараже. Однако с точки зрения водителя ситуация противоположная: лоренцево сокращение испытывает гараж и поэтому машина никак в гараже поместится не может.

Разрешение "парадокса" заключается в отсутствии понятия одновременности в специальной теории относительности. Для простоты мы предполагаем, что машина свободно пробивает заднюю стену гаража и продолжает движение. На рисунках 2.12*a,b* показаны пространственно-временны́е диаграммы событий с точки зрения дежурного по гаражу и водителя. Для удобства, мы выбрали общее начало координат (точка O) таким образом, что оно соответствует моменту закрытия ворот. С точки зрения дежурного гараж покоится и ему соответствует вертикальная полоса, границы которой пересекают ось x в точках O и B. Машине соответствует затемненная наклонная полоса, пересекающая ось x в точках O и B. В системе координат дежурного машина испытывает лоренцево сокращение и поэтому событие A расположено левее B. Это значит, что в момент закрытия ворот машина будет расположена в гараже. Через некоторое время передним бампер машины достигнет стены гаража и пробьет его, событие C.



(a) Машина и гараж с точки зрения дежурного.



(b) Машина и гараж с точки зрения водителя.

Рис. 2.12: Пространственно-временны́е диаграммы машины и гаража в различных системах координат. Общее начало систем координат, точка O, соответствует событию закрытия ворот. Затемненная полоса соответствует движению машины.

Нетрудно проверить, что события A, B и C в нештрихованной системе координат,

связанной с гаражом, на рис.2.12a имеют следующие координаты:

$$A = \left(0, L\sqrt{1 - V^2/c^2}\right),$$
$$B = (0, L)$$
$$C = \left(\frac{L}{V}\left(1 - \sqrt{1 - V^2/c^2}\right), L\right).$$

С точки зрения водителя гараж равномерно приближается к машине со скоростью V. Используя формулы для лоренцевых преобразований координат (2.196), вычислим координаты событий A, B и C в штрихованной системе координат, привязанной к водителю:

$$A' = \left(-\frac{VL}{c^2}, L\right),$$
$$B' = \left(-\frac{VL}{c^2\sqrt{1 - V^2/c^2}}, \frac{L}{\sqrt{1 - V^2/c^2}}\right)$$
$$C' = \left(-\frac{L}{V}\left(1 - \sqrt{1 - V^2/c^2}\right), L\right).$$

На рис.2.12b показана схема событий с точки зрения водителя. Машине и гаражу соответствуют вертикальная и наклонная полосы. С точки зрения водителя передний бампер машины сначала пробивает стену гаража, событие C', и лишь после этого закрываются ворота, событие O'.

Таким образом, последовательность событий с точки зрения дежурного и водителя прямо противоположная. С точки зрения дежурного сначала закрываются ворота и лишь потом машина пробивает стену. Водитель же видит, что сначала машина пробивает стену, и только потом закрываются ворота.

Заметим, что, если скорость света равна бесконечности, то все три события A, B и C имеют одинаковые координаты $(0, L)$ в обеих системах отсчета. □

Рассмотренный пример наглядно показывает, что понятие одновременности событий в специальной теории относительности отсутствует. К этому трудно привыкнуть, т.к. наша интуиция основана на повседневной жизни, где скорости движения частиц малы по сравнению со скоростью света.

2.10.4 Сложение скоростей и эффект Доплера

Пусть, как и ранее, штрихованная система координат O' движется с постоянной скоростью V вдоль оси x в покоящейся системе координат O. Следовательно, координаты событий в этих системах отсчета связаны преобразованиями Лоренца (2.196). Предположим, что в штрихованной системе координат точечная частица движется равномерно и прямолинейно с постоянной скоростью $v' = (v'_x, v'_y, v'_z)$:

$$x' = v'_x t', \quad y' = v'_y t', \quad z' = v'_z t'.$$

Тогда в покоящейся системе координат мировая линия частицы будет прямой линией, заданной уравнениями:

$$t = \frac{1 + v'_x V/c^2}{\sqrt{1 - V^2/c^2}} t', \qquad y = v'_y t',$$
$$x = \frac{v'_x + V}{\sqrt{1 - V^2/c^2}} t', \qquad z = v'_z t',$$

где время t' рассматривается, как параметр вдоль мировой линии частицы. Отсюда вытекает, что скорость частицы, наблюдаемая в покоящейся системе координат, имеет следующие компоненты:

$$\begin{aligned} v_x &:= \frac{x}{t} = \frac{v'_x + V}{1 + v'_x V/c^2}, \\ v_y &:= \frac{y}{t} = \frac{v'_y}{1 + v'_x V/c^2}\sqrt{1 - V^2/c^2}, \\ v_z &:= \frac{y}{t} = \frac{v'_z}{1 + v'_x V/c^2}\sqrt{1 - V^2/c^2}. \end{aligned} \tag{2.201}$$

Это и есть правило сложения скоростей в специальной теории относительности. Из полученных формул сразу следуют правила сложения скоростей в частных случаях, когда скорости параллельны или перпендикулярны друг другу.

Обозначим углы, которые составляют скорости v' и v с осями x' и x, соответственно через α' и α, т.е.

$$v'_x = v' \cos\alpha', \qquad v_x = v\cos\alpha.$$

Здесь мы предполагаем, что углы меняются в интервале: $0 \le \alpha', \alpha \le \pi$. Для модуля скорости $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$, наблюдаемой в покоящейся системе отсчета, из (2.201) следует равенство

$$v = \frac{\sqrt{v'^2 + V^2 + 2v'V\cos\alpha' - (v'V\sin\alpha'/c)^2}}{1 + v'V\cos\alpha'/c^2}. \tag{2.202}$$

Далее, простые вычисления дают

$$\sqrt{1 - v^2/c^2} = \frac{\sqrt{1 - v'^2/c^2}\sqrt{1 - V^2/c^2}}{1 + v'V\cos\alpha'/c^2}. \tag{2.203}$$

Предложение 2.10.1. *Если скорость частицы в какой либо инерциальной системе отсчета меньше скорости света, $v' < c$, то она будет меньше скорости света и в любой другой инерциальной системе координат, $v < c$. Если в движущейся системе координат распространяется свет, $v' = c$, то в покоящейся системе координат его скорость такая же, $v = c$.*

Доказательство. Если $v' < c$ и $V < c$, то знаменатель в (2.203) положителен, т.к. $\cos\alpha'$ ограничен сверху единицей. Следовательно, правая часть (2.203) положительна. Оценим ее сверху. Максимальное значение принимается, когда $\cos\alpha' = -1$, т.е. частица движется в сторону, противоположную движению штрихованной системы координат. Сверху правая часть (2.203) ограничена единицей, поскольку справедливо неравенство

$$\left(1 - \frac{v'^2}{c^2}\right)\left(1 - \frac{V^2}{c^2}\right) \le \left(1 - \frac{v'V}{c^2}\right),$$

которое просто проверяется. Таким образом, правая часть лежит в пределах от 0 до 1. Поэтому $0 \le v < c$. При этом $v = 0$ тогда и только тогда, когда $v' = V$ и $\cos\alpha' = -1$.

Если $v' = c$, то из уравнения (2.202) вытекает, что $v = c$. □

Из формулы (2.203) следует, что если частица движется в штрихованной системе координат со скоростью V в противоположную сторону, $\cos\alpha' = -1$, то ее скорость в неподвижной системе координат равна нулю. Это означает, что для произвольной частицы, движущейся со скоростью, меньшей скорости света, всегда найдется сопутствующая система координат, в которой она покоится. Ясно также, что для света сопутствующей системы координат не существует.

В разделе 2.7 была введена скорость относительного движения двух инерциальных систем отсчета (2.108). Из этой формулы вытекает, что скорость относительного движения двух инерциальных систем отсчета не превосходит скорости света. Вместе с предложением 2.10.1 это означает, что специальная теория относительности описывает движение частиц со скоростью, меньшей скорости света, самосогласованным образом. Если не постулировать невозможность движения частиц со скоростью, превышающей скорость света, то специальная теория относительности такое движение допускает. Ясно, что для таких частиц не может существовать сопутствующей системы координат. Частицы, движущиеся со скоростью, превышающей скорость света называются *тахионами*. Такие частицы, если они существуют, представляют большие теоретические трудности. Например, возникает проблема с положительной определенностью канонического гамильтониана. Тем не менее они возникают в различных моделях математической физики. В настоящее время тахионы экспериментально не обнаружены, и принято считать, что в природе они отсутствуют.

Из правила сложения скоростей (2.201) следует также соотношение между углами:

$$\operatorname{tg}\alpha = \frac{\sqrt{v_y^2 + v_z^2}}{v_x} = \frac{v'\sin\alpha'}{v'\cos\alpha' + V}\sqrt{1 - V^2/c^2}.$$

Теперь кратко обсудим два эффекта, которые подтверждаются экспериментально.

Если в системе отсчета O' распространяется луч света, т.е. $v' = c$, то для угла, наблюдаемого в неподвижной системе координат, получаем выражение

$$\operatorname{tg}\alpha = \frac{\sin\alpha'}{\cos\alpha' + V/c}\sqrt{1 - V^2/c^2}. \tag{2.204}$$

Эту формулу можно переписать в более симметричном виде:

$$\sin\alpha = \sin\alpha'\frac{\sqrt{1 - V^2/c^2}}{1 + \frac{V}{c}\cos\alpha'}$$

или

$$\operatorname{tg}\frac{\alpha}{2} = \operatorname{tg}\frac{\alpha'}{2}\sqrt{\frac{1 - V/c}{1 + V/c}}.$$

Из последней формулы, в частности, следует, что $\alpha' > \alpha$ при $V > 0$. Изменение угла, под которым наблюдается луч света, в движущейся и покоящейся системах координат, называется *аберрацией света*.

Замечание. Если источник света покоится в системе координат O, а наблюдатель движется вместе с системой отсчета O', то бесконечно удаленный источник света будет казаться ему смещенным на угол $\alpha' - \alpha$. Для наблюдательной астрономии это означает, что звезды на небе будут описывать эллипсы в соответствии с годовым вращением Земли вокруг Солнца. Этот эффект действительно наблюдается. □

В качестве еще одного приложения преобразований Лоренца рассмотрим *эффект Доплера*. Допустим, что луч света (плоская электромагнитная волна) распространяется в покоящейся системе координат O в плоскости x, y под углом α к оси x. В такой волне напряженность электрического поля $E(t, x)$ меняется по правилу

$$E(t, x) = E_0\, \mathsf{e}^{i\varphi},$$

где $E_0 = \mathsf{const}$ – амплитуда электрического поля, и

$$\varphi(t, x, y) = \omega \left(t - \frac{x \cos\alpha + y \sin\alpha}{c} \right)$$

– фаза электромагнитной волны. В приведенной формуле ω – это частота волны, которая наблюдается в покоящейся системе координат. Электромагнитное поле, которое описывает плоскую волну удовлетворяет уравнениям Максвелла. Поскольку уравнения Максвелла ковариантны относительно преобразований Лоренца и напряженность E подчиняется тензорному закону преобразования, то фаза волны φ является скалярным полем на пространстве Минковского. Следовательно, фазы волны в покоящейся и движущейся системах координат совпадают. Для штрихованной системы координат, которая равномерно движется вдоль оси x и была описана ранее, равенство фаз имеет вид

$$\omega \left(t - \frac{x \cos\alpha + y \sin\alpha}{c} \right) = \omega' \left(t' - \frac{x' \cos\alpha' + y' \sin\alpha'}{c} \right),$$

где ω' – частота электромагнитной волны в движущейся системе координат. Полученное уравнение должно выполняться при всех значениях t, x, y и $t', x'y'$. Подставляя выражения для t, x, y из преобразований Лоренца (2.197) и приравнивая коэффициенты при t', x' и y', получаем формулы преобразования для частоты

$$\omega' = \omega \frac{1 - \frac{V}{c} \cos\alpha}{\sqrt{1 - V^2/c^2}} \tag{2.205}$$

и углов

$$\cos\alpha' = \frac{\cos\alpha - V/c}{1 - \frac{V}{c}\cos\alpha}, \qquad \sin\alpha' = \sin\alpha \frac{\sqrt{1 - V^2/c^2}}{1 - \frac{V}{c}\cos\alpha}.$$

Из последних двух равенств следует соотношение между углами:

$$\operatorname{tg}\alpha' = \frac{\sin\alpha}{\cos\alpha - V/c} \sqrt{1 - V^2/c^2}.$$

Отличие в знаках в знаменателе полученной формулы и равенства (2.204) обусловлено тем, что углы α и α' поменялись местами, что соответствует изменению знака скорости, $V \to -V$.

Изменение частоты света (2.205), связанное с движением системы отсчета, называется эффектом Доплера. Пусть источник света находится в покоящейся системе отсчета O, и свет распространяется вдоль оси x. Допустим, что наблюдатель находящийся в движущейся системе отсчета O' и удаляется от источника света, т.е. $V > 0$ и $\cos\alpha = 1$. Тогда наблюдаемая им частота света уменьшится. Этот эффект называется *красным смещением*. Если же, наоборот, наблюдатель приближается к источнику света, $V > 0$ и $\cos\alpha = -1$, то частота света увеличится. Такое поведение частоты называется *голубым смещением*.

Замечание. Изучение спектра далеких галактик показывает, что он смещен в красную сторону. Это приводит к выводу о том, что галактики удаляются от Земли и вселенная расширяется. □

2.10.5 Равноускоренное движение

В настоящем разделе мы определим 4-скорость и наблюдаемую скорость точечной частицы, а также ее ускорение. Затем рассмотрим равноускоренное движение в специальной теории относительности и сравним его с равноускоренным движением в механике Ньютона.

Обозначим инерциальные координаты в пространстве Минковского $\mathbb{R}^{1,3}$ через x^α, $\alpha = 0, 1, 2, 3$. Пусть задана мировая линия частицы

$$\mathbb{R} \ni \sigma \mapsto \{q^\alpha(\sigma)\} \in \mathbb{R}^{1,3},$$

где σ – некоторый параметр. По-определению, мировая линия частицы времениподобна, а ее длина s при $\sigma \in [0, \tau]$ равна интегралу

$$s = \int_0^\tau d\sigma \sqrt{\eta_{\alpha\beta} \frac{dq^\alpha}{d\sigma} \frac{dq^\beta}{d\sigma}}.$$

Выберем длину траектории s, которая пропорциональна собственному времени частицы, в качестве параметра вдоль мировой линии. Тогда мировая линия задается четырьмя функциями $q^\alpha(s)$.

Для времениподобных кривых в качестве параметра вдоль кривой можно выбрать также время x^0. Действительно, поскольку для произвольной траектории частицы $dq^0/ds > 0$, то параметр s можно рассматривать как некоторую функцию координаты $s = s(x^0)$. Тогда мировая линии частицы задается функциями

$$q^0 = x^0, \qquad q^\mu = q^\mu(q^0), \qquad \mu = 1, 2, 3.$$

Замечание. Физическая интерпретация приведенных параметризаций мировых линий следующая. Время для наблюдателя, который движется вместе с частицей, совпадает с собственным временем s/c. Поэтому проведенные им измерения соответствуют сопутствующей системе координат. Во втором случае x^0 – это время наблюдателя, покоящегося в выбранной исходной инерциальной системе отсчета. Поэтому его измерения соответствуют внешнему наблюдателю. □

Определение. 4-вектор u с компонентами

$$u^\alpha := c \frac{dq^\alpha}{ds} \tag{2.206}$$

называется *4-скоростью* или просто *скоростью* частицы. 4-вектор w с компонентами

$$w^\alpha := c \frac{du^\alpha}{ds} = c^2 \frac{d^2 q^\alpha}{ds^2} \tag{2.207}$$

называется *4-ускорением* или просто *ускорением* частицы. *Наблюдаемой скоростью* частицы $\boldsymbol{v}$ в системе координат x^α называется 3-вектор с компонентами

$$v^\mu := \frac{dq^\mu}{dq^0} = \frac{u^\mu}{u^0}, \qquad \mu = 1, 2, 3. \tag{2.208}$$

3-вектор $\boldsymbol{a}$ с компонентами

$$a^\mu := \frac{dv^\mu}{dq^0} = \frac{d^2 q^\mu}{(dq^0)^2}. \tag{2.209}$$

называется *наблюдаемым ускорением* частицы. □

Скорость (2.206) и ускорение (2.207) точечной частицы являются 4-векторами относительно преобразований из группы Лоренца $\mathbb{O}(1,3)$, действующей в пространстве Минковского $\mathbb{R}^{1,3}$. Наблюдаемая скорость (2.208) и наблюдаемое ускорение (2.209) преобразуются как векторы только относительно подгруппы вращений $\mathbb{O}(3)$, действующей на пространственных сечениях $x^0 = \mathsf{const}$. Они нековариантны относительно лоренцевых преобразований и не являются компонентами каких либо 4-векторов.

Мы предполагаем, что траектория частицы времениподобна и направлена в будущее. Поэтому временна́я компонента скорости положительна, $u^0 > 0$.

В настоящем разделе полезно следить за размерностями различных величин:

$$[x^0] = \text{сек}, \qquad [x^\mu] = [s] = \text{см},$$
$$[u^0] = 1, \qquad [u^\mu] = [v^\mu] = \frac{\text{см}}{\text{сек}},$$
$$[w^0] = \frac{1}{\text{сек}}, \qquad [w^\mu] = [a^\mu] = \frac{\text{см}}{\text{сек}^2}.$$

Из определения скорости (2.206) сразу следует равенство

$$u^\alpha u_\alpha = c^2 (u^0)^2 - \boldsymbol{u}^2 = c^2, \tag{2.210}$$

где $\boldsymbol{u}^2 = -u^\mu u_\mu \geq 0$ – обычный квадрат трехмерного вектора в евклидовом пространстве. Дифференцирование этого равенства по s приводит к соотношению

$$u^\alpha w_\alpha = 0,$$

т.е. ускорение частицы всегда ортогонально ее скорости.

Умножим равенство (2.210) на $m^2 c^2$, где $m > 0$ – масса частицы, $[m] =$г, и перепишем его в виде

$$E = mc^2 \sqrt{1 + \frac{\boldsymbol{u}^2}{c^2}},$$

где введено обозначение $E := mc^2 u^0 > 0$. В нерелятивистском пределе, когда пространственные компоненты скорости малы по сравнению со скоростью света, $\boldsymbol{u}^2/c^2 \ll 1$, квадратный корень можно разложить в ряд Тейлора. В первом приближении получим равенство

$$E = mc^2 + \frac{m\boldsymbol{u}^2}{2}.$$

Поскольку второе слагаемое совпадает с выражением для кинетической энергии точечной частицы в механике Ньютона, то отсюда следует, что величина E представляет собой энергию частицы в специальной теории относительности. В сопутствующей системе координат, где $\boldsymbol{u} = 0$, получаем знаменитую формулу Эйнштейна

$$E = mc^2,$$

которая дает выражение для энергии покоя точечной частицы. Формулу (2.210) можно переписать также в виде

$$\frac{E^2}{c^2} - \boldsymbol{p}^2 = m^2 c^2,$$

где введен пространственный импульс частицы $\boldsymbol{p} := m\boldsymbol{u}$. Если ввести *4-импульс* частицы

$$p^\alpha := mu^\alpha = (mu^0, mu^\mu) = \left(\frac{E}{c^2}, p^\mu\right),$$

то равенство (2.210) примет вид

$$p^2 = p^\alpha p_\alpha = m^2c^2.$$

Согласно определению (2.208) выполнено равенство $u^\mu = u^0 v^\mu$. Поэтому из уравнения (2.210) вытекает соотношение

$$u^0 = \frac{1}{\sqrt{1 - \boldsymbol{v}^2/c^2}}. \tag{2.211}$$

Заметим, что в нерелятивистском пределе, который мы определили как $\boldsymbol{u}^2/c^2 \to 0$, из равенства (2.210) вытекает, что при этом $u^0 \to 1$. Поэтому из соотношения $u^\mu = u^0 v^\mu$ следует, что пределы $\boldsymbol{u}^2 \to 0$ и $\boldsymbol{v}^2 \to 0$ эквивалентны.

Теперь найдем формулы, связывающие 4-ускорение с наблюдаемым ускорением частицы. Для этого в определении 4-ускорения надо заменить производную

$$c\frac{d}{ds} = u^0 \frac{d}{dq^0}$$

и воспользоваться формулой (2.211). В результате получим равенства

$$\begin{aligned} w^0 &= \frac{(\boldsymbol{v}, \boldsymbol{a})}{c^2(1 - \boldsymbol{v}^2/c^2)^2}. \\ \boldsymbol{w} &= \boldsymbol{a}\frac{1}{1 - \boldsymbol{v}^2/c^2} + \boldsymbol{v}\frac{(\boldsymbol{v}, \boldsymbol{a})}{c^2}\frac{1}{(1 - \boldsymbol{v}^2/c^2)^2} \end{aligned} \tag{2.212}$$

В нерелятивистском пределе, очевидно, $\boldsymbol{w} = \boldsymbol{a}$ и $w^0 = 0$.

Для определения равноускоренного движения нам понадобятся формулы преобразования компонент наблюдаемого ускорения при переходе от покоящейся системы координат к движущейся. Как было отмечено, компоненты наблюдаемого ускорения не являются компонентами какого либо 4-вектора, и поэтому при лоренцевых бустах преобразуются не по тензорным правилам. Пусть система координат O' связана с частицей, т.е. в момент времени $t = x^0$, ее скорость совпадает с мгновенной скоростью частицы, $\boldsymbol{V} = \boldsymbol{v}$. Тогда в штрихованной системе координат скорость $\boldsymbol{v}' = 0$, и формулы (2.212) принимают вид

$$\boldsymbol{w}' = \boldsymbol{a}', \qquad w'^0 = 0. \tag{2.213}$$

Для определенности, предположим, что скорость $\boldsymbol{v}$ направлена вдоль оси x, т.е. $\boldsymbol{v} = (v^1, 0, 0)$ и $\boldsymbol{v}^2 = (v^1)^2$. Тогда из формул (2.212) следуют соотношения:

$$w^1 = \frac{a^1}{1 - \boldsymbol{v}^2/c^2} + v^1 w^0, \qquad w^2 = \frac{a^2}{1 - \boldsymbol{v}^2/c^2}, \qquad w^3 = \frac{a^3}{1 - \boldsymbol{v}^2/c^2}. \tag{2.214}$$

Теперь вспомним, что ускорение w – лоренцев вектор. Его компоненты преобразуются, как компоненты вектора (x^0, x^1, x^2, x^3). Поэтому компоненты ускорения в штрихованной и нештрихованной системах координат связаны преобразованием Лоренца (2.197):

$$w^0 = \frac{v^1 w'^1}{c\sqrt{1 - \boldsymbol{v}^2/c^2}}, \qquad w^1 = \frac{w'^1}{\sqrt{1 - \boldsymbol{v}^2/c^2}}, \qquad w^2 = w'^2, \qquad w^3 = w'^3,$$

где мы учли, что $w'^0 = 0$. Подставляя в эти формулы соотношения между 4-ускорениями и наблюдаемыми ускорениями (2.213) и (2.214) окончательно получаем формулы преобразования наблюдаемых ускорений:

$$a^1 = a'^1(1 - \boldsymbol{v}^2/c^2)^{3/2}, \quad a^2 = a'^2(1 - \boldsymbol{v}^2/c^2), \quad a^3 = a'^3(1 - \boldsymbol{v}^2/c^2). \tag{2.215}$$

Эти формулы преобразования, как и следовало ожидать, отличаются от преобразований Лоренца.

Рассмотрим равноускоренное движение.

Определение. Движение точечной частицы называется *равноускоренным*, если наблюдаемое ускорение $\boldsymbol{a}'$ в системе координат O', движущейся вместе с данной частицей, постоянно, $\boldsymbol{a}' = \mathsf{const}$. □

Теперь найдем мировую линию частицы, которая движется равноускоренно в покоящейся системе координат. Пусть система отсчета O фиксирована. Поскольку частица движется равноускоренно, то штрихованная система координат O', связанная с частицей, меняется с течением времени.

Рассмотрим случай прямолинейного движения, когда ускорение $\boldsymbol{a}'$ постоянно и параллельно скорости $\boldsymbol{v}$. Для этого достаточно, чтобы ускорение было постоянным и частица в начальный момент времени покоилась. Для определенности, предположим, что ускорение и скорость параллельны оси $x = x^1$, т.е. мировая линия частицы лежит в плоскости t, x. Положим $\boldsymbol{a}' = (b, 0, 0)$, где $b = \mathsf{const}$, и $\boldsymbol{v} = \Big(v(t), 0, 0\Big)$. Тогда из (2.215) следует одно дифференциальное уравнение

$$\frac{dv}{dt} = b\left(1 - v^2/c^2\right)^{3/2}.$$

Оно легко интегрируется:

$$b(t - t_0) = \frac{v}{\sqrt{1 - v^2/c^2}}.$$

где t_0 – постоянная интегрирования. Подставляя равенство $v = dx/dt$, получаем дифференциальное уравнение на функцию $x(t)$. Интегрирование полученного уравнения дает уравнение на мировую линию частицы:

$$(x - x_0)^2 - c^2(t - t_0)^2 = \frac{c^4}{b^2},$$

где x_0 – вторая постоянная интегрирования. Таким образом, мировая линия частицы при прямолинейном равноускоренном движении представляет собой гиперболу в плоскости движения t, x. При $t_0 = 0$ и $x_0 = 0$ мировая линия показана на рис.2.13. Для сравнения показана также мировая линия частицы в механике Ньютона, которая задана хорошо известным уравнением

$$x = \frac{c^2}{b} + \frac{1}{2}bt^2$$

и представляет собой параболу.

При равноускоренном движении наблюдаемая скорость частицы при $t \to \infty$ в специальной теории относительности стремится к скорости света, а в механике Ньютона – к бесконечности.

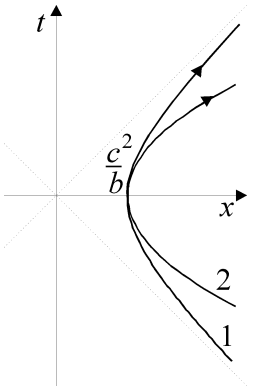


Рис. 2.13: Мировые линии частиц при прямолинейном равноускоренном движении в специальной теории относительности (1) и в механике Ньютона (2).

Глава 3

Многообразия и тензорные поля

В настоящей главе начинается изложение основных понятий дифференциальной геометрии. Изложение часто носит координатный характер, т.е. все соотношения записываются в локальной системе координат и приводятся правила преобразований геометрических объектов при переходе от одной системы координат к другой. Параллельно даются также глобальные определения, не зависящие от выбора системы координат, и показывается связь координатного и инвариантного описания. Глобальный бескоординатный подход компактен в обозначениях, прозрачен и удобен для определения геометрических объектов. Координатный подход является более громоздким, однако это не является признаком меньшей строгости. Он незаменим в моделях математической физики, основанных на дифференциальной геометрии, где требуется проведение расчетов.

3.1 Многообразие

Основным понятием дифференциальной геометрии является дифференцируемое многообразие $\mathbb{M}$, которое обобщает понятие евклидова пространства $\mathbb{R}^n$. Топологически нетривиальные многообразия не покрываются одной системой координат, однако локально устроены так же, как и евклидовы пространства. Это позволяет использовать математический анализ для построения и анализа многих важных моделей современной математической физики.

Определение. Топологическое хаусдорфово пространство $\mathbb{M}$ со счетной базой, каждая точка которого имеет окрестность, гомеоморфную открытому n-мерному шару единичного радиуса $\mathbb{B}_1^n$ в евклидовом пространстве $\mathbb{R}^n$, называется n-мерным *многообразием*. Число n называется *размерностью* многообразия. Мы пишем $\dim \mathbb{M} = n$. □

Замечание. В определении многообразия требование хаусдорфовости существенно и исключает случаи подобные примерам на с. 25. Согласно теореме 2.3.12, любая сходящаяся последовательность в хаусдорфовом пространстве имеет не более одного предела.

Предположение о счетности базы также существенно, потому что это требование обеспечивает паракомпактность многообразия, которая в рассматриваемом случае достаточна для существования разбиения единицы. В свою очередь, существование разбиения единицы важно, т.к. позволяет определять геометрические объекты на всем многообразии, исходя из их задания в локальных координатах. Оно часто используется для доказательства теорем существования.

Выбор шаров единичного радиуса $\mathbb{B}_1^n$ сделан для определенности и не является существенным, т.к. шар единичного радиуса гомеоморфен всему евклидову пространству $\mathbb{R}^n$, а также любой (связной и односвязной) области в $\mathbb{R}^n$. □

Замечание. Поскольку многообразие, по-определению, является хаусдорфовым топологическим пространством, то для многообразий понятие компактного пространства и компакта совпадают. □

Окрестность произвольной точки многообразия $\mathbb{M}$ устроена так же, как и окрестность точки в евклидовом пространстве $\mathbb{R}^n$. Однако в отличие от евклидова пространства, которое, по-определению, может быть покрыто одной картой, многообразие в общем случае одной картой не покрывается. Поэтому на многообразии общего вида нельзя ввести структуру векторного или аффинного пространства.

Предложение 3.1.1. *Многообразие представляет собой объединение конечного или счетного числа областей, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, каждая из которых гомеоморфна n-мерному шару, и, следовательно, любой области в $\mathbb{R}^n$.*

Доказательство. Поскольку каждая точка имеет окрестность, гомеоморфную $\mathbb{R}^n$, то все многообразие можно покрыть, возможно, несчетным числом карт. Выберем счетную базу на $\mathbb{M}$, которая также является покрытием. Каждая координатная окрестность $\mathbb{U}_i$ является объединением счетного числа элементов базы и, поэтому, на каждом элементе базы задан гомеоморфизм в $\mathbb{R}^n$. Теперь можно выбрать базу топологии в качестве координатного покрытия, а она – счетна. □

В приложениях рассматриваются, как правило, многообразия, которые покрываются конечным числом карт.

Замечание. Поскольку многообразие представляет собой объединение не более, чем счетного числа областей, гомеоморфных n-мерному шару, то любое многообразие является сепарабельным топологическим пространством. □

Согласно данному определению, многообразие, как объединение открытых множеств, не имеет границы, которую в случае многообразий принято называть краем (см. следующий раздел).

Поскольку каждая точка многообразия гомеоморфна шару в евклидовом пространстве, то всякое многообразие является локально связным. Согласно теореме 2.3.17 всякое многообразие представляет собой объединение связных компонент, каждая из которых является одновременно открытой и замкнутой. Число этих компонент может быть не более, чем счетным, т.к. мы предполагаем счетность базы многообразия.

Определение. Под 0-мерным многообразием мы будем понимать счетное множество точек с дискретной топологией. Одномерное и двумерное многообразия называются соответственно *линией* и *поверхностью*. □

Размерность многообразия является топологическим инвариантом: два гомеоморфных многообразия имеют одинаковую размерность. Это утверждение является прямым следствием теоремы 2.4.1 об инвариантности размерности.

Данное выше определение задает *топологическое многообразие*, т.к. в нем говорится только о непрерывности. Теперь перейдем к описанию дифференцируемых многообразий.

Определение. Из определения многообразия следует, что существует гомеоморфизм (непрерывная биекция)

$$\varphi_i : \quad \mathbb{M} \supset \mathbb{U}_i \ \rightarrow \ \varphi_i(\mathbb{U}_i) \subset \mathbb{R}^n,$$

области $\mathbb{U}_i$ на ее образ $\varphi_i(\mathbb{U}_i)$ (суммирования нет) в евклидовом пространстве $\mathbb{R}^n$. Поскольку в евклидовом пространстве есть система координат, например, декартова, то данный гомеоморфизм можно записать в виде

$$\varphi_i : \quad \mathbb{M} \supset \mathbb{U}_i \ni x \mapsto \varphi_i(x) = (x^1, \ldots, x^n) \in \mathbb{U}_i \subset \mathbb{R}^n,$$

где точку многообразия x и ее координаты $(x^1, \ldots, x^n)$ мы обозначили одной и той же буквой. Области $\mathbb{U}_i$, покрывающие многообразие, называются *координатными окрестностями*, а набор чисел $(x^1, \ldots, x^n)$ – *локальными координатами*. Пара $(\mathbb{U}_i, \varphi_i)$ называется *картой*.

Если две карты пересекаются, $\mathbb{U}_i \bigcap \mathbb{U}_j \neq 0$, то произвольная точка из пересечения $x \in \mathbb{U}_i \bigcap \mathbb{U}_j$ имеет свой набор координат в каждой карте. Отображение областей евклидова пространства,

$$\varphi_j \circ \varphi_i^{-1} : \quad \mathbb{R}^n \supset \varphi_i(\mathbb{U}_i \cap \mathbb{U}_j) \rightarrow \varphi_j(\mathbb{U}_i \cap \mathbb{U}_j) \subset \mathbb{R}^n,$$

задается набором n функций $x^{\alpha'}(x)$ от n переменных (2.55), где x^α и $x^{\alpha'}$ – координаты соответственно на $\mathbb{U}_i$ и $\mathbb{U}_j$. Они называются *функциями склейки*, поскольку склеивают между собой различные карты. Совокупность всех карт, покрывающих многообразие, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, называется *координатным покрытием* или *атласом* $\{\mathbb{U}_i, \varphi_i\}$, $i \in I$ многообразия $\mathbb{M}$. Атлас, который не содержится ни в каком другом атласе, называется *полным*. □

При проведении вычислений в одной карте точку многообразия $x \in \mathbb{M}$ и ее координаты $\varphi(x) \in \mathbb{R}^n$ можно отождествлять.

Замечание. В силу счетности топологической базы многообразия, координатное покрытие всегда можно выбрать счетным. Однако, даже евклидово пространство можно при желании покрыть несчетным числом карт. Это значит, что полный атлас любого многообразия всегда несчетен. □

Мы требуем, чтобы каждая функция склейки $\varphi_j \circ \varphi_i^{-1}$ заданного атласа определяла некоторый диффеоморфизмом, т.е. являлась элементом псевдогруппы преобразований координат $\mathsf{diff}^{\,k}(\mathbb{R}^n)$ для некоторого k, определенной в разделе 2.5. Поэтому мы говорим, что атлас $\{\mathbb{U}_i, \varphi_i\}$ совместим с псевдогруппой преобразований координат $\mathsf{diff}^{\,k}(\mathbb{R}^n)$.

Определение. Многообразие $\mathbb{M}$ вместе с полным атласом $\{\mathbb{U}_i, \varphi_i\}$ называется *дифференцируемым многообразием* класса $\mathcal{C}^k$, $k \in \mathbb{N}$, если функции склейки (2.55) для всех пересекающихся карт непрерывны вместе со своими частными производными, включая производные k-того порядка $\varphi_j \circ \varphi_i^{-1} \in \mathsf{diff}^{\,k}(\mathbb{R}^n)$. Полный атлас $\{\mathbb{U}_i, \varphi_i\}$ называется *дифференцируемой структурой* многообразия $\mathbb{M}$. □

Аналогично определяются гладкие $\mathcal{C}^\infty$ и вещественно аналитические многообразия $\mathcal{C}^\omega$. На $\mathcal{C}^\omega$-многообразиях функции склейки задаются сходящимися степенными рядами. Напомним, что функции склейки имеют ненулевой якобиан и поэтому осуществляют взаимно однозначное отображение областей в $\mathbb{R}^n$ (теорема 2.5.1). Это

означает, что обратное преобразование существует, и дифференцируемость обратных функций такая же, как и самих функций склейки. В дальнейшем, если не оговорено противное, под многообразием мы будем понимать $\mathcal{C}^\infty$ дифференцируемые многообразия. Многие из рассмотренных ниже утверждений справедливы и при более слабых ограничениях на дифференцируемость функций склейки. Для проверки достаточно следить лишь за числом производных при вычислениях.

Замечание. Отображения φ_i являются гомеоморфизмами. Это означает, что дифференцируемые структуры, согласованы с естественной топологией евклидова пространства $\mathbb{R}^n$, т.е. можно считать, что топология многообразия индуцирована отображениями областей $\mathbb{U}_i$ в евклидово пространство $\mathbb{R}^n$ с естественной топологией. □

Замечание. В определение дифференцируемого многообразия входит понятие евклидова пространства $\mathbb{R}^n$. В евклидовом пространстве можно определить много различных структур. Для определения топологического многообразия нам достаточно наличия топологии в $\mathbb{R}^n$. Для определения дифференцируемого многообразия нам достаточно дополнительно предположить наличие структуры поля на вещественной прямой $\mathbb{R}$ (для определения производной необходимо вычитать и делить числа). Конечно, в евклидовом пространстве всегда можно ввести евклидову метрику и структуру векторного пространства. Однако для определения дифференцируемого многообразия это не является необходимым условием. □

Отметим некоторые свойства функций склейки.

Предложение 3.1.2. *Функции склейки удовлетворяют тождествам:*

$$f_{ij} = f_{ji}^{-1}, \qquad \forall\ x \in \mathbb{U}_i \cap \mathbb{U}_j, \tag{3.1}$$

$$f_{ij} f_{jk} f_{ki} = \mathsf{id}\,, \qquad \forall\ x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k, \tag{3.2}$$

где id *– тождественное отображение.*

Доказательство. Прямая проверка. □

Следствие. Справедливо тождество

$$f_{ii} = \mathsf{id}\,.$$

Доказательство. Из равенства (3.2) с учетом (3.1) получаем тождество

$$f_{ii} = f_{ik} f_{ki} = \mathsf{id}\,.$$ □

Предложение 3.1.3. *Любой атлас класса* $\mathcal{C}^k$ *можно дополнить до полного атласа того же класса гладкости, и соответствующий полный атлас единственный. Если на многообразии задано два диффеоморфных атласа* $\{\mathbb{U}_i, \varphi_i\}$ *и* $\{\mathbb{V}_\text{A}, \psi_\text{A}\}$*, т.е. все отображения*

$$\psi_\text{A} \circ \varphi_i^{-1}: \qquad \varphi_i(\mathbb{U}_i \cap \mathbb{V}_\text{A}) \to \psi_j(\mathbb{U}_i \cap \mathbb{V}_\text{A})$$

являются диффеоморфизмами областей евклидова пространства $\mathbb{R}^n$*, как только пересечения* $\mathbb{U}_i \cap \mathbb{V}_\text{A}$ *не пусты, то они содержатся в одном полном атласе. Любые два атласа, содержащиеся в одном полном атласе связаны между собой диффеоморфизмом.*

Доказательство. Пусть $\{\mathbb{W}, \chi\}$ – семейство всех пар таких, что χ есть гомеоморфизм открытого подмножества $\mathbb{W} \subset \mathbb{M}$ на открытое подмножество в евклидовом пространстве $\mathbb{R}^n$ и что отображение

$$\varphi_i \circ \chi^{-1} : \quad \chi(\mathbb{W} \cap \mathbb{U}_i) \to \varphi_i(\mathbb{W} \cap \mathbb{U}_i)$$

есть элемент $\mathsf{diff}^{\,k}(\mathbb{R}^n)$, как только $\mathbb{W} \cap \mathbb{U}_i$ не пусто. Тогда $\{\mathbb{W}, \chi\}$ есть полный атлас, содержащий $\{\mathbb{U}_i, \varphi_i\}$. Этот атлас единственный по-построению.

Если заданы два диффеоморфных атласа $\{\mathbb{U}_i, \varphi_i\}$ и $\{\mathbb{V}_\alpha, \psi_\alpha\}$, то их объединение также является атласом и того же класса гладкости. Поэтому они содержатся в одном полном атласе.

Все атласы, содержащиеся в одном полном атласе, получаются путем отбрасывания некоторого числа карт. Они диффеоморфны между собой, поскольку все функции склейки полного атласа принадлежат $\mathsf{diff}^{\,k}(\mathbb{R}^n)$. □

Данное предложение означает, что для задания дифференцируемой структуры на многообразии достаточно задать один атлас, не обязательно полный. Если на многообразии $\mathbb{M}$ можно задать два атласа одного класса гладкости, которые не диффеоморфны между собой, то это означает, что на $\mathbb{M}$ существуют различные дифференцируемые структуры.

Определение. Пусть на многообразии $\mathbb{M}$ введено две дифференцируемые структуры разных классов гладкости $\mathcal{C}^k$ и $\mathcal{C}^l$, $l > k$, причем полный атлас класса $\mathcal{C}^l$ является одновременно атласом класса $\mathcal{C}^k$. Если $l = \infty$, то ограничение класса гладкости $\mathcal{C}^k$-многообразия до $\mathcal{C}^\infty$-многообразия называется *сглаживанием*. При этом сглаживания, которые связаны $\mathcal{C}^k$ автоморфизмами, рассматриваются, как эквивалентные. □

Теорема 3.1.1 (Уитни). *Каждое $\mathcal{C}^k$-многообразие имеет ровно одно (с точностью до $\mathcal{C}^k$-диффеоморфизмов) сглаживание.*

Доказательство. См. [?]. □

В силу этой теоремы во многих случаях можно игнорировать различие между $\mathcal{C}^k$ дифференцируемыми и $\mathcal{C}^\infty$ гладкими структурами.

В низших размерностях дифференцируемая структура на многообразиях единственна. Доказано, что двух- и трехмерные многообразия допускают ровно одну дифференцируемую структуру [?]. Для многообразий более высоких размерностей это не так. Это утверждение иллюстрирует пример Милнора [?], который показал, что на семимерной сфере $\mathbb{S}^7$, рассматриваемой, как топологическое многообразие, существует 28 недиффеоморфных дифференцируемых структур. Построение этих структур является сложным, и мы его не приводим. Было также доказано, что четырехмерное евклидово пространство $\mathbb{R}^4$ допускает бесконечное множество недиффеоморфных дифференцируемых структур [?]. Существует также теорема о том, что каждое гладкое многообразие имеет единственную вещественно аналитическую структуру [?].

Если в определении многообразия отбросить требование счетности базы, то получим *локально евклидово топологическое пространство*. Существуют такие топологические локально евклидовы пространства, которые вообще не допускают дифференцируемой структуры [?].

Любое открытое подмножество $\mathbb{D}$ дифференцируемого многообразия $\mathbb{M}$ само является многообразием. Дифференцируемая структура на $\mathbb{D}$ состоит из карт $(\mathbb{U}_i \cap \mathbb{D}, \varphi_i|_{\mathbb{D}})$, возникающих после сужения гомеоморфизмов φ_i на $\mathbb{D}$.

Поскольку многообразия являются топологическими пространствами, то их можно умножать как топологические пространства. При этом произведение двух многообразий $\mathbb{M} \times \mathbb{N}$ также является многообразием размерности $m+n$, где $\mathsf{dim}\,\mathbb{M} = m$ и $\mathsf{dim}\,\mathbb{N} = n$. Дифференцируемая структура на прямом произведении многообразий $\mathbb{M} \times \mathbb{N}$ строится, как прямое произведение дифференцируемых структур на $\mathbb{M}$ и $\mathbb{N}$. Пусть дифференцируемая структура на многообразии $\mathbb{M}$ определяется атласом $\{\mathbb{U}_i, \varphi_i\}$, а на $\mathbb{N}$ – атласом $\{\mathbb{V}_{\text{A}}, \phi_{\text{A}}\}$. Тогда естественная дифференцируемая структура на топологическом произведении $\mathbb{M} \times \mathbb{N}$ определяется атласом $\{\mathbb{U}_i \times \mathbb{V}_{\text{A}}, \varphi_i \times \phi_{\text{A}}\}$, где отображение

$$\varphi_i \times \phi_{\text{A}} : \quad \mathbb{U}_i \times \mathbb{V}_{\text{A}} \;\to\; \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$$

определяется естественным образом. Отметим, что этот атлас не будет полным даже если атласы $\{\mathbb{U}_i, \varphi_i\}$ и $\{\mathbb{V}_{\text{A}}, \phi_{\text{A}}\}$ полны.

Определение. Если связное многообразие покрыто совокупностью карт $\{\mathbb{U}_i, \varphi_i\}$ с координатами $\varphi_i(x) = \{x_i^\alpha\}$, причем якобианы функций перехода (2.58) для всех пересекающихся карт $\mathbb{U}_i$ и $\mathbb{U}_j$ положительны,

$$\mathsf{det}\left(\frac{\partial x_i^\alpha}{\partial x_j^\beta}\right) > 0, \qquad \forall i,j$$

то многообразие называется *ориентированным*. Многообразие называется *неориентируемым*, если атласа со всеми положительными якобианами функций склейки не существует. При неудачно выбранном атласе на ориентируемом многообразии якобианы могут быть разных знаков, однако атлас с положительными якобианами существует. Такие многообразия называются *ориентируемыми*. Несвязное многообразие $\mathbb{M}$ называется ориентируемым, если ориентируема каждая его компонента. □

Связное ориентируемое многообразие допускает в точности две ориентации. Чтобы поменять ориентацию ориентированного многообразия достаточно заменить каждую карту $(\mathbb{U}_i, \varphi_i)$ ориентированного атласа на карту $(\mathbb{U}_i, \psi_i)$, где гомеоморфизм ψ_i является композицией φ_i и отражения первой (или любой другой) координаты: $(x^1, x^2, \ldots, x^n) \mapsto (-x^1, x^2, \ldots, x^n)$. (Более подробно ориентация многообразий будет обсуждаться в разделе 11.4.) Ориентация несвязного многообразия – это выбор ориентации на каждой компоненте.

Рассмотрим простейшие примеры многообразий.

Пример 3.1.1. Все евклидово пространство $\mathbb{R}^n$ является простейшим n-мерным многообразием, которое можно покрыть одной картой (а можно и несколькими). Дифференцируемая структура – это полный атлас, содержащий естественную карту $\Big(\mathbb{U} = \mathbb{R}^n, \varphi = \mathsf{id}\,(\mathbb{R}^n)\Big)$. При этом класс гладкости многообразия определяется классом гладкости допустимых преобразований координат. Любое многообразие $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, которое покрывается одной картой, диффеоморфно (см. раздел 3.9) $\mathbb{R}^n$ и называется *тривиальным*. □

Пример 3.1.2. Рассмотрим произвольное вещественное векторное пространство $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$. Если в векторном пространстве выбран базис $\{e_a\}$, $a = 1, \ldots, n$, то каждая точка векторного пространства задается упорядоченным набором вещественных

чисел $(x^1, \ldots, x^n)$. Тогда его можно отождествить с евклидовым пространством $\mathbb{R}^n$ и рассматривать, как гладкое многообразие. Дифференцируемая структура в $\mathbb{V}$ не зависит от выбора базиса, т.к. замена базиса задается невырожденной матрицей, и соответствует преобразованию координат класса $\mathcal{C}^\infty$. В дальнейшем мы всегда будем считать, что все векторные пространства снабжены естественной дифференцируемой структурой евклидова пространства. □

Пример 3.1.3. Рассмотрим ломанную линию γ на евклидовой плоскости $\mathbb{R}^2$, заданную уравнением $y = |x|$ (см. рис. 3.1,*a*). Топология, индуцированная из $\mathbb{R}^2$, превращает ее в связное хаусдорфово топологическое пространство. Ломаную γ можно покрыть одной картой $(\mathbb{U}, \varphi)$, спроектировав ломаную линию на ось x,

$$\mathbb{U} = \{(x, y) \in \mathbb{R}^2 : \quad y = |x|\}, \qquad \varphi(x, y) = x.$$

Таким образом γ становится одномерным тривиальным многообразием класса $\mathcal{C}^\infty$. Однако вложение $\gamma \hookrightarrow \mathbb{R}^2$ (см. раздел 3.9) является только непрерывным, а не дифференцируемым, и это является причиной излома. □

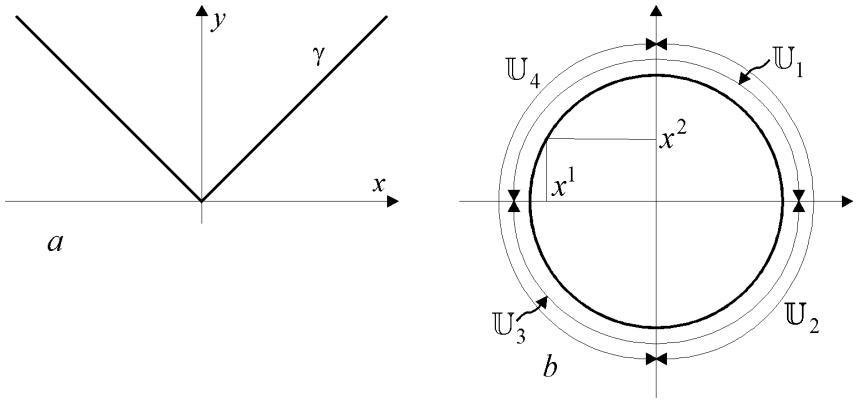


Рис. 3.1: *a)* Ломаная прямая является $\mathcal{C}^\infty$ многообразием. *b)* Покрытие окружности $\mathbb{S}^1$ четырьмя картами. Показана точка, принадлежащая пересечению $\mathbb{U}_1 \cap \mathbb{U}_4$.

Пример 3.1.4. n-мерная сфера $\mathbb{S}^n_r \hookrightarrow \mathbb{R}^{n+1}$ радиуса r с центром в начале координат, вложенная в $(n+1)$-мерное евклидово пространство, задается уравнением

$$\mathbb{S}^n_r := \{x \in \mathbb{R}^{n+1} : \quad (x^1)^2 + (x^2)^2 + \ldots + (x^{n+1})^2 = r^2\}, \qquad r = \mathsf{const} > 0. \tag{3.3}$$

Само по себе это уравнение задает только множество точек в $\mathbb{R}^{n+1}$, а никак не многообразие. Зададим на нем топологию, сказав, что топология индуцирована вложением $\mathbb{S}^n_r \hookrightarrow \mathbb{R}^{n+1}$. Пусть $N = (0, \ldots, 0, r)$ и $S = (0, \ldots, 0, -r)$ – северный и южный полюс сферы. Гладкая дифференцируемая структура на $\mathbb{S}^n_r$ – это полный атлас, содержащий две карты: $(\mathbb{S}^n_r \setminus N, \varphi_N)$ и $(\mathbb{S}^n_r \setminus S, \varphi_S)$, где φ_N и φ_S – стереографические проекции из северного и южного полюса. Тогда сфера становится n-мерным компактным ориентируемым многообразием.Это многообразие нетривиально и покрывается не менее, чем двумя картами. □

Пример 3.1.5. В примере 2.2.16, была построена сфера Римана $\overline{\mathbb{C}}$. Она представляет собой вещественное двумерное многообразие, диффеоморфное обычной сфере, $\overline{\mathbb{C}} \approx \mathbb{S}^2$. С точки зрения комплексной геометрии она является одномерным голоморфным

комплексным многообразием комплексной размерности $\dim_{\mathbb{C}}\overline{\mathbb{C}} = 1$. Действительно, покроем сферу Римана двумя картами: $\mathbb{U}_0 := \mathbb{C}$ – окрестность нуля и $\mathbb{U}_\infty := \overline{\mathbb{C}} \setminus \{0\}$ – окрестность бесконечности. Они имеют непустое пересечение $\mathbb{U}_0 \cap \mathbb{U}_\infty = \mathbb{C} \setminus \{0\}$. В качестве координат на $\mathbb{U}_0$ и $\mathbb{U}_\infty$ выберем, соответственно, z и $1/z$. В области пересечения функция склейки $f_{0\infty}(z) = 1/z$ голоморфны. Тем самым мы построили голоморфный атлас на $\overline{\mathbb{C}}$.

Как комплексное многообразие сфера Римана диффеоморфна одномерному комплексному проективному пространству (проективной прямой), $\overline{\mathbb{C}} \approx \mathbb{CP}^1$. Напомним, что проективная прямая – это множество комплексных прямых $az^1 + bz^2 = 0$, где $(z^1, z^2) \in \mathbb{C} \times \mathbb{C}$ и $a, b \in \mathbb{C}$, в двумерном комплексном многообразии $\mathbb{C}^2 := \mathbb{C} \times \mathbb{C}$. Если $b \neq 0$, то из уравнения прямой следует равенство $z^2 = -az^1/b$, т.е. каждая прямая параметризуется комплексным числом $-a/b = z \in \mathbb{C}$. При $a \neq 0$, аналогично, каждая прямая параметризуется числом $-b/a = 1/z \in \mathbb{C}$. Тем самым мы покрыли проективное пространство $\mathbb{CP}^1$ двумя картами. В области пересечения $ab \neq 0$ функция склейки имеет вид $f(z) = 1/z$. Таким образом, сфера Римана и одномерное комплексное проективное пространство имеют одинаковое координатное покрытие и, следовательно, диффеоморфны. □

Пример 3.1.6. Рассмотрим задание окружности $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2$ единичного радиуса с помощью четырех карт:

$$\begin{aligned}
\mathbb{U}_1 &= \{x \in \mathbb{S}^1 : \quad x^2 > 0\}, \quad \varphi_1(x) = x^1,\\
\mathbb{U}_2 &= \{x \in \mathbb{S}^1 : \quad x^1 > 0\}, \quad \varphi_2(x) = x^2,\\
\mathbb{U}_3 &= \{x \in \mathbb{S}^1 : \quad x^2 < 0\}, \quad \varphi_3(x) = x^1,\\
\mathbb{U}_4 &= \{x \in \mathbb{S}^1 : \quad x^1 < 0\}, \quad \varphi_4(x) = x^2.
\end{aligned}$$

Очевидно, что совокупность областей $\{\mathbb{U}_1, \mathbb{U}_2, \mathbb{U}_3, \mathbb{U}_4\}$ является конечным открытым покрытием окружности $\mathbb{S}^1$. Четыре карты $\{(\mathbb{U}_1, \varphi_1), (\mathbb{U}_2, \varphi_2), (\mathbb{U}_3, \varphi_3), (\mathbb{U}_4, \varphi_4)\}$ представляют собой атлас. В области пересечения двух карт $\mathbb{U}_1 \cap \mathbb{U}_4$ функции преобразования координат имеют вид (см. рис. 3.1,*b*)

$$\begin{aligned}
x^{2'} &= \quad \sqrt{1 - (x^1)^2} > 0,\\
x^{1'} &= -\sqrt{1 - (x^2)^2} < 0,
\end{aligned}$$

где нештрихованные и штрихованные координаты относятся соответственно к областям $\mathbb{U}_1$ и $\mathbb{U}_4$. Аналогично выписываются функции склейки для всех других пересечений карт. Все функции склейки являются бесконечно дифференцируемыми. Это значит, что построенный атлас принадлежит классу $\mathcal{C}^\infty$. □

Пример 3.1.7. Тор $\mathbb{T}^n$ представляет собой прямое произведение n окружностей,

$$\mathbb{T}^n = \underbrace{\mathbb{S} \times \mathbb{S} \times \cdots \times \mathbb{S}}_{n},$$

и является n-мерным компактным ориентируемым многообразием. Дифференцируемая структура на $\mathbb{T}^n$ задается как на прямом произведении многообразий. □

Пример 3.1.8. *Проективным пространством* $\mathbb{RP}^n$ над полем вещественных чисел называется множество прямых евклидова пространства $\mathbb{R}^{n+1}$, проходящих через начало координат. Проективное пространство $\mathbb{RP}^n$ представляет собой многообразие

размерности n. Его можно представлять себе как сферу $\mathbb{S}^n$ с отождествленными диаметрально противоположными точками. Действительно, любая прямая, проходящая через начало координат, пересекает единичную сферу с центром в начале координат ровно в двух диаметрально противоположных точках. Обратно, любая из этих двух точек однозначно определяет прямую, проходящую через начало координат. Таким образом

$$\mathbb{RP}^n \approx \frac{\mathbb{S}^n}{\mathbb{Z}_2},$$

где циклическая группа $\mathbb{Z}_2$ состоит из двух элементов $\{1, -1\}$. Проективное пространство $\mathbb{RP}^n$ можно представить также в виде полусферы $x \in \mathbb{S}^n$, $x^n \geq 0$, у которой отождествлены диаметрально противоположные "граничные" точки, т.е. точки $(n-1)$-мерной сферы

$$\mathbb{S}^{n-1} = \{x \in \mathbb{S}^n : \ x^n = 0\}.$$

Рассмотрим замкнутую кривую в проективном пространстве, проходящую через одну из "граничных" точек. На рис. 3.2,*a*, для наглядности изображена проективная плоскость $\mathbb{RP}^2$ в трехмерном евклидовом пространстве и возможная кривая. Выберем ортонормированный базис вдоль кривой, включающий единичный касательный вектор e_1. При прохождении через "граничную" точку p касательный вектор e_1 не меняет ориентации относительно кривой, в то время как все остальные базисные векторы $e_2, e_3, \ldots$ меняют направление. Это значит, что ориентация ортонормированного базиса при прохождении вдоль этой замкнутой кривой изменится при четных n и сохранится при нечетных n. Если замкнутая кривая целиком лежит в верхней полусфере, то ориентация базиса вдоль кривой сохраняется. Тем самым мы показали, что проективные пространства четного числа измерений неориентируемы, а нечетного числа измерений – ориентируемы. □

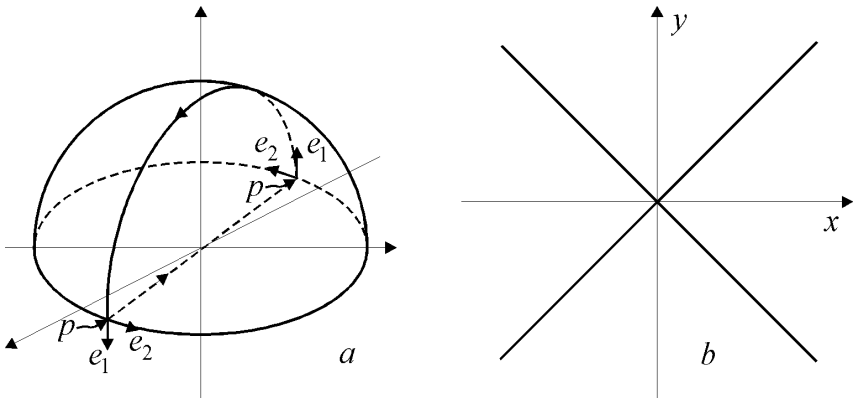


Рис. 3.2: *a)* Проективная плоскость $\mathbb{RP}^2$ как полусфера в $\mathbb{R}^3$. Показан замкнутый путь, проходящий через граничную точку p, и перенос базиса вдоль пути. *b)* Две пересекающиеся прямые не являются многообразием.

Пример 3.1.9. Продемонстрируем отличие топологического пространства от многообразия. Пусть множество точек на евклидовой плоскости состоит из двух пересекающихся прямых $y^2 - x^2 = 0$, изображенных на рис. 3.2,*b*. Пусть топология индуцирована вложением. Тогда это множество представляет собой связное хаусдорфово топологическое пространство. В то же время оно не является многообразием, потому что окрестность точки пересечения прямых нельзя взаимно однозначно отобразить на интервал вещественной прямой $\mathbb{R}$. □

Пример 3.1.10. Рассмотрим прямое произведение двух прямых $\mathbb{M} = \mathbb{R} \times \mathbb{R}_{\text{D}}$. Будем считать, что на первом сомножителе задана естественная топология а на втором – дискретная, что отмечено индексом D. Тогда, как множество, многообразие $\mathbb{M}$ совпадает с двумерным евклидовым пространством $\mathbb{R}^2$, а как многообразие – нет. База топологии $\mathbb{M}$ состоит из всех интервалов на всех прямых, параллельных оси x и проходящих через все точки оси y (см. рис. 3.3). На каждой прямой база топологии счетна. Рассмотрим открытый диск $\mathbb{D} \subset \mathbb{M}$. Так же как и в евклидовом пространстве $\mathbb{R}^2$, он является открытым множеством, но на этот раз как объединение несчетного числа интервалов. Рассмотрим отображение многообразия $\mathbb{M}$ на евклидову плоскость $\mathbb{R}^2$, которое задается простым отождествлением координат $\mathbb{M} \ni (x, y) \overset{f}{\mapsto} (x, y) \in \mathbb{R}^2$. Это отображение биективно и непрерывно. Однако обратное отображение f^{-1} не является непрерывным, т.к. на евклидовой плоскости не существует открытого множества $\mathbb{U} \subset \mathbb{R}^2$, образ которого при обратном отображении $f^{-1}(\mathbb{U})$ лежал бы в интервале. Поэтому отображение f не является гомеоморфизмом и множество $\mathbb{M} = \mathbb{R} \times \mathbb{R}_{\text{D}}$ не является двумерным многообразием. По построению, множество $\mathbb{M}$ является одномерным несвязным многообразием, состоящим из несчетного числа одномерных многообразий – прямых $\mathbb{R}$.

Рассмотренный пример показывает, что на одном и том же множестве точек можно задавать различные структуры. Мы говорим, что на евклидовой плоскости $\mathbb{R}^2$, рассматриваемой, как двумерное многообразие, задана структура слоения, т.е. мы представляем плоскость в виде объединения несчетного числа одномерных подмногообразий – прямых $\mathbb{R}$. Каждая прямая является листом слоения, которые параметризуются точками другой прямой $\mathbb{R}_{\text{D}}$. Более подробно слоения рассмотрены в разделе 3.12. □

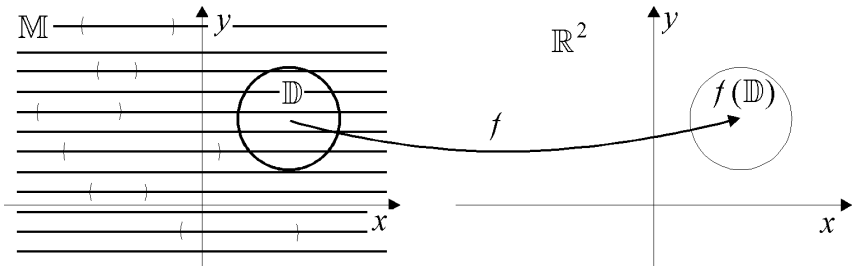


Рис. 3.3: Отображение f объединения $\mathbb{M}$ несчетного числа прямых, параллельных оси x, на евклидову плоскость $\mathbb{R}^2$.

В дальнейшем мы будем изучать различные свойства многообразий. Условимся о терминологии. Будем говорить, что данное свойство выполняется на многообразии $\mathbb{M}$ *глобально*, если оно выполнено во всех точках $x \in \mathbb{M}$. Гораздо чаще встречаются свойства, которые выполнены только *локально*. А именно, для каждой точки $x \in \mathbb{M}$ существует координатная окрестность $\mathbb{U}_x$ такая, что данное свойство выполнено на $\mathbb{U}_x$. В этом случае можно говорить, что данное свойство выполнено в фиксированной системе координат. Конечно, любое свойство, выполненное глобально, справедливо и локально, но не наоборот.

3.2 Разбиение единицы

Один из способов задания геометрических структур, например, метрики, на многообразии $\mathbb{M}$ заключается в следующем. Сначала выбирается некоторый атлас, покры-

вающий $\mathbb{M}$. Затем в каждой карте данного атласа в координатах задается некоторая геометрическая структура. Чтобы задать данную геометрическую структуру на всем многообразии $\mathbb{M}$, ее необходимо склеить в областях пересечения карт. Для этого используется разбиение единицы.

Как уже отмечалось, для существования разбиения единицы нам достаточно рассматривать паракомпактные многообразия. Счетность базы многообразия является достаточным условием паракомпактности. Напомним некоторые определения и утверждения из общей топологии, которые понадобятся для формулировки теоремы о существовании разбиения единицы.

Определение. Топологическое пространство $\mathbb{M}$ является *счетным в бесконечности*, если существует счетное семейство компактных множеств $\mathbb{U}_i$, $i \in \mathbb{N}$, таких, что

$$\mathbb{U}_1 \subset \mathsf{int}\,\mathbb{U}_2 \subset \mathbb{U}_2 \subset \ldots \subset \mathbb{U}_i \subset \mathsf{int}\,\mathbb{U}_{i+1} \subset \mathbb{U}_{i+1} \ldots$$

и

$$\mathbb{M} = \bigcup_{i=1}^{\infty} \mathbb{U}_i. \qquad \square$$

Теорема 3.2.1. *Паракомпактное топологическое пространство является объединением семейства связных паракомпактных топологических пространств, которые являются счетными в бесконечности.*

Доказательство. См., например, [**?**]. $\square$

Теорема 3.2.2. *Любое покрытие паракомпактного многообразия имеет счетное локально конечное измельчение.*

Доказательство. Существование локально конечного измельчения, которое также является покрытием, входит в определение паракомпактности и его доказывать не надо. Нетривиальность утверждения теоремы в том, что локально конечное покрытие можно выбрать счетным. Доказательство приведено в [**?**]. $\square$

Перед тем как ввести разбиение единицы напомним

Определение. *Носителем* функции f, заданной на многообразии $\mathbb{M}$, называется замыкание множества тех точек, в которых она отлична от нуля. Носитель функции обозначается $\mathsf{supp}\, f \subset \mathbb{M}$. Если носитель функции компактен, то функция называется *финитной*. $\square$

Из определения следует, что носитель произвольной функции f всегда замкнут в $\mathbb{M}$.

Определение. *Разбиением единицы*, подчиненным заданному покрытию $\{\mathbb{U}_i\}_{i\in I}$ многообразия $\mathbb{M}$, называется семейство функций $\{f_i\}$, удовлетворяющих следующим условиям:

1) $\mathsf{supp}\, f_i \subset \mathbb{U}_i$ для всех i;
2) каждая точка имеет окрестность $\mathbb{W}$ такую, что $\mathbb{W} \cap \mathsf{supp}\, f_i = \emptyset$ за исключением конечного числа функций f_i;
3) $0 \le f_i(x) \le 1, \quad \sum_i f_i(x) = 1, \quad \forall x \in \mathbb{M}.$ $\square$

Замечание. В данном определении покрытие произвольно и не обязательно счетное и локально конечное. Сумма в условии 3) определена, т.к. содержит только конечное число слагаемых в силу условия 2). $\square$

Теорема 3.2.3 (Разбиение единицы). *На любом паракомпактном многообразии класса $\mathcal{C}^k$ существует $\mathcal{C}^k$ разбиение единицы, подчиненное заданному счетному локально конечному покрытию.*

Доказательство. См., например, [?, ?]. □

Замечание. Разбиение единицы существует как для ориентируемых, так и для неориентируемых многообразий. □

Для заданного покрытия $\{\mathbb{U}_i\}$ существует много разбиений единицы. Пусть $\{f_i\}$ и $\{g_j\}$ – два разбиения единицы, подчиненные покрытию $\{\mathbb{U}_i\}$. Тогда очевидна формула

$$\sum_i f_i \sum_j g_j = \sum_{ij} f_i g_j = \sum_j g_j \sum_i f_i,$$

поскольку в каждой точке $x \in \mathbb{M}$ суммы содержат только конечное число слагаемых.

Существование разбиения единицы на многообразии является эффективным средством доказательства существования геометрических структур на многообразии путем склеивания этих структур, заданных в отдельных картах. Например, в теореме 5.1.1 доказано существование римановой метрики на произвольном многообразии.

Замечание. Условие паракомпактности, являющееся достаточным условием существования разбиения единицы можно заменить на вторую аксиому счетности, утверждающую существование счетной базы топологии. Если потребовать, чтобы многообразие $\mathbb{M}$ являлось хаусдорфовым топологическим пространством, удовлетворяющим второй аксиоме счетности, то отсюда будет следовать паракомпактность многообразий и, кроме того, нормальность и метризуемость. (Доказательство метризуемости топологических пространств, удовлетворяющих второй аксиоме счетности, приведено, например, в [?].) □

3.3 Многообразия с краем

Определение. Пусть $f(x) \in C^\infty(M)$ – гладкая функция на многообразии $\mathbb{M}$. Тогда замкнутое множество $\mathbb{A} \subset \mathbb{M}$, выделяемое в многообразии $\mathbb{M}$ неравенством $f(x) \leq 0$ (или $f(x) \geq 0$) называется *многообразием с краем*. Подмногообразие $\partial\mathbb{A} \subset \mathbb{M}$, задаваемое уравнением $f(x) = 0$, называется *краем* $\mathbb{A}$. При этом мы предполагаем, что градиент функции f на крае $\partial\mathbb{A}$ отличен от нуля. □

Если функция положительна, $f > 0$, на $\mathbb{M}$, то край пустой, $\partial\mathbb{A} = \emptyset$, и $\mathbb{A} = \mathbb{M}$. Нетривиальное многообразие с краем получается, если область значений функции f включает нуль.

Пример 3.3.1. Замыкание любой ограниченной открытой области $\mathbb{U} \subset \mathbb{R}^n$ в евклидовом пространстве $\mathbb{R}^n$ является многообразием с краем $\overline{\mathbb{U}}$. При этом краем является граница области: $\partial\overline{\mathbb{U}} = \overline{\mathbb{U}} \setminus \mathbb{U}$. □

Замечание. Условие отличия градиента функции от нуля на крае является достаточным условием того, что край является $(n-1)$-мерным подмногообразием в $\mathbb{M}$. По сути дела это определение является инвариантным обобщением понятия области и ее границы в евклидовом пространстве $\mathbb{R}^n$. Если многообразие с краем можно покрыть одной картой, то оба понятия в точности совпадают. □

Пример 3.3.2. Ориентируемым многообразием с краем является цилиндр конечной высоты, который получается при склейке двух краев прямоугольника, показанного на рис. 3.4,*a*. При этом направление склеиваемых сторон, которое показано стрелками, сохраняется. Цилиндр можно покрыть двумя картами, которые пересекаются по двум областям, и в обеих областях якобиан перехода положителен. Краем цилиндра является объединение двух окружностей. □

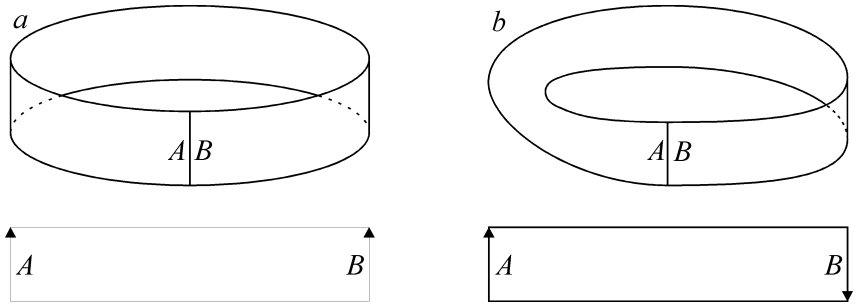


Рис. 3.4: Цилиндр *(a)* и лента Мебиуса *(b)* получаются при склеивании двух сторон прямоугольника с сохранением и изменением направления сторон при склейке, как показано стрелками.

Пример 3.3.3. Если перед склейкой прямоугольника изменить направление одной из сторон, как показано стрелками на рис. 3.4,*b*, то получится неориентируемая поверхность с краем, которая называется *листом Мебиуса*. Лист Мебиуса также можно покрыть двумя картами, пересекающимися по двум областям, но в этом случае якобиан перехода в этих областях будет иметь разный знак. У листа Мебиуса край диффеоморфен одной окружности и является связным. Окружность, возникающая при склеивании отрезка $[A, B]$, называется центральной. □

Если $\mathbb{A}$ – многообразие с краем, то разность $\mathbb{A}\backslash\partial\mathbb{A}$ также является многообразием, но уже без края.

Дадим эквивалентное определение многообразия с краем. Обозначим полупространство евклидова пространства $\mathbb{R}^n$, определяемое уравнением $x^n > 0$, через $\mathbb{R}^n_+$. Будем считать, что на замыкании $\overline{\mathbb{R}}^n_+ \subset \mathbb{R}^n$ задана индуцированная топология. При этом открытые множества в $\overline{\mathbb{R}}^n_+$ могут как содержать, так и не содержать точки края, определяемого уравнением $x^n = 0$.

Определение. Топологическое хаусдорфово пространство $\mathbb{M}$ со счетной базой, каждая точка которого имеет окрестность, гомеоморфную открытому множеству в $\overline{\mathbb{R}}^n_+$, называется n-мерным *многообразием с краем*. Точки, которые имеют окрестность, гомеоморфную $\mathbb{R}^n$, называются *внутренними*. Остальные точки называются *краевыми*. □

Дифференцируемая структура на многообразии с краем вводится так же как и на многообразии без края. Если не оговорено противное, то мы рассматриваем гладкие дифференцируемые структуры класса $\mathcal{C}^\infty$.

Теорема 3.3.1. *Пусть* $\mathbb{M}$, $\dim\mathbb{M} = n$, *– многообразие с краем. Если край* $\partial\mathbb{M}$ *не является пустым, то* $\partial\mathbb{M}$ *представляет собой многообразие размерности* $n-1$ *и без края* $\partial(\partial\mathbb{M}) = \emptyset$.

Доказательство. См., например, []. □

Определение. Многообразие с краем $\mathbb{M}$ называется *ориентируемым*, если ориентируемо соответствующее ему многообразие без края $\mathbb{M} \setminus \partial\mathbb{M}$. □

Теорема 3.3.2. *Если многообразие с краем $\mathbb{M}$ ориентируемо, то его край $\partial\mathbb{M}$ также является ориентируемым многообразием. Ориентация на $\mathbb{M}$ индуцирует каноническую ориентацию края $\partial\mathbb{M}$.*

Доказательство. Пусть на $\mathbb{M} \setminus \partial\mathbb{M}$ задана какая либо ориентация, которую назовем положительной. Рассмотрим естественное вложение края $h : \partial\mathbb{M} \to \mathbb{M}$. Тогда дифференциал отображения действует на касательные векторы $h_* : \mathbb{T}_x(\partial\mathbb{M}) \to \mathbb{T}_x(\mathbb{M})$. Для каждой точки края $x \in \partial\mathbb{M}$ выберем первый базисный вектор $e_1 \in \mathbb{T}_x(\mathbb{M})$, $e_1 \notin \mathbb{T}_x(\partial\mathbb{M})$ таким образом, чтобы он был ориентирован наружу. Это значит, что для любой дифференцируемой функции, удовлетворяющей условиям $f(\partial\mathbb{M}) = 0$ и $f(\mathbb{M} \setminus \partial\mathbb{M}) \leq 0$, производная вдоль e_1 неотрицательна $e_1 f \geq 0$. Дополним этот вектор базисными векторами края $\{e_2, \ldots, e_n\}$ таким образом, чтобы базис $\{e_1, \ldots, e_n\}$ имел положительную ориентацию в $\mathbb{M}$. Тогда ориентация $\{e_2, \ldots, e_n\}$ задает *каноническую ориентацию* края. □

На языке дифференциальных форм задание согласованной ориентации многообразия $\mathbb{M}$ и его края $\partial\mathbb{M}$ означает следующее. Пусть $\{e^a\}$, $a = 1, \ldots, n$, – набор 1-форм, дуальных к базису e_a, построенному в доказательстве теоремы 3.3.2: $e^a(e_b) = \delta^a_b$. Поскольку 1-формы линейно независимы, то n-форма $e^1 \wedge e^2 \wedge \ldots \wedge e^n$ отлична от нуля и задает ориентацию на $\mathbb{M}$. Формы $e^2, \ldots, e^n$, по-построению, линейно независимы на крае $\partial\mathbb{M}$ и задают каноническую ориентацию края. В дальнейшем мы всегда предполагаем, что ориентация края $\partial\mathbb{M}$ индуцирована ориентацией самого многообразия $\mathbb{M}$.

Из определения края следует выражение для края прямого произведения двух многообразий (правило Лейбница)

$$\partial\Big(\mathbb{M}_1 \times \mathbb{M}_2\Big) = \Big(\partial\mathbb{M}_1 \times \mathbb{M}_2\Big) \cup \Big(\mathbb{M}_1 \times \partial\mathbb{M}_2\Big).$$

Определение. В моделях гравитации принято называть вселенную *замкнутой*, если она представляет собой компактное многообразие без края. *Открытая вселенная* является некомпактным многообразием без края. В общем случае будем называть компактное многообразие без края *замкнутым*. Такие многообразия являются, конечно, замкнутыми множествами в топологическом смысле. Однако термин замкнутый в данном определении и в определении замкнутого множества в топологии имеют разный смысл. □

Пример 3.3.4. Прямая представляет собой замкнутое (и одновременно открытое) множество в естественной топологии. В то же время она не является замкнутым многообразием, поскольку некомпактна. □

3.4 Расслоения

В разделе 3.1 было определено прямое произведение $\mathbb{M} \times \mathbb{F}$ двух дифференцируемых многообразий $\mathbb{M}$ и $\mathbb{F}$, которое также является дифференцируемым многообразием. При этом дифференцируемая структура на прямом произведении многообразий определяется дифференцируемыми структурами сомножителей. В настоящем разделе мы обобщим понятие прямого произведения многообразий.

Определение. *Расслоением* называется четверка $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$, где $\mathbb{E}, \mathbb{M}, \mathbb{F}$ – многообразия, а $\pi : \mathbb{E} \to \mathbb{M}$ – дифференцируемое отображение, которые удовлетворяют следующим условиям:

1) каждая точка $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}_x$ такую, что существует диффеоморфизм $\chi : \pi^{-1}(\mathbb{U}_x) \to \mathbb{U}_x \times \mathbb{F}$ (локальная тривиальность);
2) композиция отображений $\pi \circ \chi^{-1} : \mathbb{U}_x \times \mathbb{F} \to \mathbb{M}$ есть проекция на первый сомножитель: $(y, v) \mapsto y$ для всех $y \in \mathbb{U}_x$ и $v \in \mathbb{F}$.

Многообразие $\mathbb{E}$ называется *пространством расслоения*, $\mathbb{M}$ – *базой* расслоения, $\mathbb{F}$ – *типичным слоем* и π – *проекцией*. □

Поскольку отображение χ в условии 1) является диффеоморфизмом, то

$$\dim \mathbb{E} = \dim \mathbb{M} + \dim \mathbb{F}.$$

Мы предполагаем, что все многообразия являются дифференцируемыми. Кроме того база предполагается связным многообразием. В противном случае расслоения можно рассматривать над каждой связной компонентой. Расслоение также называется *расслоенным пространством* и обозначается $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$. Мы считаем, что проекция π является дифференцируемым отображением на базу $\mathbb{M}$, т.е. сюрьективным. В противном случае можно рассмотреть расслоение с базой $\pi(\mathbb{E}) \subset \mathbb{M}$. Поскольку отображение π непрерывно, то согласно теореме 2.4.3 прообраз $\pi^{-1}(x)$ является замкнутым подмногообразием в $\mathbb{E}$, которое называется *слоем* над точкой базы $x \in \mathbb{M}$.

В дальнейшем, для краткости, мы иногда будем обозначать расслоение одной буквой $\mathbb{E}$. На рис. 3.5 схематично показана структура расслоения.

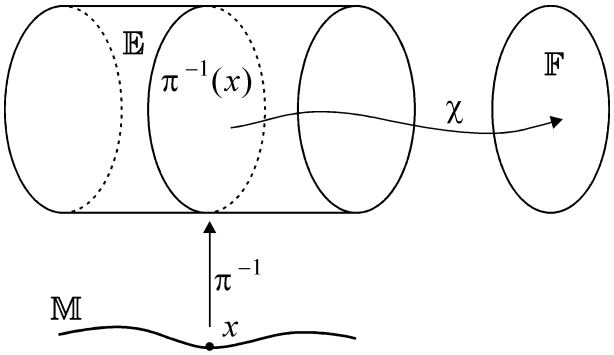
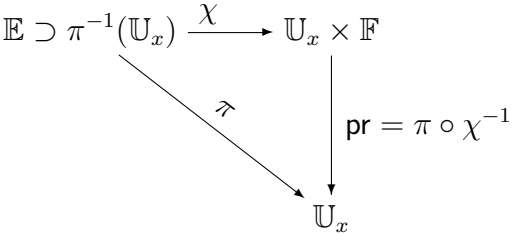


Рис. 3.5: Схематичное изображение структуры расслоения. $\mathbb{E}$ – пространство расслоения, $\mathbb{M}$ – база, $\mathbb{F}$ – типичный слой, $\pi : \mathbb{E} \to \mathbb{M}$ – проекция, $\chi : \pi^{-1}(x) \to \mathbb{F}$ – диффеоморфизм.

Замечание. Атласы, заданные на базе и типичном слое, определяют атлас на пространстве расслоения в силу локальной тривиальности расслоения. Тем самым первое условие в определении расслоения является достаточным для того, чтобы дифференцируемые структуры на трех многообразиях $\mathbb{E}, \mathbb{M}$ и $\mathbb{F}$ были согласованы между собой. Второе требование является условием коммутативности диаграммы



где pr – естественная проекция прямого произведения $\mathbb{U}_x \times \mathbb{F} \xrightarrow{\mathsf{pr}} \mathbb{U}_x$ на первый сомножитель. В общем случае это могло бы быть не так. В приведенной диаграмме окрестность $\mathbb{U}_x$ нельзя заменить на все многообразие $\mathbb{M}$, так как отображение χ определено локально. □

Пример 3.4.1. Прямое произведение двух многообразий $\mathbb{E} = \mathbb{M} \times \mathbb{F}$ с проекцией на первый сомножитель $\mathbb{M} \times \mathbb{F} \xrightarrow{\pi} \mathbb{M}$ является расслоением с базой $\mathbb{M}$ и типичным слоем $\mathbb{F}$, которое называется *тривиальным*. С равным успехом четверка $\mathbb{M} \times \mathbb{F} \xrightarrow{\pi'} \mathbb{F}$ с проекцией на второй сомножитель является тривиальным расслоением с базой $\mathbb{F}$ и типичным слоем $\mathbb{M}$. □

Многообразия $\mathbb{M}, \mathbb{F}, \mathbb{E}$ могут быть как с краем, так и без края. Мы допускаем также в качестве базы или типичного слоя 0-мерные многообразия, т.е. конечные или счетные наборы точек с дискретной топологией.

По-построению, пространство расслоения $\mathbb{E}$ представляет собой объединение несчетного числа слоев, $\mathbb{E} = \bigcup_{x\in\mathbb{M}} \pi^{-1}(x)$, каждый из которых диффеоморфен типичному слою $\mathbb{F}$ и "нумеруется" точкой базы. Это – частный случай слоений, рассмотренных в разделе 3.12.

Как было отмечено, дифференцируемые структуры на базе и слое согласованы с дифференцируемой структурой на пространстве расслоения $\mathbb{E}$. Опишем это более подробно. Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ – расслоение, $\mathsf{dim}\,\mathbb{M} = n$, $\mathsf{dim}\,\mathbb{F} = k$. По-определению, отображение χ действует на каждую точку $p \in \mathbb{E}$: $\chi(p) = \Big(x = \pi(p), v\Big) \in \mathbb{M} \times \mathbb{F}$. Поскольку база и типичный слой – многообразия, то существуют карты:

$$(\mathbb{U}_x, \varphi): \quad \mathbb{M} \supset \mathbb{U}_x \xrightarrow{\varphi} \varphi(\mathbb{U}_x) \subset \mathbb{R}^n,$$
$$(\mathbb{V}_v, \phi): \quad \mathbb{F} \supset \mathbb{V}_v \xrightarrow{\phi} \phi(\mathbb{V}_v) \subset \mathbb{R}^k.$$

Локально определено отображение из пространства расслоения $\mathbb{E}$ в евклидово пространство $\mathbb{R}^{n+k}$:

$$\rho: \ \chi^{-1}(\mathbb{U}_x \times \mathbb{V}_v) \ \to \ \varphi(\mathbb{U}_x) \times \phi(\mathbb{V}_v) \in \mathbb{R}^{n+k}$$

Таким образом, для каждой точки расслоения $p \in \mathbb{E}$ определена карта $\Big(\chi^{-1}(\mathbb{U}_x \times \mathbb{V}_v), \rho\Big)$. Определено также отображение, действующее из $\mathbb{R}^n \times \mathbb{R}^k$ в $\mathbb{R}^n$:

$$\tilde{\pi} = \varphi \circ \pi \circ \rho^{-1}: \quad \varphi(\mathbb{U}_x) \times \phi(\mathbb{V}_v) \ni (y, w) \ \mapsto \ y \in \varphi(\mathbb{U}_x),$$

где $y \in \varphi(\mathbb{U}_x) \subset \mathbb{R}^n$ и $w \in \phi(\mathbb{V}_f) \subset \mathbb{R}^k$.

Определение. Дифференцируемое отображение $\mathbb{M} \xrightarrow{\sigma} \mathbb{E}$ называется *сечением* или *глобальным сечением* расслоения $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$, если $\pi \circ \sigma = \mathsf{id}\,(\mathbb{M})$. Аналогичным образом, дифференцируемое отображение $\mathbb{U} \xrightarrow{\sigma} \mathbb{E}$, где $\mathbb{U}$ – открытое подмножество базы $\mathbb{M}$, называется *локальным сечением* расслоения $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$, если $\pi \circ \sigma = \mathsf{id}\,(\mathbb{U})$. □

Локальные сечения существуют у любого расслоения – это функции со значениями в пространстве расслоения $\mathbb{E}$, которые определены в областях $\mathbb{U} \subset \mathbb{M}$. Глобальные сечения расслоений, как мы увидим в дальнейшем, существуют далеко не всегда.

Пример 3.4.2. Пусть $\mathbb{E} = \mathbb{M} \times \mathbb{R}$ – тривиальное расслоение, типичным слоем которого является поле вещественных чисел. Тогда множество всех гладких сечений

$$f: \quad \mathbb{M} \ni x \mapsto \Big(x, f(x)\Big) \in \mathbb{E} = \mathbb{M} \times \mathbb{R}$$

совпадает с множеством графиков всех вещественнозначных бесконечно дифференцируемых функций $\mathcal{C}^\infty(\mathbb{M})$ □

Пример 3.4.3 (**Цилиндр**). Цилиндр конечной высоты – это прямое произведение окружности на единичный отрезок $\mathbb{S}^1 \times [0,1]$. Его можно рассматривать как расслоение с базой $\mathbb{M} = \mathbb{S}^1$ и типичным слоем $[0,1]$. У этого расслоения существуют глобальные сечения – гладкие функции на окружности со значениями в единичном отрезке $[0,1]$. Например, $f(\varphi) = \sin\varphi$, где $\varphi \in [0, 2\pi]$ – координата на окружности. □

Пример 3.4.4 (**Лист Мёбиуса**). Лист Мебиуса, изображенный на рис.**??**, является расслоением с базой $\mathbb{S}^1$, типичным слоем которого, как и у цилиндра, является единичный отрезок $[0,1]$. Это расслоение нетривиально, т.к. не имеет вида прямого произведения. □

Пример 3.4.5 (**Бутылка Клейна**). Построение бутылки Клейна изображено на рис.**??**. Мы берем цилиндр конечной высоты и отождествляем точки граничных окружностей, предварительно отобразив точки окружности с одной стороны цилиндра относительно произвольного диаметра. Эту поверхность нельзя вложить в трехмерное евклидово пространство $\mathbb{R}^3$ и поэтому трудно представить. Бутылку Клейна можно рассматривать, как расслоение с базой $\mathbb{M} = \mathbb{S}^1$ и типичным слоем $\mathbb{F} = \mathbb{S}^1$. Это расслоение нетривиально. □

Пример 3.4.6. Пусть задана группа Ли $\mathbb{G}$ и ее нормальная подгруппа $\mathbb{H} \subset \mathbb{G}$. Рассмотрим отображение группового многообразия $\mathbb{G}$ на пространство правых (или левых) смежных классов $\pi : \mathbb{G} \to \mathbb{G}/\mathbb{H}$, определенное правилом $\pi(g) = \mathbb{H}g$, где $g \in \mathbb{G}$. Тогда $\mathbb{G} \xrightarrow{\pi} \mathbb{G}/\mathbb{H}$ – расслоение. Пространством расслоения является группа Ли $\mathbb{G}$, базой – факторгруппа $\mathbb{G}/\mathbb{H}$ и типичным слоем – нормальная подгруппа $\mathbb{H}$. Дифференцируемая структура на фактор пространстве $\mathbb{G}/\mathbb{H}$ будет определена позже в теореме 10.1.1. Если размерность группы равна размерности нормальной подгруппы, $\dim\mathbb{G} = \dim\mathbb{H}$, то базой является конечный или счетный набор точек, т.е. 0-мерное многообразие. □

Замечание. В определении расслоения общего вида мы не предполагаем наличия каких либо структур в типичном слое, кроме структуры дифференцируемого многообразия. В дальнейшем мы рассмотрим частные случаи расслоенных пространств, когда типичным слоем является векторное пространство (векторное расслоение) или группа Ли (главное расслоение). □

3.5 Скалярные поля и плотности

В моделях математической физики, как правило, постулируется, что пространство и пространство-время, в котором мы живем, являются многообразиями. Само по себе это очень глубокое предположение. Однако для построения физических моделей его недостаточно. Для описания движения и взаимодействия различных физических объектов в пространстве-времени необходимо задание дополнительных структур на многообразии. Такими структурами являются скалярные и векторные поля, r-формы, тензорные поля, метрика и связность, которые, в частности, характеризуются различными трансформационными свойствами при преобразовании координат. Начнем с простейшего объекта – скалярного поля (функции).

Рассмотрим вещественнозначную функцию f на многообразии $\mathbb{M}$, $\dim\mathbb{M} = n$, т.е. отображение

$$f: \quad \mathbb{M} \to \mathbb{R}. \tag{3.4}$$

Это отображение часто, особенно в физических приложениях, называют *скалярным полем* на $\mathbb{M}$. По-определению отображения (2.54) скалярное поле должно быть однозначно. Функция называется *дифференцируемой* класса $\mathcal{C}^k$, если отображение (3.4), заданное в координатах,

$$f \circ \varphi_i^{-1} : \quad \mathbb{R}^n \supset \varphi_i(\mathbb{U}_i) \ni x \mapsto f(x) \in \mathbb{R}$$

k раз непрерывно дифференцируемо в каждой карте атласа $\{\mathbb{U}_i, \varphi_i\}$. Конечно, не имеет смысла говорить о степени гладкости функции, которая превышает степень гладкости дифференцируемой структуры многообразия. Поэтому мы предполагаем, что степень гладкости функции меньше или равна степени гладкости многообразия.

В двух областях $\mathbb{U}_i$ и $\mathbb{U}_j$ скалярное поле задается, соответственно, двумя функциями $f(x)$ и $f'(x')$ от n переменных $x = \{x^\alpha\}$ и $x' = \{x^{\alpha'}\}$, $\alpha, \alpha' = 1, \ldots, n$. Если области пересекаются, то в области пересечения согласно, (2.55), справедливо равенство

$$f'\Big(x'(x)\Big) = f(x), \tag{3.5}$$

поскольку в каждой точке функция имеет только одно значение. Формулу (3.5) можно интерпретировать, как правило преобразования функции при замене координат $x^\alpha \mapsto x^{\alpha'}(x)$. Другими словами, значение функции после преобразования в точке x' равно ее прежнему значению в точке x.

Замечание. В дальнейшем мы будем позволять себе некоторую вольность в обозначениях. Запись $f(x)$ в зависимости от контекста будет пониматься двояко. Во-первых, $f(x)$ обозначает значение функции f в произвольной точке многообразия $x \in \mathbb{M}$ безотносительно какой либо карты. Это не есть функция n вещественных переменных. Во-вторых, запись $f(x)$ обозначает также значение функции $f \circ \varphi^{-1}$ в точке $\{x^\alpha\} \in \mathbb{R}^n$. Это – обычная функция от n вещественных переменных (координат). Строго говоря, частная производная от функции на многообразии $\partial_\alpha f$ не определена, т.к. мы не знаем, что такое разность двух точек $x_1 - x_2$ на многообразии. Тем не менее мы будем употреблять запись $\partial_\alpha f$, принимая

$$\partial_\alpha f := \partial_\alpha (f \circ \varphi^{-1}).$$

Фактически это означает, что в каждой отдельно взятой карте мы отождествляем точки многообразия с точками евклидова пространства: $x \sim \varphi(x) = \{x^\alpha\} \in \mathbb{R}^n$, и функции: $f \sim f \circ \varphi^{-1}$. При проведении вычислений в одной карте это не приводит к какой либо путанице. □

Множество всех функций класса $\mathcal{C}^k$ на многообразии $\mathbb{M}$ обозначим $\mathcal{C}^k(\mathbb{M})$. Множество гладких (бесконечно дифференцируемых) функций обозначим $\mathcal{C}^\infty(\mathbb{M})$. Скалярное поле называется *тривиальным*, если оно равно нулю на $\mathbb{M}$.

На множестве функций $\mathcal{C}^k(\mathbb{M})$ определим две поточечные операции: сложение и умножение:

$$(f + g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x), \qquad \forall x \in \mathbb{M}.$$

То есть значения суммы и произведения двух функций в данной точке равно соответственно сумме и произведению значений этих функций в той же точке. Очевидно, что сумма и произведение двух функций снова дает функцию. По отношению к этим операциям функции образуют коммутативное кольцо. Кроме этого функции

можно умножать на действительные числа. Умножение на числа вместе с операцией сложения превращает множество функций в векторное пространство над полем вещественных чисел. Если на множестве функций рассматривать все три операциями (умножение на числа, сложение и умножение функций), то оно образует коммутативную ассоциативную алгебру с единицей над полем вещественных чисел, которую так же будем обозначать $\mathcal{C}^k(\mathbb{M})$. Эта алгебра является бесконечномерной.

В разделе 2.5 было показано, что координаты точки евклидова пространства сами можно рассматривать, как набор функций. Для каждой карты $(\mathbb{U},\varphi)$ многообразия $\mathbb{M}$ определен набор функций $x^\alpha(x)$, $\alpha=1,\dots,n$ от точки многообразия $x\in\mathbb{U}$,

$$\varphi(x)=\{x^\alpha(x)\}\in\mathbb{R}^n, \tag{3.6}$$

которые называются *координатными функциями*. Эти функции свои для каждой карты. Во многих случаях координатные функции, определенные на $\mathbb{U}\subset\mathbb{M}$, упрощают запись, позволяя опускать символ отображения $\varphi(x)$.

Определим новый геометрический объект – скалярную плотность h. С этой целью рассмотрим две пересекающиеся карты $\mathbb{U}_i\cap\mathbb{U}_j\neq\emptyset$. При преобразовании координат можно умножить функцию h_i, заданную в карте $\mathbb{U}_i$, на якобиан преобразования (2.58) в степени p:

$$h_j\Big(x'(x)\Big)=J^p_{ji}(x)h_i(x), \tag{3.7}$$

Пусть $\{\mathbb{U}_i,\varphi_i\}$ – некоторый атлас на многообразии $\mathbb{M}$. В каждой карте $(\mathbb{U}_i,\varphi_i)$ зададим отображение

$$h_i:\quad \mathbb{R}^n\supset\varphi_i(\mathbb{U}_i)\ni x=\{x^\alpha\}\mapsto h_i(x)\in\mathbb{R}$$

таким образом, что в области пересечения карт $\mathbb{U}_i\cap\mathbb{U}_j$ выполнен закон преобразования (3.7) для всех $x\in\mathbb{U}_i\cap\mathbb{U}_j$. Это определение корректно, т.к. в области пересечения трех карт $\mathbb{U}_i\cap\mathbb{U}_j\cap\mathbb{U}_k$ с координатами $x=\{x^\alpha\}$, $x'=\{x^{\alpha'}\}$ и $x''=\{x^{\alpha''}\}$ выполнено равенство

$$h_k\Big(x''\Big(x'(x)\Big)\Big)=J^p_{kj}h_j\Big(x'(x)\Big)=J^p_{kj}J^p_{ji}h(x)=J^p_{ki}h(x),$$

где мы воспользовались равенством $J_{kj}J_{ji}=J_{ki}$ для якобианов преобразования координат $x\mapsto x'(x)$, $x'\mapsto x''(x')$ и $x\mapsto x''(x)$.

Преобразования (3.7) образуют группу. Действительно, якобианы преобразований координат, по-определению, отличны от нуля, якобиан двух последовательных преобразований равен произведению якобианов, а якобиан обратного преобразования равен J^{-1}.

Определение. Геометрический объект $h=\{h_i\}$, заданный в некотором атласе $\{\mathbb{U}_i,\varphi_i\}$ многообразия $\mathbb{M}$ с правилом преобразования (3.7) в области пересечения любых двух карт называется скалярной *плотностью степени* p, и мы будем писать $\mathsf{deg}\,h=p$. □

Строго говоря, скалярная плотность не является функцией в смысле определения (3.4), и мы не можем писать $h(x)$, где x – точка многообразия. Имеет смысл лишь запись $h(x^\alpha)$ для функции $h\circ\varphi^{-1}$, которая задана в координатном евклидовом пространстве. Несмотря на это, использование скалярных плотностей, например, при интегрировании, бывает удобным. Кроме того, производить вычисления с плотностями часто бывает проще, чем с тензорами, как, например, в общей теории относительности.

Забегая вперед, отметим, что поскольку определитель репера $\det e_\alpha{}^a \neq 0$ является плотностью веса -1, то произвольную скалярную плотность степени p можно представить в виде

$$h = e^{-p} f,$$

где $f = e^p h$ – скалярное поле (функция).

Замечание. Множество скалярных плотностей фиксированной степени алгебры не образует, т.к. произведение двух плотностей степеней p_1 и p_2 дает скалярную плотность степени $p_1 + p_2$. □

3.6 Векторные поля и 1-формы

3.6.1 Локальное определение

Начнем с локального определения векторных полей и 1-форм, которое является более наглядным. Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$. Ограничим наше рассмотрение двумя пересекающимися картами $(\mathbb{U}_i, \varphi_i)$ и $(\mathbb{U}_j, \varphi_j)$ с координатами x^α и $x^{\alpha'}$ соответственно. В области пересечения этих карт (или при преобразовании координат $x^\alpha \to x^{\alpha'}(x)$) дифференциалы умножаются на матрицу Якоби, а частные производные – на ее обратную:

$$dx^{\alpha'} = dx^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha}, \tag{3.8}$$

$$\partial_{\alpha'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \partial_\alpha. \tag{3.9}$$

Матрицы преобразования дифференциалов и частных производных являются взаимно обратными по правилу дифференцирования сложных функций:

$$\frac{\partial x^\alpha}{\partial x^\beta} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^{\alpha'}}{\partial x^\beta} = \delta^\alpha_\beta. \tag{3.10}$$

Векторные поля и 1-формы на многообразии определяются, исходя из правила преобразования дифференциалов и частных производных. А именно, n достаточно гладких функций $X^\alpha(x)$, заданных на карте $(\mathbb{U}_i, \varphi_i)$ и преобразующихся по правилу (3.8),

$$X^{\alpha'} = X^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha}, \tag{3.11}$$

при преобразовании координат, называются *компонентами векторного* или *контравариантного векторного* поля. Аналогично, n достаточно гладких функций $A_\alpha(x)$, преобразующихся по правилу (3.9),

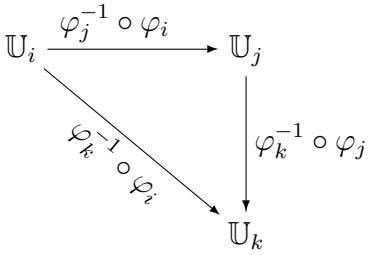
$$A_{\alpha'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} A_\alpha. \tag{3.12}$$

называются *компонентами ковекторного*, или *ковариантного векторного* поля, или 1-*формы*. 1-формы называются также *формами Пфаффа*. В общем случае (ко-)векторное поле имеет n независимых компонент. Каждая из функций X^α или A_α является компонентой векторного или ковекторного поля относительно координатных базисов $e_\alpha = \partial_\alpha$ и $e^\alpha = dx^\alpha$. Смысл обозначения координатных базисов частными производными и дифференциалами не случаен и будет ясен из дальнейшего.

Для того, чтобы задать компоненты (ко-)векторного поля на всем многообразии, их необходимо задать в каждой карте некоторого атласа $\{\mathbb{U}_i, \varphi_i\}$ таким образом, чтобы во всех областях пересечения карт $\mathbb{U}_i \cap \mathbb{U}_j$ они были связаны преобразованием (3.11). Это определение непротиворечиво. Действительно, если точка x лежит в пересечении трех карт $(\mathbb{U}_i, \varphi_i)$, $(\mathbb{U}_j, \varphi_j)$ и $(\mathbb{U}_k, \varphi_k)$ с координатами x^α, $x^{\alpha'}$ и $x^{\alpha''}$ соответственно, то компоненты векторов преобразуются по правилу:

$$X^{\alpha''} = X^{\alpha'} \frac{\partial x^{\alpha''}}{\partial x^{\alpha'}} = X^{\alpha} \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \frac{\partial x^{\alpha''}}{\partial x^{\alpha'}} = X^{\alpha} \frac{\partial x^{\alpha''}}{\partial x^{\alpha}}.$$

Данное равенство является следствием правила дифференцирования сложных функций и означает коммутативность следующей диаграммы



для всех точек $x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k$.

Само векторное поле X на многообразии $\mathbb{M}$ является инвариантным геометрическим объектом и не зависит от выбора системы координат. Выше мы определили *компоненты* векторного поля на многообразии $\mathbb{M}$ путем их задания в некотором атласе. Далее мы должны показать, что таким образом определенный геометрический объект – векторное поле – не зависит от выбора атласа. Это также обеспечено правилом преобразования (3.11). Таким образом, для задания векторного поля X на многообразии $\mathbb{M}$ достаточно задать его компоненты в некотором атласе и указать правило преобразования (3.11) или (3.12). Аналогично дается глобальное определение ковекторного поля.

Пример 3.6.1. Если векторное поле имеет нулевые компоненты в одной системе координат, то они равны нулю и во всех других системах. Нулевое векторное поле, компоненты которого равны нулю во всех картах, называется *тривиальным*. Это единственное векторное поле, компоненты которого инвариантны относительно преобразований координат. Аналогично определяется нулевая 1-форма. □

Пример 3.6.2. Дифференциалы dx^α, рассматриваемые как функции от точки $x \in \mathbb{U}$, являются компонентами гладкого векторного поля. Это векторное поле определено в произвольной карте, а в областях пересечения карт справедливо равенство (3.8). Сами координатные функции $x^\alpha(x)$, хотя и имеют векторный индекс, векторного поля не образуют. Это – набор скалярных функций. □

Пример 3.6.3. Частные производные от произвольной функции $f \circ \varphi$ на образе $\varphi(\mathbb{U}) \subset \mathbb{R}^n$ являются компонентами ковариантного векторного поля $\{\partial_\alpha f\}$, которое называется градиентом функции. Это ковекторное поле определено в произвольной карте и имеет правильный закон преобразования (3.12). □

Пример 3.6.4. Примером векторного поля на кривой $\gamma = \{x^\alpha(t)\} \in \mathbb{M}$ является вектор скорости (2.9). Действительно, при преобразовании координат компоненты вектора скорости $\dot{x}^\alpha$ преобразуются, как дифференциалы. При этом вектор скорости

рассматривается в точке кривой $\gamma \in \mathbb{M}$. В то же время сами координатные функции $x^\alpha(t)$ определены на отрезке $t \in [0,1]$, а не на многообразии и векторного поля не образуют. Векторное поле скорости называется также касательным векторным полем к кривой γ. □

Преобразования векторных полей (3.11) и 1-форм (3.12) различны, поэтому контравариантные и ковариантные индексы необходимо различать и они всегда будут писаться, соответственно, сверху и снизу.

Преобразования векторов (3.11) и 1-форм (3.12) линейны и однородны, причем матрица $\partial x^\alpha/\partial x^{\alpha'}$ в каждой точке многообразия невырождена и поэтому принадлежит группе невырожденных матриц $\mathbb{GL}(n,\mathbb{R})$. То есть каждому преобразованию координат соответствует $\mathbb{GL}(n,\mathbb{R})$ преобразование компонент векторного поля в касательном пространстве. Поскольку элементы группы $\mathbb{GL}(n,\mathbb{R})$ зависят от точки многообразия, то такие преобразования называются локальными. Обратное утверждение, вообще говоря, неверно. Не каждому локальному $\mathbb{GL}(n,\mathbb{R})$ преобразованию компонент векторного поля можно сопоставить некоторое преобразование координат. Это видно из подсчета параметров: локальное $\mathbb{GL}(n,\mathbb{R})$ преобразование параметризуется n^2 функциями по числу элементов $n\times n$ матрицы, в то время как преобразования координат параметризуются n функциями.

Из закона преобразования частных производных (3.9) и дифференциалов (3.8) следует, что суммы

$$X = X^\alpha \partial_\alpha = X^\alpha e_\alpha, \tag{3.13}$$

$$A = dx^\alpha A_\alpha = e^\alpha A_\alpha \tag{3.14}$$

инвариантны относительно преобразований координат. Эти формулы представляют собой разложения векторов и 1-форм по координатному базису, определенному в разделе 3.6.4.

Замечание. Скажем несколько слов об обозначениях. Там, где это возможно, мы будем записывать индексы суммирования по правилу "с десяти до четырех", т.е. сначала будем писать верхний индекс, а затем – нижний. Это правило является следствием записи *оператора внешнего дифференцирования* в виде

$$d = dx^\alpha \partial_\alpha. \tag{3.15}$$

В обратном порядке запись $d = \partial_\alpha dx^\alpha$ выглядит чрезвычайно неуклюже. В дифференциальной геометрии, когда все координаты являются вещественными числами, это правило не играет существенной роли, т.к. компоненты векторов и 1-форм можно менять местами. Однако, если часть координат антикоммутирует, то порядок индексов существенен, и необходимо придерживаться какого-либо фиксированного правила. Правило "с десяти до четырех" не является единственным, однако оно широко используется при построении моделей супергравитации в суперпространстве, где часть координат является антикоммутирующей. □

Дифференцируемое векторное поле, заданное в какой нибудь одной карте $\mathbb{U} \subset \mathbb{M}$, не всегда может быть продолжено до дифференцируемого векторного поля на всем многообразии.

Пример 3.6.5. Рассмотрим сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ единичного радиуса, вложенную в евклидово пространство (рис.**??**):

$$x^2 + y^2 + z^2 = 1, \qquad (x,y,z) \in \mathbb{R}^3.$$

Рассмотрим открытую верхнюю полусферу:

$$\mathbb{U} = \{(x, y, z) \in \mathbb{S}^2 : \ z > 0\}.$$

Зададим на $\mathbb{U}$ систему координат, спроектировав точки полусферы на плоскость x, y: $(x, y, z) \to (x, y)$. Таким образом мы построили карту $(\mathbb{U}, \varphi)$. Зададим в этой карте векторное поле $X = \partial_y$. В координатном базисе оно имеет компоненты $(0, 1)$ и поэтому является гладким векторным полем на $\mathbb{U}$. Точки сферы $(1, 0, 0)$ и $(-1, 0, 0)$ являются предельными для $\mathbb{U}$ и существенно особыми точками для векторного поля X, поскольку предел зависит от пути, по которому мы стремимся к данным точкам. Пока речь идет об открытом подмножестве $\mathbb{U}$ проблем не возникает, т.к. указанные точки не принадлежат $\mathbb{U}$. Однако любое продолжение векторного поля на окрестность, содержащую любую из точек $(1, 0, 0)$ и $(-1, 0, 0)$ приведет к векторному полю с особенностью. В данном случае особенность означает не обращение компонент векторного поля в бесконечность, а то, что в указанных точках векторное поле не будет однозначно определено. □

3.6.2 Глобальное определение векторных и ковекторных полей

Дадим глобальное определение векторных полей и 1-форм на многообразии $\mathbb{M}$, как это обычно делается в современных курсах дифференциальной геометрии. Рассмотрим дифференцируемую кривую $\gamma_1 : [-a, a] \ni t \to \mathbb{M}$, $a \in \mathbb{R}_+$, проходящую через некоторую точку $p \in \mathbb{M}$. Пусть $(\mathbb{U}, \varphi)$ – карта, содержащая точку $p \in \mathbb{U} \subset \mathbb{M}$. Тогда в координатах кривая задается набором функций:

$$\varphi \circ \gamma_1 = \{x_1^\alpha(t)\}, \quad p = \gamma_1(0).$$

Для определенности мы выбрали такую параметризацию кривой, что точке p соответствует значение $t = 0$. Рассмотрим другую дифференцируемую кривую γ_2, также проходящую через точку p,

$$\varphi \circ \gamma_2 = \{x_2^\alpha(t)\}, \quad p = \gamma_2(0).$$

Мы говорим, что две кривые касаются друг друга в точке p, если векторы скорости кривых в этой точке совпадают:

$$\dot{x}_1^\alpha|_{t=0} = \dot{x}_2^\alpha|_{t=0}, \qquad \alpha = 1, \ldots, n.$$

Поскольку векторы скорости при преобразованиях координат преобразуются одинаково, то данное определение не зависит от карты, покрывающей точку $p \in \mathbb{M}$. Касание кривых в точке p является отношением эквивалентности в классе всех кривых, проходящих через эту точку. Обозначим класс эквивалентности кривых, проходящих через точку p, который соответствует некоторому представителю γ, квадратными скобками $[\gamma_p]$. Каждый класс эквивалентности в координатах взаимно однозначно характеризуется набором чисел $\{\dot{x}^\alpha(0)\}$.

Определение. *Касательным вектором* $X_{(\gamma)}(p)$ к многообразию $\mathbb{M}$ в точке $p \in \mathbb{M}$ называется класс эквивалентности кривых $[\gamma_p]$, проходящих через эту точку. Множество всех касательных векторов в точке p называется *касательным пространством*

к многообразию в точке p и обозначается $\mathbb{T}_p(\mathbb{M})$. Объединение всех касательных пространств

$$\mathbb{T}(\mathbb{M}) = \bigcup_{p\in\mathbb{M}} \mathbb{T}_p(\mathbb{M}). \tag{3.16}$$

называется *касательным расслоением* с базой $\mathbb{M}$ и естественной проекцией $\pi : \mathbb{T}(\mathbb{M}) \to \mathbb{M}$, которая задана отображением $(p, [\gamma_p]) \to p$. Слоем касательного расслоения в точке p является касательное пространство $\pi^{-1}(p) = \mathbb{T}_p(\mathbb{M})$. *Векторным полем* на многообразии $\mathbb{M}$ называется сечение касательного расслоения $\mathbb{T}(\mathbb{M})$. □

То, что каждый слой является многообразием и диффеоморфен n-мерному векторному пространству, мы покажем чуть ниже.

Пусть f – функция на многообразии $\mathbb{M}$. Тогда каждому вектору $X_{(\gamma)}$ в точке p мы ставим в соответствие производную функции вдоль вектора

$$X_{(\gamma)}f|_p = \left.\frac{df\big(\gamma(t)\big)}{dt}\right|_{t=0} = X^\alpha(p)\partial_\alpha(f\circ\varphi^{-1})\Big|_{t=0}, \tag{3.17}$$

где $X^\alpha(p) = \dot{x}^\alpha(0)$ – компоненты вектора $X(p)$ в точке p в некоторой карте. Поскольку векторы скорости для всех кривых из одного класса эквивалентности совпадают, то вектор $X(p)$ в точке p взаимно однозначно характеризуется своими компонентами $\{X^\alpha(p)\}$. Отсюда следует, что векторное поле в произвольной карте $(\mathbb{U}, \varphi)$ взаимно однозначно задается своими компонентами $X(x) = \{X^\alpha(x)\}$, $x \in \mathbb{U}$. При этом компоненты векторного поля при переходе от одной системы координат к другой преобразуются по-правилу (3.11), что является следствием правила дифференцирования сложных функций. Таким образом, локальное определение векторного поля, данное ранее, согласуется с глобальным определением настоящего раздела.

В каждой точке многообразия $p \in \mathbb{M}$ множество векторов $\mathbb{T}_p(\mathbb{M})$ обладает естественной структурой векторного пространства.

Теорема 3.6.1. *Касательное пространство $\mathbb{T}_p(\mathbb{M})$ имеет естественную структуру вещественного векторного пространства той же размерности, что и само многообразие,* $\dim \mathbb{T}_p(\mathbb{M}) = \dim \mathbb{M} = n$.

Доказательство. Рассмотрим два вектора $X_1 = [\gamma_1] \in \mathbb{T}_p(\mathbb{M})$ и $X_2 = [\gamma_2] \in \mathbb{T}_p(\mathbb{M})$ в произвольной точке $p \in \mathbb{M}$. Пусть $(\mathbb{U}, \varphi)$ – координатная окрестность точки p. Пусть $\gamma_1 \in [\gamma_1]$ и $\gamma_2 \in [\gamma_2]$ – две произвольные кривые из классов эквивалентности $[\gamma_1]$ и $[\gamma_2]$. Тогда в евклидовом пространстве определены две кривые $\varphi\circ\gamma_1$ и $\varphi\circ\gamma_2$, которые задаются набором функций $\varphi\circ\gamma_1 = \{x_1^\alpha(t)\}$ и $\varphi\circ\gamma_2 = \{x_2^\alpha(t)\}$. Поскольку координаты являются вещественными числами, то их можно складывать и умножать. Определим сумму двух векторов и умножение на число $a \in \mathbb{R}$ как следующие классы эквивалентности

$$X_1 + X_2 = [\varphi^{-1}\circ(\varphi\circ\gamma_1 + \varphi\circ\gamma_2)],$$
$$aX = [\varphi^{-1}\circ(a\varphi\circ\gamma)].$$

Касательные векторы к кривым $\varphi\circ\gamma_1+\varphi\circ\gamma_2$ и $a\varphi\circ\gamma$ в точке p имеют, соответственно, компоненты: $\{\dot{x}_1^\alpha(0)+\dot{x}_2^\alpha(0)\}$ и $\{a\dot{x}^\alpha(0)\}$. Это определение суммы векторов и умножения на числа не зависит от выбора карты $(\mathbb{U}, \varphi)$ и представителя $\gamma \in [\gamma]$. Тем самым касательное пространство $\mathbb{T}_p(\mathbb{M})$ снабжается структурой векторного пространства. Поскольку каждый вектор $X(p)$ в карте взаимно однозначно задается набором n чисел $\{X^\alpha(p)\}$, то размерность касательного пространства совпадает с размерностью самого многообразия, $\dim \mathbb{T}_p(\mathbb{M}) = \dim \mathbb{M} = n$. □

Таким образом, мы установили, что касательное пространство $\mathbb{T}_p(\mathbb{M})$ в каждой точке многообразия имеет естественную структуру векторного пространства $\mathbb{R}^n$. Определим в этом векторном пространстве евклидову топологию, которая является единственной топологией, согласованной с линейной структурой. Тем самым касательное пространство $\mathbb{T}_p(\mathbb{M})$ так же является многообразием. В дальнейшем мы всегда предполагаем, что касательное пространство $\mathbb{T}_p(\mathbb{M})$ снабжено естественной структурой векторного пространства $\mathbb{R}^n$ и евклидовой топологией.

Типичным слоем касательного расслоения $\mathbb{T}(\mathbb{M}) \xrightarrow{\pi} \mathbb{M}$ является евклидово пространство $\mathbb{R}^n$, на котором введена структура векторного пространства. Это частный случай векторных расслоений, рассмотренных в разделе **??**.

Касательное расслоение является многообразием размерности $2n$. При этом, если на базе $\mathbb{M}$ задана дифференцируемая структура класса $\mathcal{C}^k$, $k \geq 1$, то касательное расслоение $\mathbb{T}(\mathbb{M})$ является дифференцируемым многообразием класса $\mathcal{C}^{k-1}$. Понижение класса дифференцируемости связано с тем, что, если функции преобразования координат принадлежат классу $\mathcal{C}^k$, то матрица Якоби преобразования координат, которая действует в касательном пространстве, принадлежит классу $\mathcal{C}^{k-1}$.

Пусть $\mathbb{M}$ – гладкое многообразие. Векторное поле $X(x)$ называется гладким, если в любой карте атласа $\{\mathbb{U}_i, \varphi_i\}$ компоненты $X^\alpha(x)$, задающие векторное поле, являются гладкими функциями. Множество всех гладких векторных полей на многообразии $\mathbb{M}$ обозначим $\mathcal{X}(\mathbb{M})$. Это множество, так же как и множество всех векторов в фиксированной точке $\mathbb{T}_p(\mathbb{M})$, обладает структурой вещественного линейного пространства с поточечным определением сложения и умножения на числа. Более того, вместо умножения на числа можно рассматривать умножение на гладкие функции $f \in \mathcal{C}^\infty(\mathbb{M})$. Легко проверить, что, если $X(x)$ – векторное поле, то $fX(x)$ так же является векторным полем. Таким образом, множество векторных полей $\mathcal{X}(\mathbb{M})$ является модулем над алгеброй гладких функций $\mathcal{C}^\infty(\mathbb{M})$. Как линейное пространство множество векторных полей является бесконечномерным.

Теперь нетрудно дать глобальное определение ковариантных векторных полей.

Определение. Множество линейных функционалов на касательном пространстве $\mathbb{T}_p(\mathbb{M})$ в точке $p \in \mathbb{M}$ называется *кокасательным векторным пространством* и обозначается $\mathbb{T}^*_p(\mathbb{M})$. Объединение всех кокасательных пространств

$$\mathbb{T}^*(\mathbb{M}) = \bigcup_{p\in\mathbb{M}} \mathbb{T}^*_p(\mathbb{M}). \tag{3.18}$$

называется *кокасательным расслоением* с базой $\mathbb{M}$ и естественной проекцией $\pi : \mathbb{T}^*(\mathbb{M}) \to \mathbb{M}$. Слоем кокасательного расслоения в точке p является кокасательное пространство $\pi^{-1}(p) = \mathbb{T}^*_p(\mathbb{M})$. *Кокасательным векторным полем* или 1-*формой* на многообразии $\mathbb{M}$ называется сечение кокасательного расслоения $\mathbb{T}^*(\mathbb{M})$ или линейное отображение множества векторных полей

$$A: \ \mathcal{X}(\mathbb{M}) \ni X \mapsto A(X) \in \mathcal{C}^\infty(\mathbb{M}). \qquad \square$$

В координатах ковекторное поле задается набором n функций $A_\alpha(x)$, которые при преобразовании координат преобразуются по правилу (3.12). Тогда линейное отображение задается простым суммированием компонент:

$$A(X) = X^\alpha A_\alpha.$$

Кокасательное векторное поле $A(x)$ называется гладким, если его компоненты $\{A_\alpha(x)\}$ являются гладкими функциями во всех картах. Кокасательное пространство $\mathbb{T}^*_p(\mathbb{M})$

в точке $p \in \mathbb{M}$ снабжается естественной структурой векторного пространства $\mathbb{R}^n$ и евклидовой топологией. Множество всех кокасательных векторных полей, которое обозначим $\Lambda_1(\mathbb{M})$, так же как и множество векторных полей $\mathcal{X}(\mathbb{M})$, образует модуль над алгеброй функций $\mathcal{C}^\infty(\mathbb{M})$.

Предложение 3.6.1. *Касательное пространство к прямому произведению двух многообразий $\mathbb{M} \times \mathbb{N}$ в точке $(p,q) \in \mathbb{M} \times \mathbb{N}$ естественно изоморфно прямой сумме касательных пространств:*

$$\mathbb{T}_{(p,q)}(\mathbb{M} \times \mathbb{N}) \ni \ X \simeq \Big(\pi_{1*}(X), \pi_{2*}(X)\Big) \ \in \mathbb{T}_p(\mathbb{M}) \oplus \mathbb{T}_q(\mathbb{N}),$$

где введены проекции на первый и второй сомножитель

$$\pi_1 : \ \mathbb{M} \times \mathbb{N} \ni (p,q) \mapsto p \in \mathbb{M},$$
$$\pi_2 : \ \mathbb{N} \times \mathbb{N} \ni (p,q) \mapsto q \in \mathbb{N},$$

и π_{1*}, π_{2*} – дифференциалы соответствующих отображений (см. следующий раздел).

Доказательство. Достаточно спроектировать кривую в прямом произведении $\mathbb{M}\times\mathbb{N}$ на каждый из сомножителей. □

3.6.3 Кокасательные векторные поля и ростки

Дадим также независимое определение кокасательного пространства 1-форм $\mathbb{T}^*_p(\mathbb{M})$ в точке $p \in \mathbb{M}$, без обращения к понятию дуального пространства как это было сделано в предыдущем разделе. Рассмотрим алгебру гладких функций $\mathcal{C}^\infty(\mathbb{M})$. Зафиксируем точку $p \in \mathbb{M}$. Будем считать две функции $f, g \in \mathcal{C}^\infty(\mathbb{M})$ эквивалентными $f \sim g$, если существует окрестность $\mathbb{U} \ni x$ такая, что $f|_{\mathbb{U}} = g|_{\mathbb{U}}$. Очевидно, что отношение $\sim$ является отношением эквивалентности в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$. Класс эквивалентности функции f обозначается $[f_p]$ и называется $\mathcal{C}^\infty$ *ростком* в точке $p \in \mathbb{M}$. Обозначим множество всех ростков в точке p через

$$\mathcal{C}_p = \mathcal{C}^\infty(\mathbb{M})/\sim \ = \{[f_p] : \quad f \in \mathcal{C}^\infty(\mathbb{M})\}. \tag{3.19}$$

Это множество естественным образом снабжается структурой векторного пространства, которая переносится из алгебры функций,

$$[f_p] + [g_p] = [f_p + g_p], \qquad a[f_p] = [af_p], \qquad a \in \mathbb{R}.$$

Таким образом, множество $\mathcal{C}^\infty$ ростков $\mathcal{C}_p$ в точке $p \in \mathbb{M}$ превращается в бесконечномерное вещественное векторное пространство.

Обозначим множество гладких кривых γ, которые задаются набором функций $\{x^\alpha(t)\}$, проходящих через точку $\gamma(0) = p$, символом Γ_p. Тогда производная функции вдоль кривой в точке p равна (3.17). Выражение $\partial_\alpha(f \circ \varphi^{-1})$ в правой части равенства зависит только от ростка $[f] \in \mathcal{C}_p$, но не от представителя $f \in [f_p]$. Поэтому будем писать

$$X_{(\gamma)}[f_p] : \quad \mathcal{C}_p \to \mathbb{R}.$$

Это отображение линейно по росткам:

$$X_{(\gamma)}([f_p] + [g_p]) = X_{(\gamma)}[f_p] + X_{(\gamma)}[g_p], \qquad X_{(\gamma)}(a[f_p]) = aX_{(\gamma)}[f_p], \qquad a \in \mathbb{R}.$$

Введем обозначение для тех ростков, производные которых вдоль всех кривых $\gamma \in \Gamma_p$, проходящих через точку p, равны нулю

$$\mathcal{H}_p = \{[f_p] \in \mathcal{C}_p : \quad X_{(\gamma)}[f_p] = 0\}.$$

Множество $\mathcal{H}_p$ является линейным подпространством в $\mathcal{C}_p$. В координатах принадлежность $[f_p] \in \mathcal{H}_p$ задается равенством $\partial_\alpha(f \circ \varphi^{-1} = 0)$. То есть подпространство $\mathcal{H}_p$ состоит в точности из тех ростков, у которых все частные производные равны нулю.

Определение. *Кокасательным пространством* в точке $p \in \mathbb{M}$ называется фактор пространство $\mathbb{T}^*_p(\mathbb{M}) = \mathcal{C}_p/\mathcal{H}_p$. □

Из данного определения следует, что кокасательное пространство состоит из тех ростков $[f_p]$, для которых хотя бы одна частная производная была отлична от нуля. Поскольку две функции принадлежат одному ростку $f \sim g$ тогда и только тогда, когда совпадают все их частные производные, то каждый росток взаимно однозначно определяется градиентом функции. Это доказывает эквивалентность независимого определения кокасательного пространства определению, данному ранее, т.к. градиент функции есть 1-форма в смысле прежнего определения. В основе данного определения, так же, как и в определении вектора, лежит понятие кривой на многообразии и вектора скорости.

В дальнейшем, в целях упрощения обозначений, мы почти всегда будем писать

$$\partial_\alpha f := \frac{\partial f}{\partial x^\alpha} = \frac{\partial (f \circ \varphi^{-1})}{\partial x^\alpha},$$

несмотря на то, что функция $f(x)$ определена на точках многообразия $x \in \mathbb{M}$, а не в евклидовом пространстве. Это является общепринятой записью и не приводит к путанице. В тех местах, где нужно подчеркнуть различие точки многообразия и точки евклидова пространства, мы будем использовать полную запись.

3.6.4 Векторные поля и дифференцирования

Дадим второе, теперь уже алгебраическое, глобальное определение векторных полей на многообразии $\mathbb{M}$. Зафиксируем произвольную точку $p \in \mathbb{M}$ и рассмотрим некоторую координатную окрестность этой точки $(\mathbb{U}, \varphi)$. С каждым векторным полем естественным образом связывается оператор дифференцирования (3.13) в алгебре гладких функций $\mathcal{C}^\infty(\mathbb{U})$. Его действие на функцию в произвольной карте

$$Xf = X^\alpha \partial_\alpha f \tag{3.20}$$

представляет собой *дифференцирование вдоль векторного поля*. Это дифференцирование не зависит от выбора карты, т.к. запись (3.20) инвариантна относительно преобразований координат, и удовлетворяет свойствам:

1) $X(af + bg) = aXf + bXg$ – линейность,
2) $X(fg) = (Xf)g + f(Xg)$ – правило Лейбница,

для всех $a, b \in \mathbb{R}$ и $f, g \in \mathcal{C}^\infty(\mathbb{U})$.

Определение. Непрерывное линейное отображение алгебры функций $\mathcal{C}^\infty(\mathbb{U})$ в $\mathbb{R}$:

$$X : \quad \mathcal{C}^\infty(\mathbb{U}) \ \to \ \mathbb{R},$$

удовлетворяющее правилу Лейбница 2), называется *дифференцированием* в точке $p \in \mathbb{M}$. Множество всех дифференцирований обозначим $\mathbb{D}_p(\mathbb{M})$. □

Замечание. Каждое векторное поле на $\mathbb{U} \subset \mathbb{M}$ отображает алгебру функций $\mathcal{C}^\infty(\mathbb{U})$ в себя. Обратим внимание, что в определении дифференцирования в точке стоит отображение не в алгебру функций $\mathcal{C}^\infty(\mathbb{U})$, а в вещественную прямую $\mathbb{R}$. Ясно, что если два векторных поля касаются друг друга в точке p, то они порождают одно и то же дифференцирование. □

Множество дифференцирований $\mathbb{D}_p(\mathbb{M})$ снабжается естественной структурой вещественного векторного пространства:

$$\begin{aligned}(X_1 + X_2)f &= X_1 f + X_2 f,\\ (aX)f &= a(Xf),\end{aligned} \tag{3.21}$$

где $X_1, X_2, X \in \mathcal{X}(\mathbb{M})$ и $a \in \mathbb{R}$.

Теорема 3.6.2. *Касательное пространство $\mathbb{T}_p(\mathbb{M})$ и пространство дифференцирований $\mathbb{D}_p(\mathbb{M})$ изоморфны как векторные пространства.*

Доказательство. То, что каждому вектору $X \in \mathbb{T}_p(\mathbb{M})$ однозначно ставится в соответствие дифференцирование было показано выше. Нетрудно проверить, что это отображение сохраняет линейную структуру.

Докажем обратное утверждение. С этой целью рассмотрим важный пример отображений $\mathcal{C}^\infty(\mathbb{U}) \to \mathbb{R}$:

$$\left(\frac{\partial}{\partial x^\alpha}\right)_p f(x) := \partial_\alpha (f \circ \varphi^{-1})|_{\varphi(p)}, \qquad \alpha = 1, \ldots, n, \tag{3.22}$$

Эти отображения линейны, удовлетворяют правилу Лейбница и, следовательно, являются дифференцированиями. Подчеркнем, что символ $\left(\frac{\partial}{\partial x^\alpha}\right)_p$ не является частной производной, т.к. определен в точке многообразия $p \in \mathbb{M}$, а не евклидова пространства $\mathbb{R}^n$. В правой же части равенства (3.22) стоит частная производная от функции, определенной в евклидовом пространстве. Теперь докажем два утверждения.

Лемма 3.6.1. *Пусть X – дифференцирование в $\mathcal{C}^\infty(\mathbb{U})$ и $f_c(x) = c =$ const – постоянная функция на $\mathbb{M}$. Тогда $Xc = 0$.*

Доказательство. Представим постоянную функцию на $\mathbb{M}$ в виде $f_c = cf_1 = cf_1 \cdot f_1$, где $f_1(x) = 1$ – функция, равная единице на всем многообразии $\mathbb{M}$. Используем линейность дифференцирования и правило Лейбница,

$$Xc = cX(1 \cdot 1) = cX(1) + cX(1) = 2cX(1) = 2Xc,$$

что возможно только при $Xc = 0$. Это – нетривиальное использование, казалось бы, тривиального тождества: $1 \cdot 1 = 1$. □

Лемма 3.6.2. *Для любой гладкой функции $f \in \mathbb{C}^\infty(\mathbb{U})$ существует такой набор функций $f_\alpha \in \mathcal{C}^\infty(\mathbb{U})$, $\alpha = 1, \ldots, n$, что для любой точки x в некоторой окрестности точки $p \in \mathbb{U}$ выполнены равенства:*

$$f_\alpha(p) = \left(\frac{\partial}{\partial x^\alpha}\right)_p f(p), \tag{3.23}$$

$$f(x) = f(p) + x^\alpha(x) f_\alpha(x), \tag{3.24}$$

где $\{x^\alpha(x)\} : \mathbb{U} \to \mathbb{R}^n$ – координатные функции.

Доказательство. Для определенности будем считать, что образ точки $p \in \mathbb{U}$ совпадает с началом координат евклидова пространства, $\varphi(p) = (0, \dots, 0) \in \mathbb{R}^n$. Пусть $F(x^1, \dots, x^n) = f \circ \varphi^{-1}$ – координатное представление функции f в некоторой окрестности точки p. Тогда справедливо тождество

$$\begin{aligned} F(x^1, \dots, x^n) = {} & F(x^1, \dots, x^n) - F(x^1, \dots, x^{n-1}, 0) + \\ & + F(x^1, \dots, x^{n-1}, 0) - F(x^1, \dots, x^{n-2}, 0, 0) + \\ & + \dots \\ & + F(x^1, 0, \dots, 0) - F(0, \dots, 0) + \\ & + F(0, \dots, 0). \end{aligned}$$

Это тождество перепишем в виде

$$\begin{aligned} F(x^1, \dots, x^n) = {} & F(0, \dots, 0) + \sum_{\alpha=1}^{n} F(x^1, \dots, tx^\alpha, 0, \dots, 0)|_{t=0}^{t=1} = \\ = {} & F(0, \dots, 0) + \sum_{\alpha=1}^{n} \int_0^1 dt \frac{\partial F}{\partial (tx^\alpha)}(x^1, \dots, x^{\alpha-1}, tx^\alpha, 0, \dots, 0) x^\alpha = \\ = {} & F(0, \dots, 0) + x^\alpha F_\alpha(x^1, \dots, x^n), \end{aligned}$$

где

$$F_\alpha(x^1, \dots, x^n) = \int_0^1 dt \frac{\partial F}{\partial (tx^\alpha)}(x^1, \dots, x^{\alpha-1}, tx^\alpha, 0, \dots, 0)$$

– набор гладких функций в некоторой окрестности начала координат евклидова пространства. Теперь вернемся на многообразие и определим набор функций $f_\alpha = F \circ \varphi$. Отсюда сразу следует равенство (3.24).

Теперь надо определить вид функций $f_\alpha(x)$. С этой целью применим дифференцирование $\left(\frac{\partial}{\partial x^\alpha}\right)_x$ к равенству (3.24)

$$\left(\frac{\partial}{\partial x^\alpha}\right)_x f(x) = \left(\frac{\partial}{\partial x^\alpha}\right)_x f(p) + \left(\frac{\partial}{\partial x^\alpha}\right)_x x^\beta(x) f_\beta(x) + x^\beta(x) \left(\frac{\partial}{\partial x^\alpha}\right)_x f_\beta(x).$$

Первое слагаемое равно нулю, как следствие леммы 3.6.1. Поскольку $\left(\frac{\partial}{\partial x^\alpha}\right)_x x^\beta(x) = \delta_\alpha^\beta$, то в точке p имеем равенство (3.23) т.к. $x^\alpha(p) = 0$. □

Следствие. Если $X \in \mathbb{D}_p(\mathbb{U})$ – дифференцирование в точке $p \in \mathbb{M}$, то

$$X(p) = X\Big(x^\alpha(x)\Big)_p \left(\frac{\partial}{\partial x^\alpha}\right)_p = X^\alpha(p) \left(\frac{\partial}{\partial x^\alpha}\right)_p. \tag{3.25}$$

Доказательство. Пусть $(\mathbb{U}, \varphi)$ – произвольная карта в окрестности точки $p \in \mathbb{M}$. Тогда, возможно, $x^\alpha(p) \neq 0$. В этом случае сдвинем начало координат в евклидовом пространстве: $y^\alpha = x^\alpha - x^\alpha(p)$. Тогда из леммы 3.6.2 следует представление

$$f(x) = f(p) + \Big(x^\alpha(x) - x^\alpha(p)\Big) f_\alpha(x).$$

Применяя дифференцирование X к этому равенству и переходя в точку p, получим (3.25). □

Таким образом, множество всех дифференцирований $\mathbb{D}_p(\mathbb{U})$ в произвольной точке $p \in \mathbb{M}$ представляет собой конечномерное векторное пространство, $\dim \mathbb{D}_p(\mathbb{U}) = n$, с базисом $\left(\frac{\partial}{\partial x^\alpha}\right)_p$. Это пространство изоморфно касательному пространству $\mathbb{T}_p(\mathbb{M})$, и теорема доказана. □

Определение. *Координатным базисом* векторных полей $\mathcal{X}(\mathbb{U})$ на карте $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ называется набор гладких векторных полей

$$\left\{ e_\alpha(x) := \left(\frac{\partial}{\partial x^\alpha}\right)_x \right\} \in \mathcal{X}(\mathbb{U}), \qquad \alpha = 1, \ldots, n,$$

определенных формулой (3.22). Дуальный базис $\{e^\alpha(x)\} \in \Lambda_1(\mathbb{U})$, $e^\alpha(e_\beta) = \delta^\alpha_\beta$, называется *координатным базисом* ковекторных полей (1-форм). Координатный базис для ковекторных полей обозначается dx^α_p или просто dx^α. Координатные базисы называют также *голономными*. □

Замечание. Подчеркнем, что координатный базис – это не набор частных производных, а векторные поля на многообразии. Их действие, как дифференцирований, определено только для достаточно гладких функций $\mathcal{C}^k(\mathbb{U})$. Действие векторных полей $e_\alpha(x)$ на тензоры более высокого ранга не определено. □

Замечание. Обозначение координатных базисов векторных и ковекторных полей через ∂_α и dx^α оправдано простой формулой из математического анализа

$$\frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta.$$

Для тривиальных многообразий $\mathbb{M} \approx \mathbb{R}^n$, покрытых одной картой, векторы e_α можно отождествить с операторами частных производных ∂_α, которые действуют на дифференцируемые функции. Тогда дуальный базис e^α ковекторных полей естественным образом отождествляется с дифференциалами координатных функций dx^α. □

Из разложения по базису (3.25) следует, что компонента векторного поля X в точке $p \in \mathbb{M}$ – это результат действия векторного поля на координатную функцию $X^\alpha(p) = Xx^\alpha(x)|_{x=p}$.

В дальнейшем мы будем писать сокращенно $X = X^\alpha(x)\partial_\alpha$, имея в виду, что на функции $f \circ \varphi$ в евклидовом пространстве координатный базис действительно действует, как частная производная. Для 1-форм в координатном базисе мы часто будем использовать общепринятую запись $A = dx^\alpha A_\alpha$, $e^\alpha = dx^\alpha$.

Пусть $D_x(\mathbb{U}) \in \mathbb{D}_x(\mathbb{M})$ – некоторое дифференцирование в точке $x \in \mathbb{M}$. Рассмотрим объединение $D(\mathbb{M}) = \bigcup_{x\in\mathbb{M}} D_x(\mathbb{U})$ по всем точкам многообразия, которое соответствует некоторому векторному полю $X \in \mathcal{X}(\mathbb{M})$. Оно задает отображение

$$D(\mathbb{M}) : \ \mathcal{C}^\infty(\mathbb{M}) \to \mathcal{C}^\infty(\mathbb{M}), \tag{3.26}$$

которое называется дифференцированием в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$, т.е. непрерывный линейный эндоморфизм в $\mathcal{C}^\infty(\mathbb{M})$, удовлетворяющий правилу Лейбница. Поэтому каждому векторному полю $X \in \mathcal{X}(\mathbb{M})$ ставится в соответствие некоторое дифференцирование $D(\mathbb{M}) \in \mathcal{D}(\mathbb{M})$, где $\mathcal{D}(\mathbb{M})$ – множество всех дифференцирований а алгебре функций. Верно также и обратное утверждение: любому дифференцированию в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$ соответствует единственное векторное поле. Мы доказали аналогичное утверждение в фиксированной точке многообразия, где пространство

$\mathbb{D}_x(\mathbb{U})$ является конечномерным. Доказательство в рассматриваемом бесконечномерном случае сложнее и приведено в [?]. Линейная структура на $\mathcal{D}(\mathbb{M})$ вводится также, как и в точке (3.21). Таким образом множество векторных полей $\mathcal{X}(\mathbb{M})$ биективно отображается на множество дифференцирований $\mathcal{D}(\mathbb{M})$, при этом как векторные пространства эти множества изоморфны. Эта биекция позволяет дать эквивалентное определение векторного поля.

Определение. *Векторным полем* $X \in \mathcal{X}(\mathbb{M})$ на многообразии $\mathbb{M}$ называется дифференцирование в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$. □

3.6.5 Векторные поля и интегральные кривые

Начнем с локального описания. Рассмотрим векторное поле $X = X^\alpha \partial_\alpha$, все компоненты которого отличны от нуля в окрестности некоторой точки $x \in \mathbb{M}$, $\dim \mathbb{M} = n$. Бесконечно малые перемещения dx^α точки x вдоль этого векторного поля должны быть пропорциональны компонентам X^α и поэтому удовлетворять системе уравнений

$$\frac{dx^1}{X^1} = \frac{dx^2}{X^2} = \cdots = \frac{dx^n}{X^n}. \tag{3.27}$$

Эту систему уравнений можно переписать в виде равенства нулю 1-форм:

$$A^{\text{м}} = \frac{dx^{\text{м}}}{X^{\text{м}}} - \frac{dx^1}{X^1} = 0, \qquad \text{м} = 2, \ldots, n.$$

В каждой точке $x \in \mathbb{M}$ векторное поле X задает одномерное подпространство в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, а совокупность 1-форм $\{A^{\text{м}}\}$ – $(n-1)$-мерное ортогональное дополнение в сопряженном пространстве $\mathbb{T}^*_x(\mathbb{M})$, поскольку $A^{\text{м}}(X) = 0$.

Согласно теории дифференциальных уравнений, уравнения (3.27) допускают $n-1$ функционально независимых решений

$$\varphi^{\text{м}}(x^\alpha) = c^{\text{м}} = \mathsf{const}, \qquad \text{м} = 2, \ldots, n, \tag{3.28}$$

для которых $A^{\text{м}} = d\varphi^{\text{м}} = 0$. При этом прямоугольная $n \times (n-1)$-матрица, составленная из производных $\partial_\alpha \varphi^{\text{м}}$, имеет ранг $n-1$, и каждая из функций $\varphi^{\text{м}}$ является решением уравнения в частных производных

$$X^\alpha \partial_\alpha \varphi^{\text{м}} = 0. \tag{3.29}$$

Совершим преобразование координат $x^\alpha \to y^\alpha(x)$, выбрав в качестве последних $n-1$ координат функции (3.28) $(y^2, \ldots, y^n) = (\varphi^2, \ldots, \varphi^n)$, а координату x^1 оставим без изменения. Якобиан этого преобразования отличен от нуля в силу функциональной независимости функций $\varphi^{\text{м}}$. Тогда из закона преобразования векторного поля (3.11) и уравнения (3.29) следует, что в новой системе координат все компоненты векторного поля, кроме первой, равны нулю:

$$X^1 \neq 0, \quad X^2 = \ldots = X^n = 0.$$

Заменим теперь координату x^1 на функцию $y^1(x)$, которая удовлетворяет дифференциальному уравнению

$$\frac{dy^1}{dx^1} X^1 = 1.$$

Это уравнение локально разрешимо, и, значит, в новой системе координат $X^1 = 1$.

Если у векторного поля X^α часть компонент равнялась нулю до преобразования координат, то все, сказанное выше, можно повторить для ненулевых компонент. Отсюда следует частный случай теоремы Фробениуса, которая будет сформулирована в разделе 3.11.

Теорема 3.6.3. *Для произвольного нетривиального векторного поля $X \in \mathcal{X}(\mathbb{M})$ в некоторой окрестности $\mathbb{U}_x$ произвольной точки $x \in \mathbb{M}$ существует такая система координат, в которой все компоненты X, кроме одной (например, первой) обращаются в нуль. Координатную функцию $x^1(x)$, $x \in \mathbb{U}_x$ можно подобрать таким образом, чтобы первая компонента векторного поля X^1 была равна единице в этой окрестности, т.е. $X = \partial_1$.*

Замечание. Эта теорема показывает, что векторное поле в окрестности точки, в которой оно отлично от нуля, устроено довольно просто. Если в некоторой точке векторное поле обращается в нуль, то в окрестности этой точки оно может быть устроено очень сложно [**?**]. Нули гладкого касательного поля к многообразию связаны с топологическими свойствами. Например, на четномерной сфере не существует векторного поля, нигде не обращающегося в нуль (см. теорему 11.2.1). В то же время такое поле всегда можно задать на торе. □

Если в уравнения (3.28) подставить координаты некоторой фиксированной точки $p = \{x^\alpha_p\} \in \mathbb{M}$, то определятся значения постоянных $c^{\text{м}}$. При этих значениях постоянных система $n-1$ трансцендентных уравнений относительно x^α (3.28) определяет кривую, проходящую через точку p. Если кривая γ параметризуется параметром t, $\gamma = \{x^\alpha(t)\}$, то уравнения (3.27) эквивалентны системе обыкновенных дифференциальных уравнений

$$\frac{dx^\alpha}{dt} = X^\alpha(x), \tag{3.30}$$

с начальными условиями

$$x^\alpha|_{t=t_p} = x^\alpha_p, \tag{3.31}$$

где t_p – значение параметра вдоль кривой, при котором она проходит через точку p.

Из теории дифференциальных уравнений хорошо известна

Теорема 3.6.4. *Если векторное поле X на многообразии $\mathbb{M}$ дифференцируемо, то через каждую точку $p \in \mathbb{M}$ проходит одна и только одна интегральная кривая этого векторного поля.*

На языке теории дифференциальных уравнений это утверждение означает, что решение задачи Коши (3.30), (3.31) существует и единственно.

Общее решение системы уравнений (3.30) зависит от n постоянных интегрирования. Одна из постоянных соответствует сдвигу параметра вдоль кривой $t \mapsto t + \mathsf{const}$, а оставшиеся $n-1$ постоянных определяются положением точки p на гиперповерхности (3.28), проходящей через точку p.

Если решение системы уравнений (3.30) представимо в виде ряда, то вблизи точки p оно выглядит очень просто

$$x^\alpha(t) = x^\alpha_p + X^\alpha_p(t - t_p) + \ldots, \qquad (t - t_p) \ll 1.$$

То есть компоненты векторного поля определяют главную линейную часть интегральной кривой.

Замечание. Интегральные кривые векторного поля являются ни чем иным, как *характеристиками* (линиями уровня) для решений дифференциального уравнения в частных производных первого порядка (3.29). □

Условие дифференцируемости векторного поля в теореме 3.6.4 можно ослабить.

Определение. Функция (отображение)

$$f: \quad \mathbb{R}^n \supset \mathbb{U} \ni x \mapsto f(x) \in \mathbb{R}^m$$

удовлетворяет *условию Липшица*, если существует такая положительная постоянная $\mathbb{C}$, что

$$|f(x_2) - f(x_1)| \leq C|x_2 - x_1|,$$

для всех $x_{1,2} \in \mathbb{U}$. □

В условии Липшица знак модуля обозначает обычный модуль вектора в евклидовом пространстве.

Теорема 3.6.5. *Пусть правая часть системы уравнений (3.30) непрерывна и удовлетворяет условию Липшица в области $\mathbb{U}$. Тогда через каждую точку $\mathbb{U}$ проходит одна и только одна интегральная кривая системы уравнений (3.30).*

Если правая часть системы уравнений (3.30) только непрерывна, то и тогда через каждую точку $p \in \mathbb{U}$ проходит хотя бы одна интегральная кривая. Однако единственность может быть нарушена.

Пример 3.6.6. Рассмотрим обыкновенное дифференциальное уравнение $\frac{2}{3}\dot{x} = x^{1/3}$. Его решения имеют вид $x = 0$ и

$$x = \pm(x + c)^{3/2}.$$

Отсюда следует, что через каждую точку $x \neq 0$ проходят две интегральные кривые. Через точку $x = 0$ проходит даже три интегральные кривые: $x = 0$, $x = \pm t^{3/2}$. Если положить в условии Липшица $x_1 = 0$, то оно примет вид

$$|x_2|^{1/3} \leq C|x_2|.$$

Ясно, что такой постоянной C не существует, т.к. правая часть неравенства при $x_2 \to 0$ стремится к нулю быстрее. □

В дальнейшем мы всегда предполагаем, что векторное поле по крайней мере дифференцируемо.

Предложение 3.6.2. *Если дифференцируемое векторное поле X обращается в нуль в некоторой точке многообразия $p \in \mathbb{M}$, то эта точка является неподвижной относительно потока векторного поля, т.е. $x(t) = p$ для всех $t \in \mathbb{R}$. Обратно. Если точка $p \in \mathbb{M}$ является неподвижной на интегральной кривой некоторого векторного поля, то в этой точке векторное поле обращается в нуль, $X(p) = 0$.*

Доказательство. Постоянные функции $x^\alpha(t) = x^\alpha_p$ удовлетворяют системе уравнений (3.30), если $X^\alpha_p = 0$. Обратное утверждение очевидно. □

Если векторное поле умножить на произвольную достаточно гладкую отличную от нуля функцию $X \mapsto fX$, то уравнение для интегральной кривой $x(\tau)$ примет вид

$$\frac{dx^\alpha}{d\tau} = fX^\alpha.$$

Введем новый параметр $t(\tau)$ вдоль кривой, который является решением уравнения

$$\frac{dt}{d\tau} = f.$$

Решение этого уравнения существует и является монотонным, т.к. $f \neq 0$. В новой параметризации уравнение для интегральной кривой принимает прежний вид (3.30). Таким образом, два векторных поля X и fX, отличающиеся умножением на отличную от нуля функцию, имеют интегральные кривые, которые совпадают, как подмножества в $\mathbb{M}$. Отличие сводится только к различным параметризациям кривых.

Перейдем к глобальному описанию.

Определение. *Интегральной кривой* векторного поля $X \in \mathcal{X}(\mathbb{M})$, проходящей через точку $p \in \mathbb{M}$ называется кривая $\gamma: \ (a,b) \ni t \mapsto x(t) \in \mathbb{M}$ такая, что $\gamma(t_p) = p$ и

$$\gamma_* \partial_t = X\Big(x(t)\Big), \tag{3.32}$$

где γ_* – дифференциал отображения некоторого открытого интервала $(a,b) \subset \mathbb{R}$, содержащего точку t_p, и ∂_t – касательный вектор к интервалу в точке $t \in (a,b)$. Вектор $X\Big(x(t)\Big)$ называется *касательным вектором* к кривой γ в точке $x(t)$ (вектором скорости). □

Нетрудно проверить, что в каждой карте уравнение (3.32) записывается в виде системы обыкновенных дифференциальных уравнений (3.30).

Вообще говоря, интегральные кривые существуют только локально, даже для гладких векторных полей, заданных на всем многообразии. Другими словами, параметр t в общем случае определен лишь на некотором конечном или полубесконечном интервале $(a,b) \subset \mathbb{R}$.

Определение. Векторное поле $X \in \mathcal{X}(\mathbb{M})$ называется *полным*, если все интегральные кривые этого поля определены при всех значениях $t \in \mathbb{R}$. □

Теорема 3.6.6. *На компактном многообразии* $\mathbb{M}$ *любое гладкое векторное поле* $X \in \mathcal{X}(\mathbb{M})$*, не обращающееся в нуль, является полным.*

Доказательство. См., например, [**?**]. □

Замечание. Для многообразий понятие компакта и компактного пространства совпадают, т.к. многообразие, по-определению, является хаусдорфовым пространством.

На некомпактном многообразии векторное поле может быть либо полным, либо неполным.

Пример 3.6.7. Векторному полю ∂_x на вещественной прямой $\mathbb{R}$ соответствуют интегральные кривые $x = t + \mathsf{const}$. Они полны на всей прямой $\mathbb{R}$. Однако они неполны на полупрямой $\mathbb{R}_+$ или любом конечном открытом интервале $(a,b) \subset \mathbb{R}$. □

Пример 3.6.8. Рассмотрим гладкое векторное поле $X = x^2\partial_x$ на вещественное прямой $\mathbb{R}$. Общее решение уравнения интегральной кривой $\dot{x} = x^2$ имеет вид

$$x = -\frac{1}{t-c}, \qquad c = \mathsf{const}.$$

Таким образом, для каждого значения постоянной c, имеются две никак не связанные между собой интегральные кривые:

$$\begin{aligned} \gamma_1 :&\quad (-\infty, c) \;\to\; (0, \infty), \\ \gamma_2 :&\quad (c, \infty) \;\to\; (-\infty, 0). \end{aligned}$$

При этом точка $x = 0$ соответствует бесконечному значению параметра вдоль интегральной кривой $t = \pm\infty$, а бесконечно удаленные точки $x = \pm\infty$ – конечному значению параметра $t = c$. Поэтому векторное поле $X = x^2\partial_x$ неполно. □

Определение. Точка $p \in \mathbb{M}$, в которой векторное поле обращается в нуль, $X(p) = 0$, называется *особой точкой векторного поля*. Система уравнений (3.30) в координатной окрестности особой точки в линейном приближении имеет вид

$$\frac{dx^\alpha}{dt} = x^\beta B_\beta{}^\alpha, \qquad B_\beta{}^\alpha := \partial_\beta X^\alpha\big|_{x=p}. \tag{3.33}$$

Если $\mathsf{det}\, B \neq 0$, то особая точка называется *невырожденной*. □

При преобразовании координат $x^\alpha \mapsto x^{\alpha'}(x)$ уравнение (3.33) сохраняет свой вид. При этом компоненты матрицы B преобразуются по-правилу

$$B_\beta{}^\alpha \mapsto B_{\beta'}{}^{\alpha'} = \frac{\partial x^\beta}{\partial x^{\beta'}} B_\beta{}^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha},$$

т.е. подвергаются преобразованию подобия.

Напомним, что любое дифференциальное уравнение n-того порядка можно записать в виде эквивалентной ей системы уравнений первого порядка, состоящей из n уравнений (см. раздел **??**). Следовательно, теория обыкновенных дифференциальных уравнений произвольного порядка, разрешенных относительно старшей производной, сводится к нахождению и исследованию свойств интегральных кривых векторных полей. При этом поведение интегральных кривых в окрестностях особых точек представляет исключительный интерес, т.к. позволяет понять качественное поведение решений.

В общем случае поведение интегральных кривых в окрестности особой точки довольно сложно. Особые точки можно классифицировать, приведя матрицу B к какому либо каноническому виду с помощью преобразования подобия, что означает переход в новую систему координат.

Пример 3.6.9. Рассмотрим двумерное многообразие (поверхность) $\mathbb{M}$, на котором задано дифференцируемое векторное поле $X \in \mathcal{X}(\mathbb{M})$. Пусть $p \in \mathbb{M}$ – невырожденная особая точка векторного поля X. Выберем систему координат в окрестности особой точки так, чтобы она находилась в начале координат. Обозначим собственные числа матрицы B через λ_1 и λ_2. В общем случае они комплексны. Из невырожденности следует, что $\lambda_1\lambda_2 \neq 0$. Невырожденные особые точки в рассматриваемом случае делятся на шесть классов.

a). Седло. $\lambda_1\lambda_2 < 0$. Пусть собственные числа вещественны и разных знаков. Тогда существует система координат в которой матрица B диагональна и уравнения для интегральных кривых (3.33) примут вид

$$\dot{x}^1 = x^1\lambda_1, \qquad \dot{x}^2 = x^2\lambda_2. \tag{3.34}$$

Они легко интегрируются:

$$x^1 = C_1\,\mathsf{e}^{\lambda_1 t}, \qquad x^2 = C_2\,\mathsf{e}^{\lambda_2 t}, \qquad t \in \mathbb{R}, \tag{3.35}$$

где $C_{1,2}$ – постоянные интегрирования. Касательная к интегральной кривой имеет угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2\lambda_2}{C_1\lambda_1}\,\mathsf{e}^{(\lambda_2-\lambda_1)t}.$$

Соответствующие интегральные кривые для $\lambda_1 < 0 < \lambda_2$ показаны на рис.3.6,*a*. Стрелки указывают направление возрастания параметра t. Ни одна из интегральных кривых не проходит через начало координат, которое является неподвижной точкой.

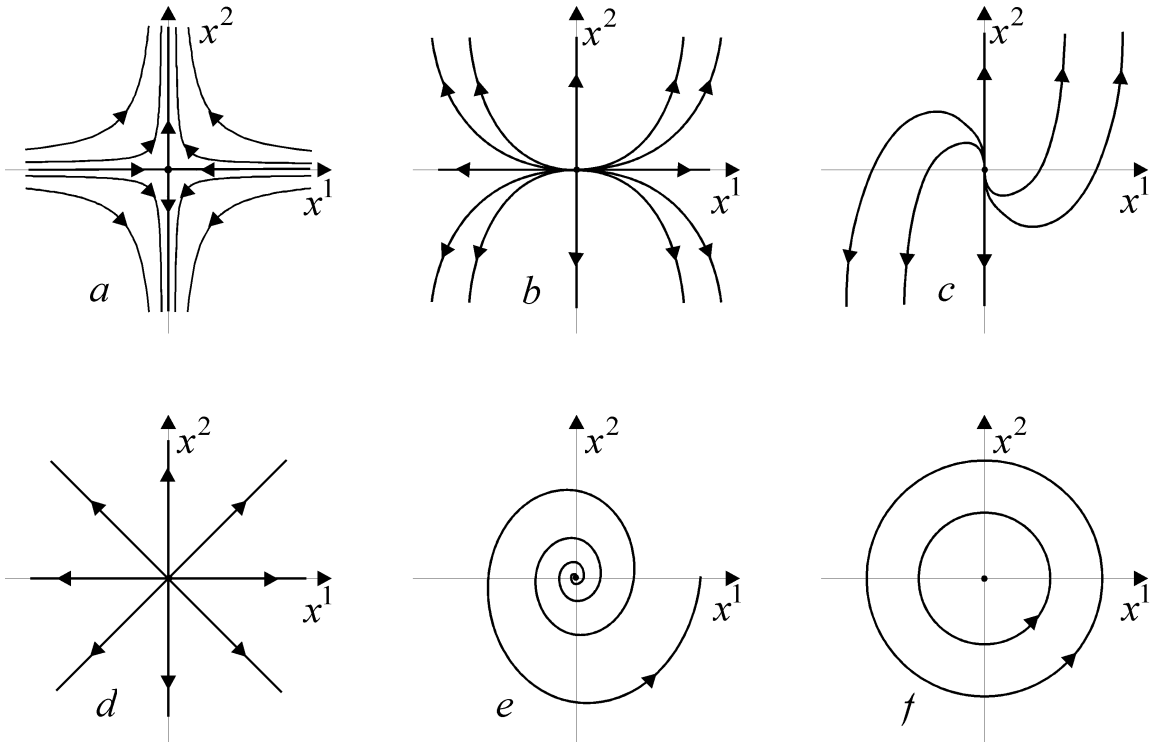


Рис. 3.6: Седло (a), узел (b), жорданов узел (c), дикритический узел (d), фокус (e), центр (f).

b). Узел. $\lambda_1\lambda_2 > 0$, $\lambda_1 \neq \lambda_2$. Если собственные числа матрицы B вещественны одного знака и различны, то ее также можно диагонализировать. При этом уравнения для интегральных кривых и решений имеют прежний вид (3.34), (3.35). Меняется только знак одного из собственных чисел. На рис.3.6,*b* показаны интегральные кривые для $0 < \lambda_1 < \lambda_2$.

c). Жорданов узел. $\lambda_1 = \lambda_2$, матрица B недиагонализируема. Пусть собственные числа матрицы B вещественны и равны, и матрицу B нельзя диагонализировать

преобразованием подобия. Тогда матрицу B можно привести к жордановой клетке (см. дополнение 28.1)

$$B = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

Соответствующие уравнения интегральных кривых

$$\dot{x}^1 = x^1\lambda_1, \quad \dot{x}^2 = x^1 + x^2\lambda_1,$$

легко интегрируются:

$$x^1 = C_1\,\mathsf{e}^{\lambda_1 t}, \quad x^2 = (C_2 + C_1 t)\,\mathsf{e}^{\lambda_1 t}, \quad t \in \mathbb{R},$$

Касательная к интегральной кривой имеет угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2}{C_1} + \frac{1}{\lambda_1} + t.$$

На рис.3.6,*c* показаны интегральные кривые для $0 < \lambda_1$.

d). Дикритический узел. $\lambda_1 = \lambda_2$, матрица B кратна единичной. Пусть собственные числа матрицы B вещественны и равны, и матрица B кратна единичной,

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}.$$

При преобразовании координат матрица B не меняется, т.к. пропорциональна единичной матрице. В этом случае уравнения для интегральных кривых,

$$\dot{x}^1 = x^1\lambda_1, \quad \dot{x}^2 = x^2\lambda_1,$$

легко интегрируются:

$$x^1 = C_1\,\mathsf{e}^{\lambda_1 t}, \quad x^2 = C_2\,\mathsf{e}^{\lambda_1 t}, \quad t \in \mathbb{R},$$

Касательная к интегральной кривой имеет постоянный угол наклона

$$\frac{dx^2}{dx^1} = \frac{C_2}{C_1}.$$

На рис.3.6,*d* показаны интегральные кривые для $0 < \lambda_1$.

e). Фокус. $\lambda_{1,2}$ комплексны, $\mathsf{re}\,\lambda_{1,2} \neq 0$. Поскольку комплексные собственные значения могут встречаться только комплексно сопряженными парами, то $\lambda_1 = \mu + i\tilde{\omega}$ и $\lambda_2 = \mu - i\tilde{\omega}$, где $\mu, \tilde{\omega} \in \mathbb{R}$ и $\mu \neq 0$. В этом случае матрицу B можно привести к виду

$$B = \begin{pmatrix} 0 & -\omega^2 \\ 1 & 2\mu \end{pmatrix},$$

где $\omega^2 := \tilde{\omega}^2 + \mu^2$. Соответствующие уравнения для интегральных кривых принимают вид

$$\begin{aligned} \dot{x}^1 &= x^2, \\ \dot{x}^2 &= -\omega^2 x^1 + 2\mu x^2. \end{aligned}$$

Дифференцирование первого уравнения приводит к уравнению для осциллятора с трением

$$\ddot{x}^1 - 2\mu\dot{x}^1 + \omega^2 x^1 = 0.$$

Отсюда следует, что интегральные кривые имеют вид

$$\begin{aligned} x^1 &= C_1\, \mathsf{e}^{\mu t} \cos(\tilde{\omega}t + C_2), \\ x^2 &= C_1\, \mathsf{e}^{\mu t} \left[\mu \cos(\tilde{\omega}t + C_2) - \tilde{\omega} \sin(\tilde{\omega}t + C_2)\right]. \end{aligned}$$

На рис.3.6,*д* показаны интегральные кривые для $\tilde{\omega} > 0$, $\mu > 0$.

f). Центр. $\lambda_{1,2}$ комплексны, $\mathsf{re}\,\lambda_{1,2} = 0$. Пусть собственные числа чисто мнимые: $\lambda_1 = -\lambda_2 = i\omega$, $\omega \in \mathbb{R}$. В этом случае матрицу B можно привести к виду

$$B = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

Уравнения для интегральных кривых,

$$\dot{x}^1 = \omega x^2, \qquad \dot{x}^2 = -\omega x^1,$$

соответствуют гармоническому осциллятору

$$\ddot{x}^1 + \omega^2 x^1 = 0$$

и легко интегрируются:

$$x^1 = C_1 \cos(\omega t + C_2), \qquad x^2 = -C_1 \sin(\omega t + C_2), \qquad t \in \mathbb{R}.$$

Интегральные кривые являются окружностями,

$$(x^1)^2 + (x^2)^2 = C_1^2.$$

Они изображены на рис.3.6,*f* при $\omega > 0$.

Во всех случаях интегральные кривые определены при всех значениях параметра t. Это означает, что векторные поля полны. Для узлов и фокуса интегральные кривые стремятся к началу координат при $t \to -\infty$. Начало координат в соответствии с предложением 3.6.2 является неподвижной точкой. □

Множество всех полных векторных полей на некомпактном многообразии $\mathbb{M}$ является подмножеством в $\mathcal{X}(\mathbb{M})$. В отличие от всего множества $\mathcal{X}(\mathbb{M})$ это подмножество не образует абелеву группу (модуль) по отношению к сложению.

Пример 3.6.10. Рассмотрим два векторных поля $X = y^2\partial_x$ и $Y = x^2\partial_y$ на евклидовой плоскости $\mathbb{R}^2$. Оба поля являются полными, однако их сумма $X + Y$ неполна. Действительно, векторное поле X определяет интегральные кривые:

$$\begin{aligned} \dot{x} = y^2 \quad &\Rightarrow \quad x = y^2 t + C_0, \qquad\qquad C_0, C_1 = \mathsf{const}, \\ \dot{y} = 0 \quad &\Rightarrow \quad y = C_1. \end{aligned}$$

Эти интегральные кривые параллельны оси x и проходят через все точки y, кроме $y = 0$, см. рис.3.7,*a* (стрелки показывают возрастание параметра t). Поскольку все интегральные прямые определены при всех $t \in (-\infty, \infty)$, то векторное поле X

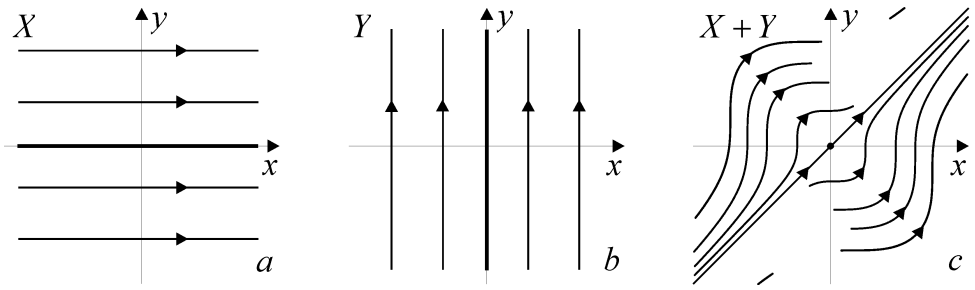


Рис. 3.7: Интегральные кривые для векторного поля X (a), Y (b) и $X+Y$ (c).

полно. Все точки оси абсцисс $(x,0) \in \mathbb{R}^2$ для векторного поля X являются вырожденными особыми точками и неподвижны. Через них не проходит ни одна интегральная кривая.

Аналогично, векторное поле Y определяет интегральные кривые:

$$\begin{aligned} &\dot{x} = 0 \quad \Rightarrow \quad x = C_2, && C_0, C_2 = \mathsf{const}, \\ &\dot{y} = x^2 \quad \Rightarrow \quad y = x^2 t + C_0. \end{aligned}$$

Интегральные кривые параллельны оси y и проходят через все точки x, кроме $x = 0$, см. рис.3.7,b. Соответствующее векторное поле Y полно. Все точки оси ординат $(0,y) \in \mathbb{R}^2$ являются вырожденными особыми точками и неподвижны.

Уравнения для интегральных кривых, определяемых суммой векторных полей, имеют вид:

$$\dot{x} = y^2, \tag{3.36}$$
$$\dot{y} = x^2. \tag{3.37}$$

При $x \neq 0$ и $y \neq 0$ отсюда следует дифференциальное уравнение для формы интегральной кривой:

$$x^2 dx = y^2 dy \quad \Rightarrow \quad y^3 = x^3 + c, \qquad c = \mathsf{const}.$$

Подставляя это решение в уравнение (3.36), получаем

$$\dot{x} = (x^3 + c)^{2/3} \quad \Leftrightarrow \quad t = \int \frac{dx}{(x^3 + c)^{2/3}}. \tag{3.38}$$

При $c = 0$ это уравнение легко интегрируется

$$x = -\frac{1}{t - c_0}, \qquad c_0 = \mathsf{const}. \tag{3.39}$$

После подстановки в (3.37) получаем уравнение

$$y = -\frac{1}{t - c_0} + c_1, \qquad c_1 = \mathsf{const}. \tag{3.40}$$

Таким образом мы получили интегральные кривые для векторного поля $X + Y$, которые неполны, т.к. уходят в бесконечность $x = \pm\infty$, $y = \pm\infty$ при конечном значении параметра $t = c_0$. Остальные интегральные кривые можно не исследовать, поскольку, по-определению, векторное поле полно, если *все* интегральные кривые определены при всех $t \in \mathbb{R}$.

Интегральные кривые для суммы векторных полей показаны на рис.3.7,*c*. Все интегральные кривые неполны в бесконечности. Это легко видеть, т.к. при $x \to \pm\infty$ постоянной c в интеграле (3.38) можно пренебречь, и, следовательно, справедлива асимптотика (3.39). Матрица B (3.33), определяющая линейное приближение, для суммы векторных полей $X + Y$ имеет вид

$$B = \begin{pmatrix} 0 & 2x \\ 2y & 0 \end{pmatrix}.$$

Она вырождена только в начале координат. Поэтому точка $(0, 0)$ является вырожденной особой точкой. К ней подходят интегральные кривые (3.39), (3.40) при бесконечном значении параметра t. □

Ранее мы показали, что всюду отличное от нуля векторное поле, заданное на многообразии $\mathbb{M}$, определяет семейство интегральных кривых, проходящих через каждую точку $\mathbb{M}$, причем через каждую точку проходит единственная кривая.

Определение. Пусть $s(t, p)$ – интегральная кривая векторного поля $X \in \mathcal{X}(\mathbb{M})$, проходящая через точку p, $s(0, p) = p$. Будем считать, что векторное поле полно для всех $p \in \mathbb{M}$. Тогда отображение

$$s: \quad \mathbb{R} \times \mathbb{M} \ni t, p \mapsto s(t, p) \in \mathbb{M}, \tag{3.41}$$

генерируемое векторным полем X, называется *потоком векторного поля.* □

Предложение 3.6.3. *Отображение (3.41) удовлетворяет тождеству*

$$s\Big(t_1, s(t_2, p)\Big) = s(t_1 + t_2, p), \tag{3.42}$$

для всех значений $t_1, t_2 \in \mathbb{R}$, для которых формула (3.42) имеет смысл.

Доказательство. Предложение следует из единственности решения системы дифференциальных уравнений. Действительно, в произвольной карте выполнено равенство

$$\frac{d}{dt_1} s^\alpha\Big(t_1, s(t_2, p)\Big) = X^\alpha\Big(s(t_2, p)\Big)$$
$$s\Big(0, s(t_2, p)\Big) = s(t_2, p).$$

С другой стороны

$$\frac{d}{dt_1} s^\alpha(t_1 + t_2, p) = \frac{d}{d(t_1 + t_2)} s^\alpha(t_1 + t_2, p) = X^\alpha\Big(s(t_1 + t_2, p)\Big),$$
$$s(0 + t_2, p) = s(t_2, p).$$

Тем самым и правая, и левая часть равенства (3.42) удовлетворяют одной и той же системы уравнений с одинаковыми начальными условиями. □

Теорема 3.6.7. *Пусть на многообразии $\mathbb{M}$ задано гладкое полное векторное поле $X \in \mathcal{X}(\mathbb{M})$. Тогда для любой точки $x \in \mathbb{M}$ существует гладкое отображение (3.41) со следующими свойствами:*

1) $s(0, x) = x$;

2) $s(t, x)$ есть решение уравнения (3.32);

3) $s\Big(t_1, s(t_2, x)\Big) = s(t_1 + t_2, x)$.

Замечание. Поток векторного поля можно представлять себе, как стационарный поток жидкости. В этом случае параметр t является временем, а X – векторным полем скорости частиц жидкости. □

Пример 3.6.11. Рассмотрим гладкое векторное поле $K = -y\partial_x + x\partial_y$ на евклидовой плоскости $\mathbb{M} = \mathbb{R}^2$. Нетрудно проверить, что поток этого векторного поля имеет вид

$$s: \quad \mathbb{R} \times \mathbb{R}^2 \ni t \times (x, y) \mapsto (x\cos t - y\sin t, x\sin t + y\cos t) \in \mathbb{R}^2.$$

Интегральная кривая, проходящая через точку (x, y), представляет собой окружность с центром в начале координат. В начале координат векторное поле обращается в нуль, и интегральная кривая вырождается в точку. Матрица (3.33) имеет вид

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Ее собственные числа являются мнимыми $\lambda_{1,2} = \pm i$. Поэтому начало координат является для векторного поля K невырожденной особой точкой – центром.

Векторное поле K является ничем иным, как векторным полем Киллинга двумерных вращений евклидовой плоскости и принимает особо простой вид в полярных координатах, $K = \partial_\varphi$. Интегральные кривые поля K являются в данном случае траекториями Киллинга. Поток векторного поля на $\mathbb{M}$ определяется независимо от наличия метрики. Однако интерпретация векторного поля K, как поля Киллинга уже связана с наличием на плоскости евклидовой метрики. □

Пример 3.6.12. Рассмотрим гладкое векторное поле $K = y\partial_x + x\partial_y$ на плоскости $\mathbb{M} = \mathbb{R}^2$. Поток этого векторного поля имеет вид

$$s: \quad \mathbb{R} \times \mathbb{R}^2 \ni t \times (x, y) \mapsto (x\,\mathsf{ch}\,t + y\,\mathsf{sh}\,t, x\,\mathsf{sh}\,t + y\,\mathsf{ch}\,t) \in \mathbb{R}^2.$$

Интегральная кривая, проходящая через точку (x, y), является ветвью гиперболы $x^2 - y^2 = \mathsf{const}$ с центром в начале координат. При $y = \pm x$ гиперболы вырождаются в прямые линии, проходящие через начало координат под углом $\pm\pi/4$. Матрица (3.33) имеет вид

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Она невырождена и имеет различные вещественные собственные значения $\lambda_{1,2} = \pm 1$. Поэтому начало координат является невырожденной особой точкой – седлом.

Векторное поле K является векторным полем Киллинга для метрики Минковского, заданной на плоскости $\mathbb{R}^2$, а интегральные кривые – траекториями Киллинга. □

Определение. При фиксированном значении параметра t поток $s(t, x)$ представляет собой диффеоморфизм, обозначаемый также

$$s_t: \quad \mathbb{M} \to \mathbb{M}.$$

Он представляет собой абелеву группу:

1) $s_{t_1} s_{t_2} = s_{t_1+t_2}$;
2) s_0 – единичный элемент;
3) $s_t^{-1} = s_{-t}$ – обратный элемент.

Эта группа называется *однопараметрической группой преобразований*, генерируемой

векторным полем X. Действительно, из системы уравнений (3.30) следует, что при малых значениях параметра поток имеет вид

$$s_\epsilon : \quad x^\alpha \to x^\alpha + \epsilon X^\alpha.$$

То есть векторное поле X является генератором бесконечно малых преобразований многообразия $\mathbb{M}$. □

Замечание. Псевдогруппа гладких преобразований координат $\mathsf{diff}\,\mathbb{M}$ на многообразии $\mathbb{M}$, бесконечномерна. С соответствующими оговорками множество гладких векторных полей $\mathcal{X}(\mathbb{M})$ можно рассматривать, как бесконечномерную алгебру Ли для $\mathsf{diff}\,\mathbb{M}$. □

Сам поток часто обозначают

$$s^\alpha(t,x) = \mathsf{exp}\,(tX)x^\alpha$$

и называют *экспоненциальным отображением*. Это обозначение оправдано следующим образом. Разложим функцию $s^\alpha(t,x)$ в ряд Тейлора по t:

$$\begin{aligned} s^\alpha(t,x) &= x^\alpha + t\,\frac{\partial s^\alpha(\tau,x)}{\partial\tau}\bigg|_{\tau=0} + \frac{t^2}{2!}\,\frac{\partial^2 s^\alpha(\tau,x)}{\partial\tau^2}\bigg|_{\tau=0} + \dots \\ &= \left[1 + t\frac{\partial}{\partial\tau} + \frac{t^2}{2!}\frac{\partial^2}{\partial\tau^2} + \dots\right] s^\alpha(\tau,x)\bigg|_{\tau=0} \\ &= \mathsf{exp}\left(t\frac{\partial}{\partial\tau}\right) s^\alpha(\tau,x)\bigg|_{\tau=0} = \mathsf{exp}\,(tX)x^\alpha. \end{aligned}$$

При этом выполнены формальные свойства экспоненты:

$$\begin{gathered} \mathsf{exp}\,(0X)x = x, \\ \mathsf{exp}\,(t_1X)\,\mathsf{exp}\,(t_2X)x = \mathsf{exp}\,[(t_1+t_2)X]x, \\ \frac{d}{dt}\,\mathsf{exp}\,(tX)x = X\,\mathsf{exp}\,(tX)x. \end{gathered}$$

Выше было показано, что любое полное векторное поле генерирует единственную однопараметрическую группу преобразований. Верно и обратное утверждение: любая однопараметрическая группа $s^\alpha(t,x)$ определяет векторное поле. Для этого достаточно положить

$$X^\alpha = \frac{ds^\alpha}{dt}\bigg|_{t=0}.$$

Если векторное поле $X(\mathbb{M})$ является неполным, то понятие потока и однопараметрической группы преобразований можно ввести только локально (см., например, [**?**].)

3.6.6 1-формы и гиперповерхности

Рассмотрим отличную от нуля в каждой точке 1-форму $dx^\alpha A_\alpha(x)$ в некоторой карте на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$. Тогда линейное алгебраическое уравнение

$$dx^\alpha A_\alpha = 0 \tag{3.43}$$

относительно дифференциалов dx^α имеет $n-1$ линейно независимых решений в каждой точке $x \in \mathbb{M}$. При этом любое решение уравнения (3.43) является линейной комбинацией данных решений. Отсюда следует, что совокупность векторов dx^α, удовлетворяющих уравнению (3.43), задает $(n-1)$-мерное подпространство в касательном пространстве $\mathbb{T}_x(\mathbb{M})$. Таким образом, отличная от нуля 1-форма задает распределение $(n-1)$-мерных подпространств в касательном расслоении $\mathbb{T}(\mathbb{M})$.

Казалось бы, что существуют такие $(n-1)$-мерные подмногообразия в $\mathbb{M}$, что касательные векторы к ним образуют $(n-1)$-мерное распределение, задаваемое 1-формой. Однако в общем случае это не так. Критерий существования таких подмногообразий дает теорема Фробениуса 3.11.4, которая будет рассмотрена несколько позже. Сейчас мы остановимся на простейшем случае.

Пусть 1-форма (3.43) является точной, т.е. имеет вид,

$$A_\alpha = \partial_\alpha f, \tag{3.44}$$

для некоторой функции $f(x) \in \mathcal{C}^1(\mathbb{M})$. Тогда уравнение (3.43) можно рассмотреть, как дифференциальное уравнение на f, любое решение которого имеет вид

$$f(x) = \mathsf{const}. \tag{3.45}$$

Уравнение (3.45) при разных значениях константы определяет семейство $(n-1)$-мерных подмногообразий в $\mathbb{M}$, которые называются *гиперповерхностями*. При $n = 2$ эти подмногообразия называются *линиями уровня*. В этом случае касательные векторы к гиперповерхностям определяют те же $(n-1)$-мерные подпространства в касательном пространстве, что и 1-форма. Говорят также, что гиперповерхность имеет коразмерность один.

Не ограничивая общности, для непостоянной функции можно считать, что только $\partial_1 f \neq 0$ в некоторой области. Поэтому функцию f можно выбрать в качестве первой координаты $f = x^1$. Тогда семейство гиперповерхностей будет задано уравнением $x^1 = \mathsf{const}$, а направления вдоль остальных координатных осей будут определять касательные к гиперповерхности направления.

3.6.7 Алгебра Ли векторных полей

Ранее мы дали два эквивалентных определения векторных полей $\mathcal{X}(\mathbb{M})$ как сечений касательного расслоения $\mathbb{T}(\mathbb{M})$ и как дифференцирований в алгебре функций $\mathcal{C}^\infty(\mathbb{M})$. Это – разные определения, одно из которых может иметь определенные преимущества в той или иной ситуации. При рассмотрении алгебраических вопросов, как правило, удобнее использовать алгебраическое определение векторных полей через дифференцирования. Используя это определение, мы введем на множестве векторных полей $\mathcal{X}(\mathbb{M})$ структуру алгебры Ли.

Определение. Последовательное применение двух двух дифференцирований (векторных полей) X и Y к некоторой функции f снова дает функцию из $\mathcal{C}^\infty(\mathbb{M})$. Определим композицию двух дифференцирований $X \circ Y$ формулой $(X \circ Y)f = X(Yf)$. Отображение $(X \circ Y)$ является линейным, однако Правило Лейбница для него не выполнено:

$$(X \circ Y)fg = XfYg + fX(Yg) + XgYf + gX(Yf) \neq f(X \circ Y)g + g(X \circ Y)f.$$

Это означает, что композиция векторных полей $X \circ Y$ не является векторным полем. По другому, отображение $X \circ Y$ в координатах содержит не только первые, но и

вторые производные. Рассмотрим композицию этих дифференцирований в другом порядке,

$$(Y \circ X)fg = YfXg + fY(Xg) + YgXf + gY(Xf).$$

Нетрудно проверить, что разность $Y \circ X - Y \circ X$ удовлетворяет правилу Лейбница

$$(X \circ Y - Y \circ X)fg = f(X \circ Y - Y \circ X)g + g(X \circ Y - Y \circ X)f,$$

т.е. является векторным полем. Эта разность называется *коммутатором векторных полей* или *скобкой Ли* и обозначатся

$$[X, Y] = X \circ Y - Y \circ X. \qquad \square$$

Из определения следует, что коммутатор двух векторных полей антисимметричен,

$$[X, Y] = -[Y, X], \tag{3.46}$$

и коммутаторы трех произвольных векторных полей удовлетворяют *тождеству Якоби*

$$\Big[[X, Y], Z\Big] + \Big[[Y, Z], X\Big] + \Big[[Z, X], Y\Big] = 0, \tag{3.47}$$

где слагаемые отличаются циклической перестановкой.

Замечание. Антисимметрия коммутатора (3.46) эквивалентна условию $[X, X] = 0$. Действительно, то, что это условие вытекает из (3.46) очевидно. Для доказательства обратного утверждения достаточно рассмотреть уравнение $[X + Y, X + Y] = 0$. $\square$

Алгебра Ли является неассоциативной алгеброй, при этом условие ассоциативности заменяется на тождества Якоби.

Рассмотрим векторные поля в произвольной карте $X = X^\alpha \partial_\alpha$, $Y = Y^\alpha \partial_\alpha$. Тогда коммутатор дает новое векторное поле

$$Z = [X, Y] = (X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha)\partial_\alpha. \tag{3.48}$$

Используя закон преобразования компонент векторных полей (3.11), нетрудно проверить, что выражение в правой части инвариантно относительно преобразования координат.

Определение. Множество векторных полей с операцией сложения и коммутирования, которое удовлетворяет условиям (3.46) и (3.47) образует *кольцо Ли*. Умножение векторного поля на числа снова дает векторной поле, при этом коммутатор (3.48) билинеен

$$[aX + bY, Z] = a[X, Z] + b[Y, Z].$$
$$[X, aY + bZ] = a[X, Y] + b[X, Z],$$

где $a, b \in \mathbb{R}$. Множество векторных полей с операциями умножения на числа, сложения и коммутирования образует *алгебру Ли* над полем вещественных чисел. Эта алгебра бесконечномерна и также обозначается $\mathcal{X}(\mathbb{M})$. $\square$

Пример 3.6.13. Двумерное векторное пространство с базисом e_1 и e_2 становится алгеброй Ли, если положить

$$[e_1, e_1] = [e_2, e_2] = 0, \qquad [e_1, e_2] = e_2$$

и продолжить эту операцию по линейности. $\square$

Пример 3.6.14. Трехмерное векторное пространство $\mathbb{R}^3$ с ортонормальным базисом e_i, $i = 1, 2, 3$, является алгеброй Ли, если в качестве коммутатора двух векторов $X, Y \in \mathbb{R}^3$ выбрать их векторное произведение

$$[X, Y]^i = -\varepsilon^{ijk} X_j Y_k,$$

где ε^{ijk} – полностью антисимметричный тензор третьего ранга и опускание индексов производится с помощью евклидовой метрики δ_{ij}. Эта алгебра Ли совпадает с алгеброй Ли группы трехмерных вращений $\mathbb{SO}(3)$ (2.119). □

Пример 3.6.15. Векторное пространство $\mathfrak{gl}(n, \mathbb{R})$ всех вещественных $n \times n$-матриц образует конечномерную алгебру Ли, если положить

$$[A, B] = AB - BA, \qquad A, B \in \mathfrak{gl}(n, \mathbb{R}),$$

где AB – обычное произведение матриц. При этом $\dim \mathfrak{gl}(n, \mathbb{R}) = n^2$. □

Забегая вперед (см. раздел 7.2), заметим, что коммутатор векторных полей (3.48) можно записать в эквивалентном виде, используя ковариантную производную и тензор кручения

$$[X, Y] = (X^\beta \nabla_\beta Y^\alpha - Y^\beta \nabla_\beta X^\alpha - X^\beta Y^\gamma T_{\beta\gamma}{}^\alpha)\partial_\alpha.$$

В таком виде правая часть явно ковариантна.

3.7 Тензорные поля

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$. В каждой точке $x \in \mathbb{M}$ у нас есть два n-мерных векторных пространства: касательное $\mathbb{T}_x(\mathbb{M})$ и кокасательное $\mathbb{T}^*_x(\mathbb{M})$ пространства. Рассмотрим их тензорное произведение

$$\mathbb{T}^r_{s,x}(\mathbb{M}) = \underbrace{\mathbb{T}_x(\mathbb{M}) \otimes \ldots \otimes \mathbb{T}_x(\mathbb{M})}_{r} \otimes \underbrace{\mathbb{T}^*_x(\mathbb{M}) \otimes \ldots \otimes \mathbb{T}^*_x(\mathbb{M})}_{s}, \tag{3.49}$$

где мы взяли r экземпляров касательного и s экземпляров кокасательного пространства. Для определенности мы фиксировали порядок сомножителей. Таким образом в каждой точке многообразия мы построили векторное пространство размерности $\dim \mathbb{T}^r_{s,x}(\mathbb{M}) = n^{r+s}$.

Определение. Объединение

$$\mathbb{T}^r_s(\mathbb{M}) = \bigcup_{x \in \mathbb{M}} \mathbb{T}^r_{s,x}(\mathbb{M}),$$

взятое по всем точкам многообразия, называется *расслоением тензоров типа* (r, s) на многообразии $\mathbb{M}$. Сечение этого расслоения $T^r_s(x)$ называется *тензорным полем типа* (r, s) или r раз контравариантным и s раз ковариантным тензорным полем на многообразии $\mathbb{M}$. Число $r + s$ называется *рангом* тензорного поля. □

Базой этого расслоения является многообразие $\mathbb{M}$, типичным слоем – векторное пространство

$$\underbrace{\mathbb{R}^n \otimes \ldots \otimes \mathbb{R}^n}_{r} \otimes \underbrace{\mathbb{R}^n \otimes \ldots \otimes \mathbb{R}^n}_{s},$$

где $\mathbb{R}^n$ – типичный слой касательного расслоения. Слоем над $x \in \mathbb{M}$ является векторное пространство (3.49) (тем самым мы определили проекцию). Дифференцируемая структура на расслоении тензоров задается дифференцируемыми структурами на базе и в типичном слое аналогично тому, как она была построена для касательного расслоения.

Координатные базисы в касательном и кокасательном пространствах, $e_\alpha = \partial_\alpha$ и $e^\alpha = dx^\alpha$, индуцируют координатный базис в тензорном произведении, который мы обозначим

$$e_{\alpha_1} \otimes \ldots \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes \cdots \otimes e^{\beta_s}. \tag{3.50}$$

Напомним, что тензорное произведение векторов не является коммутативным,

$$e_\alpha \otimes e^\beta \neq e^\beta \otimes e_\alpha,$$

поэтому порядок следования базисных векторов в произведении (3.50) фиксирован: сначала мы пишем базисные векторы касательного, а затем кокасательного пространств.

Рассмотрим произвольную карту $(\mathbb{U}, \varphi)$ на многообразии. Тогда тензорное поле типа (r, s) в координатах имеет вид

$$T^r_s(x) = T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x)\, e_{\alpha_1} \otimes \ldots \otimes e_{\alpha_r} \otimes e^{\beta_1} \otimes \ldots \otimes e^{\beta_s}. \tag{3.51}$$

Нижние и верхние индексы называют соответственно *ковариантными* и *контравариантными*. Общее число индексов $r + s$ равно рангу тензорного поля.

Замечание. Для определенности, у компонент $T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x)$ мы сначала выписали все ковариантные индексы, а затем – все контравариантные. Порядок индексов зафиксирован порядком сомножителей в правой части (3.49) и принятом нами соглашением для записи тензорного поля типа $(1, 1)$:

$$dx^\beta T_\beta{}^\alpha \partial_\alpha.$$

Контравариантные индексы, так же, как и ковариантные, упорядочены между собой. В разделе 5 будет введена операция опускания и подъема индексов с помощью метрики. Она будет неоднозначной, если порядок контравариантных и ковариантных индексов не фиксирован.

Ниже мы построим тензорную алгебру для тензорных полей вида (3.51). Аналогично можно построить тензорную алгебру для произвольного расположения сомножителей в правой части (3.49), когда касательные и кокасательные пространства чередуются в произвольном порядке. Мы будем предполагать, что все индексы упорядочены определенным образом. В этом случае обозначение векторного поля $T^r_s(x)$ является грубым, т.к. учитывает только общее число ковариантных и контравариантных индексов, а не их последовательность. □

Набор функций $X_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x)$ с r верхними и s нижними индексами называется *компонентами* тензорного поля типа (r, s) в карте $(\mathbb{U}, \varphi)$. Тензорное поле называется гладким, если все компоненты – гладкие функции. При преобразованиях координат каждый контравариантный индекс умножается на матрицу Якоби (2.57), так же, как и компоненты вектора, а каждый ковариантный индекс – на обратную матрицу Якоби, так же, как и 1-форма.

Пример 3.7.1. Компоненты тензорного поля типа (1,1) преобразуются по правилу:

$$X_{\beta'}{}^{\alpha'} = \frac{\partial x^{\beta}}{\partial x^{\beta'}} X_{\beta}{}^{\alpha} \frac{\partial x^{\alpha'}}{\partial x^{\alpha}}. \tag{3.52}$$

Аналогично преобразуются компоненты тензорных полей произвольного типа. □

В дальнейшем, для краткости, тензорные поля мы часто будем называть просто тензорами.

Пример 3.7.2. Очевидно, что, если все компоненты тензорного поля равны нулю в какой-то одной системе координат, то они равны нулю во всех остальных системах отсчета. У нетривиальных тензоров хотя бы одна компонента должна быть отлична от нуля. В общем случае у тензора типа (r, s) на многообразии размерности n имеется n^{r+s} независимых компонент в каждой точке. □

Обозначим множество гладких тензорных полей типа (r, s) символом $\mathcal{T}^r_s(\mathbb{M})$. При этом $\mathcal{T}^0_0(\mathbb{M}) = \mathcal{C}^\infty(\mathbb{M})$, $\mathcal{T}^1_0 = \mathcal{X}(\mathbb{M})$ и $\mathcal{T}^0_1 = \Lambda_1(\mathbb{M})$. В дальнейшем индекс 0 у множеств тензорных полей, имеющих только контравариантные или ковариантные индексы, писаться не будет: $\mathcal{T}^r_0(\mathbb{M}) = \mathcal{T}^r(\mathbb{M})$ и $\mathcal{T}^0_s(\mathbb{M}) = \mathcal{T}_s(\mathbb{M})$.

Тензорные поля фиксированного типа в каждой точке можно складывать и умножать на числа, т.е. они образуют (бесконечномерное) векторное пространство над полем вещественных чисел. Кроме того, тензорное поле произвольного типа можно поточечно умножать на произвольные функции, при этом получится новое тензорное поле того же типа. Таким образом они образуют модуль над алгеброй гладких функций $\mathcal{C}^\infty(\mathbb{M})$.

Введем обозначение для прямой суммы тензорных полей

$$\mathcal{T}(\mathbb{M}) = \bigoplus_{r,s=0}^{\infty} \mathcal{T}^r_s(\mathbb{M}).$$

На множестве $\mathcal{T}(\mathbb{M})$ можно ввести поточечное тензорное умножение, которое двум тензорам типа (r, s) и (p, q) ставит в соответствие тензорное поле типа $r + p, s + q$. А именно, зафиксируем точку $x \in \mathbb{M}$. Из универсального факторизационного свойства тензорного произведения следует, что существует единственное билинейное отображение из $\mathbb{T}^r_{s,x} \times \mathbb{T}^p_{q,x}$ в $\mathbb{T}^{r+p}_{s+q,x}$, которое отображает пару тензоров

$$(X_1 \otimes \cdots \otimes X_r \otimes A_1 \otimes \cdots \otimes A_s) \in \mathbb{T}^r_{s,x},$$
$$(Y_1 \otimes \cdots \otimes Y_p \otimes B_1 \otimes \cdots \otimes B_q) \in \mathbb{T}^p_{q,x}$$

в тензор типа $r + p, s + q$:

$$X_1 \otimes \cdots \otimes X_r \otimes Y_1 \otimes \cdots \otimes Y_p \otimes A_1 \otimes \cdots A_s \otimes B_1 \otimes \cdots \otimes B_q \quad \in \mathbb{T}^{r+p}_{s+q,x}. \tag{3.53}$$

Это отображение называется тензорным произведением тензоров в данной точке.

Замечание. Мы зафиксировали порядок сомножителей в произведении (3.53), который соответствует тензорному произведению (3.49).

Определение. *Тензорным произведением* тензорных полей $\mathcal{T}^r_s(\mathbb{M})$ и $\mathcal{T}^p_q(\mathbb{M})$ называется тензорное поле типа $\mathcal{T}^{r+p}_{s+q}(\mathbb{M})$, полученное поточечным тензорным произведением (3.53). □

Чтобы получить выражение для компонент тензорного произведения в определенной карте, достаточно в качестве векторных X, Y и ковекторных A, B полей в определении тензорного произведения выбрать координатный базис. Пусть в некоторой карте задано два тензорных поля:

$$K = K_{\beta_1\cdots\beta_s}{}^{\alpha_1\cdots\alpha_r}\, e_{\alpha_1}\otimes e_{\alpha_r}\otimes e^{\beta_1}\cdots\otimes e^{\beta_s} \quad \in \mathcal{T}^r_s(\mathbb{M}),$$
$$L = L_{\delta_1\cdots\delta_q}{}^{\gamma_1\cdots\gamma_p}\, e_{\gamma_1}\otimes e_{\gamma_p}\otimes e^{\delta_1}\cdots\otimes e^{\delta_q} \quad \in \mathcal{T}^p_q(\mathbb{M}).$$

Тогда компоненты их тензорного произведения

$$(K\otimes L)_{\beta_1\cdots\beta_s\delta_1\cdots\delta_q}{}^{\alpha_1\cdots\alpha_r\gamma_1\cdots\gamma_p} = K_{\beta_1\cdots\beta_s}{}^{\alpha_1\cdots\alpha_r} L_{\delta_1\cdots\delta_q}{}^{\gamma_1\cdots\gamma_p}$$

просто равны произведению компонент каждого сомножителя, как чисел.

Пример 3.7.3. Произведение двух векторных полей $X\otimes Y$ дает контравариантный тензор второго ранга с компонентами

$$T(x) = T^{\alpha\beta}(x)e_\alpha\otimes e_\beta = X(x)\otimes Y(x) = X^\alpha(x)Y^\beta(x)e_\alpha\otimes e_\beta,$$

Эта операция является некоммутативной, поскольку первый индекс поля $T(x)$ относится к векторному полю X, а не Y. □

Вместе с тензорным умножением, множество тензорных полей $\mathcal{T}(\mathbb{M})$ образует некоммутативную ассоциативную *тензорную алгебру* над полем вещественных чисел. Эта алгебра бесконечномерна, поскольку векторное пространство тензоров фиксированного типа бесконечномерно само по себе и, вдобавок, ранг тензоров неограничен. Алгебра тензоров имеет естественную градуировку, как прямая сумма тензоров фиксированного типа. Образующими тензорной алгебры являются векторные поля и 1-формы.

Определение. Каждой паре индексов (i, j) таких, что $1\le i\le r$ и $1\le j\le s$, мы ставим в соответствие линейное отображение $C^{ij} : \mathcal{T}^r_s(\mathbb{M}) \to \mathcal{T}^{r-1}_{s-1}(\mathbb{M})$ с помощью следующей формулы

$$X_1\otimes\ldots\otimes X_r\otimes A_1\otimes\ldots\otimes A_s \quad\mapsto$$
$$\to (A_j, X_i)X_1\otimes\ldots X_{i-1}\otimes X_{i+1}\otimes\ldots\otimes X_r\otimes A_1\otimes\ldots\otimes A_{j-1}\otimes\ldots\otimes A_{j+1}\otimes A_s,$$

где $(A_j, X_i) = A_j(X_i)$ – значение 1-формы A_j на векторе X_i. Это отображение называется *сверткой* и обозначается C^{ij}. □

Компоненты свернутого тензора $K\in\mathcal{T}^r_s(\mathbb{M})$ имеют вид

$$K_{\beta_1\cdots\beta_{j-1}\gamma\beta_{j+1}\cdots\beta_s}{}^{\alpha_1\cdots\alpha_{i-1}\gamma\alpha_{i+1}\cdots\alpha_r} = \delta^{\beta_j}_{\alpha_i} K_{\beta_1\cdots\beta_s}{}^{\alpha_1\cdots\alpha_r},$$

где произведена свертка по одному верхнему и одному нижнему индексу.

Пример 3.7.4. Тензору типа $(1,1)$ ставится в соответствие скалярное поле $\mathsf{tr}\, K = K_\alpha{}^\alpha$, которое называется *следом* тензора $K = K_\beta{}^\alpha e_\alpha\otimes e^\beta$. □

Пример 3.7.5. Значением 1-формы $A = dx^\alpha A_\alpha$ на векторном поле $X = X^\alpha\partial_\alpha$ является свертка тензорного произведения $A\otimes X$: $\langle X, A\rangle = A(X) = X^\alpha A_\alpha$. □

Определение. Тензорное поле $K \in \mathcal{T}_s^r(\mathbb{M})$ называется *разложимым*, если его можно представить в виде

$$K = X_1 \otimes \cdots \otimes X_r \otimes A_1 \otimes \cdots \otimes A_s,$$

для некоторых векторов: $X_i \in \mathcal{T}^1(\mathbb{M})$, $i = 1, \cdots, r$ и 1-форм: $A_j \in \mathcal{T}_1(\mathbb{M}) = \Lambda_1(\mathbb{M})$, $j = 1, \cdots, s$. □

Пример 3.7.6. Сумма двух разложимых контравариантных тензоров второго ранга

$$X_1 \otimes Y_1 + X_2 \otimes Y_2 = (X_1^\alpha Y_1^\beta + X_2^\alpha Y_2^\beta) e_\alpha \otimes e_\beta$$

может не быть разложимым тензором. □

Если тензор имеет два или более индексов одного типа, то с помощью симметризации или антисимметризации по верхним или нижним индексам можно строить новые тензорные поля. Поскольку преобразование координат действует одинаково на каждый ковариантный и контравариантный индекс, то симметризация и антисимметризация индексов является инвариантной операцией и свойство симметрии по индексам сохраняется при преобразовании координат.

Тензорное поле называется *неприводимым*, если нельзя найти такие линейные комбинации его компонент с постоянными коэффициентами, которые сами образовывали бы тензор.

Пример 3.7.7. Скалярные, векторные поля и 1-формы являются неприводимыми тензорными полями. □

Пример 3.7.8. Ковариантные или контравариантные тензорные поля второго ранга приводимы, т.к. их компоненты можно разложить на симметричную и антисимметричную неприводимую части

$$X_{\alpha\beta} = X_{(\alpha\beta)} + X_{[\alpha\beta]}, \tag{3.54}$$

где

$$X_{(\alpha\beta)} = \frac{1}{2}(X_{\alpha\beta} + X_{\beta\alpha}), \qquad X_{[\alpha\beta]} = \frac{1}{2}(X_{\alpha\beta} - X_{\beta\alpha}).$$ □

Замечание. Выделение следа у тензоров со всеми ковариантными или контравариантными индексами невозможно без наличия метрики. □

Пример 3.7.9. Тензорное поле $X_\alpha{}^\beta$ типа (1,1) также приводимо, поскольку у него можно выделить след X и бесследовую часть $Y_\alpha{}^\beta$ ($Y_\alpha{}^\alpha = 0$)

$$X_\alpha{}^\beta = Y_\alpha{}^\beta + \frac{1}{n}\delta_\alpha^\beta \,\mathsf{tr}\, X, \tag{3.55}$$

где

$$\mathsf{tr}\, X = X_\alpha{}^\alpha, \qquad Y_\alpha{}^\beta = X_\alpha{}^\beta - \frac{1}{n}\delta_\alpha^\beta \,\mathsf{tr}\, X.$$ □

Тензорные поля более высокого ранга, в общем случае, приводимы. Если на многообразии не задано никаких других объектов, кроме тензорного поля, то разложение на неприводимые компоненты может осуществляться только с помощью взятия следа, симметризации или антисимметризации по индексам.

Замечание. При проведении вычислений с тензорными полями важно иметь ввиду следующее обстоятельство. Если тензорное поле приводимо, то его разложение на неприводимые компоненты в большинстве случаев упрощает вычисления и проясняет математическую структуру модели. □

Пример 3.7.10. *Символ Кронекера*, компоненты которого в каждой карте многообразия $\mathbb{M}$ составляют n-мерную единичную матрицу,

$$\delta^\alpha_\beta = \begin{cases} 1, & \alpha = \beta, \\ 0, & \alpha \neq \beta, \end{cases} \tag{3.56}$$

и имеют один верхний и один нижний индекс, определяет тензорное поле типа $(1,1)$. Он инвариантен относительно преобразований координат

$$\delta^{\alpha'}_{\beta'} = \frac{\partial x^\beta}{\partial x^{\beta'}} \delta^\alpha_\beta \frac{\partial x^{\alpha'}}{\partial x^\alpha} = \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{\partial x^{\alpha'}}{\partial x^\beta} = \begin{cases} 1, & \alpha' = \beta', \\ 0, & \alpha' \neq \beta', \end{cases}$$

т.к. верхний и нижний индексы преобразуется с помощью взаимно обратных матриц. Символ Кронекера представляет собой исключение в двух отношениях. Во-первых, он инвариантен относительно преобразований координат и, во-вторых, его индексы можно писать один под другим, поскольку подъем и опускание индексов с помощью метрики приводит к симметричным тензорам. □

Тензорное поле $T_\alpha{}^\beta \in \mathcal{T}^1_1(\mathbb{M})$ типа $(1,1)$ можно рассматривать, как линейный оператор (эндоморфизм), действующий в пространстве векторов $\mathcal{T}^1(\mathbb{M})$ и 1-форм $\mathcal{T}_1(\mathbb{M}) = \Lambda_1(\mathbb{M})$. В компонентах действие оператора задается правилами:

$$A'_\alpha = T_\alpha{}^\beta A_\beta, \qquad X'^\alpha = X^\beta T_\beta{}^\alpha.$$

Действие оператора $T_\alpha{}^\beta$ естественным образом распространяется на тензоры произвольного типа.

Пример 3.7.11. Символ Кронекера представляет собой тождественный оператор. □

Пример 3.7.12. Проекционные операторы (5.25) являются тензорными полями типа $(1,1)$. □

На многообразии $\mathbb{M}$ можно также определить *тензорные плотности* степени p и ранга (r,s), если при преобразовании координат все их компоненты умножить на якобиан преобразования в степени p, как и в случае скалярных полей (3.7). Например, тензорная плотность типа $(1,1)$ и степени p преобразуется по-правилу

$$X_{\alpha'}{}^{\beta'} = J^p \frac{\partial x^\alpha}{\partial x^{\alpha'}} X_\alpha{}^\beta \frac{\partial x^{\beta'}}{\partial x^\beta}. \tag{3.57}$$

В каждой точке многообразия тензорные плотности фиксированного типа и степени образуют векторное пространство над полем вещественных чисел и модуль над алгеброй гладких функций. По аналогии с тензорным произведением тензоров можно ввести тензорное произведение плотностей, которое двум плотностям типа (r_1,s_1), (r_2,s_2) и степеней p_1 и p_2 ставит в соответствие тензорную плотность типа (r_1+r_2, s_1+s_2) и степени p_1+p_2. Если ограничиться случаем, когда степени плотностей p пробегают счетное множество значений, то множество тензорных плотностей

и всех их линейных комбинаций в фиксированной точке образует некоммутативную ассоциативную алгебру над полем вещественных чисел. Эта алгебра имеет естественную градуировку, как прямая сумма тензорных плотностей фиксированного типа и степени.

Так же, как и в случае скалярных плотностей тензорные плотности ${X_{\beta_1\dots\beta_s}}^{\alpha_1\dots\alpha_r}$ степени p можно представить в виде

$$X_{\beta_1\dots\beta_s}{}^{\alpha_1\dots\alpha_r} = \sqrt{|g|}^{-p} Y_{\beta_1\dots\beta_s}{}^{\alpha_1\dots\alpha_r},$$

где $\sqrt{|g|} = \det {e_\alpha}^a$ – определитель репера, а ${Y_{\beta_1\dots\beta_s}}^{\alpha_1\dots\alpha_r}$ – тензорное поле того же типа, что и исходная тензорная плотность.

Множество тензорных полей можно интерпретировать, как множество полилинейных отображений.

Теорема 3.7.1. *Множество ковариантных тензорных полей $\mathcal{T}_s(\mathbb{M})$ можно рассматривать, как s линейное отображение $\big(\mathcal{C}(\mathbb{M})$-модуль$\big)$ из $\underbrace{\mathcal{X}\times\ldots\times\mathcal{X}}_{s}$ в алгебру непрерывных функций $\mathcal{C}(\mathbb{M})$ такое, что*

$$K(f_1X_1,\dots,f_sX_s) = f_1\dots f_sK(X_1,\dots,X_s), \qquad K(x)\in\mathcal{T}_s(\mathbb{M}),$$

для всех $f_i\in\mathcal{C}(\mathbb{M})$ и $X_i\in\mathcal{X}(\mathbb{M})$. Обратно, каждое такое отображение можно рассматривать, как тензорное поле типа $(0,s)$.

Доказательство. См., например, [?]. □

Аналогично можно интерпретировать тензоры произвольного типа (r,s), как множество всех $r+s$ линейных отображений:

$$\mathcal{T}^r_s(\mathbb{M}):\ \underbrace{\mathcal{X}\times\ldots\times\mathcal{X}}_{s}\times\underbrace{\Lambda_1\times\ldots\times\Lambda_1}_{r}\ \to\ \mathcal{C}(\mathbb{M}). \tag{3.58}$$

Поскольку между тензорными полями и полилинейными отображениями существует взаимно однозначное соответствие, то некоторые авторы принимают эти отображения в качестве глобального определения тензорных полей.

3.8 Полностью антисимметричные тензоры

В настоящем разделе мы рассмотрим полностью антисимметричные тензоры, которые играют очень важную роль в различных приложениях дифференциальной геометрии.

Определение. Рассмотрим тензорные поля типа $(r,0)$ или $(0,r)$ при $r\le n$ на многообразии $\mathbb{M}$, $\dim\mathbb{M}=n$, компоненты которых антисимметричны относительно перестановки любой пары индексов. В инвариантном виде условие антисимметричности для ковариантных тензоров записывается в виде:

$$T(X_1,\dots,X_i,\dots,X_j,\dots,X_r) = -T(X_1,\dots,X_j,\dots,X_i,\dots,X_r), \qquad 1\le i<j\le r,$$

где $T\in\mathcal{T}_r(\mathbb{M})$, $X_1,\dots,X_r\in\mathcal{X}(\mathbb{M})$. Эти тензоры неприводимы и называются *полностью антисимметричными* ковариантными тензорами ранга r. □

Компонента полностью антисимметричного тензорного поля может быть отлична от нуля только в том случае, если все индексы различны, поскольку при совпадении двух или более индексов соответствующая компонента равна нулю. На многообразии размерности n не существует полностью антисимметричного тензора ранга большего, чем размерность многообразия, т.к. в этом случае по крайней мере два индекса будут совпадать.

Очевидно, что число независимых компонент полностью антисимметричного тензора ранга r равно числу выборок r различных индексов из n: $C_n^r = n!/r!(n-r)!$. В частности, полностью антисимметричный тензор $X_{\alpha_1\ldots\alpha_n}$ максимального ранга n имеет только одну независимую компоненту. Нетрудно проверить, что при преобразовании координат $x \to y(x)$ полностью антисимметричный тензор типа $(0,n)$ преобразуется по закону

$$X'_{\alpha_1\ldots\alpha_n} = \frac{\partial x^{[\beta_1}}{\partial y^{\alpha_1}} \cdots \frac{\partial x^{\beta_n]}}{\partial y^{\alpha_n}} X_{\beta_1\ldots\beta_n} = X_{\alpha_1\ldots\alpha_n} J^{-1}, \tag{3.59}$$

где $J = \det(\partial y^\alpha/\partial x^\beta)$ – якобиан преобразования координат. То есть каждая компонента полностью антисимметричного ковариантного тензора максимального ранга умножается на якобиан преобразования в минус первой степени, и ее фиксированную компоненту можно рассматривать как скалярную плотность веса -1.

В каждой карте $(\mathbb{U}, \varphi)$ можно построить полностью антисимметричный объект, компоненты которого равны по модулю единице

$$\hat{\varepsilon}_{\alpha_1\ldots\alpha_n} = \operatorname{sgn}\sigma(\alpha_1\ldots\alpha_n), \qquad \hat{\varepsilon}_{1\ldots n} = 1, \tag{3.60}$$

где $\operatorname{sgn}\sigma$ – *знак перестановки* σ, который равен $+1$ или -1, если для получения последовательности индексов $\alpha_1, \ldots, \alpha_n$ из последовательности натуральных чисел $1, \ldots, n$ необходимо переставить соответственно четное и нечетное число пар индексов. Объект $\hat{\varepsilon}_{\alpha_1\ldots\alpha_n}$ не может быть тензором, т.к. в общем случае якобиан преобразования J отличен от единицы. Из закона преобразования тензорных плотностей (3.7) следует, что каждую фиксированную компоненту $\hat{\varepsilon}_{\alpha_1\ldots\alpha_n}$ можно рассматривать, как скалярную тензорную плотность степени 1. Поскольку компоненты антисимметричной тензорной плотности постоянны, то в произвольной системе координат имеем

$$\partial_\alpha \hat{\varepsilon}_{\alpha_1\ldots\alpha_n} = 0. \tag{3.61}$$

В римановом пространстве при наличии метрики $g_{\alpha\beta}$ (см. раздел 5.1) можно построить полностью антисимметричный тензор

$$\varepsilon_{\alpha_1\ldots\alpha_n} = \sqrt{|g|}\,\hat{\varepsilon}_{\alpha_1\ldots\alpha_n}, \tag{3.62}$$

где введено сокращенное обозначение для определителя метрики, которое будет часто использоваться в дальнейшем

$$g = \det g_{\alpha\beta}. \tag{3.63}$$

Полностью антисимметричный тензор (3.62) преобразуется по стандартному закону (3.59).

Замечание. Здесь и в дальнейшем мы примем следующее обозначение: шляпка над символом означает, что рассматривается тензорная плотность, а не тензор. □

Полностью антисимметричные тензоры типа $(k,0)$ со всеми контравариантными индексами называются *поливекторами*. Поливектор максимального ранга n имеет одну независимую компоненту, которая преобразуется по правилу

$$X'^{\alpha_1\ldots\alpha_n} = X^{\alpha_1\ldots\alpha_n} J, \tag{3.64}$$

В пространстве контравариантных тензоров можно ввести полностью антисимметричную тензорную плотность степени -1 с компонентами, равными по модулю единице, аналогично тому, как это было сделано для ковариантных тензоров. Мы положим

$$\hat{\varepsilon}^{\alpha_1\ldots\alpha_n} = \hat{\varepsilon}_{\alpha_1\ldots\alpha_n}\,\mathsf{sgn}\,, \tag{3.65}$$

где множитель sgn зависит от того задана ли на многообразии метрика или нет

$$\mathsf{sgn} = \begin{cases} 1, & \text{если метрика не задана,} \\ \mathsf{sgn}\,(\,\mathsf{det}\, g_{\alpha\beta}), & \text{если метрика задана.} \end{cases} \tag{3.66}$$

Компоненты этой плотности также равны по модулю единице и постоянны $\partial_\alpha \hat{\varepsilon}^{\alpha_1\ldots\alpha_n} = 0$. При наличии метрики можно ввести полностью антисимметричный контравариантный тензор

$$\varepsilon^{\alpha_1\ldots\alpha_n} = \frac{1}{\sqrt{|g|}}\hat{\varepsilon}^{\alpha_1\ldots\alpha_n}. \tag{3.67}$$

Замечание. Отметим, что равенство (3.65) имеет смысл, несмотря на то, что индексы слева контравариантны, а справа ковариантны, т.к. компоненты тензорных плотностей $\hat{\varepsilon}$ не зависят от выбора системы координат. Появление множителя (3.66) в определении (3.65) объясняется тем, что на (псевдо-)римановом многообразии мы требуем, чтобы тензор с контравариантными индексами можно было бы получить из тензора с ковариантными индексами простым подъемом индексов:

$$\varepsilon^{\alpha_1\ldots\alpha_n} = g^{\alpha_1\beta_1}\ldots g^{\alpha_n\beta_n}\varepsilon_{\beta_1\ldots\beta_n}.$$

Отсюда следует равенство (3.65). □

Наличие тензора $\varepsilon_{\alpha_1\ldots\alpha_n}$ позволяет представить компоненты произвольного ковариантного антисимметричного тензора максимального ранга в виде

$$X_{\alpha_1\ldots\alpha_n} = X^*(x)\varepsilon_{\alpha_1\ldots\alpha_n}, \qquad X^* = \hat{\varepsilon}^{\alpha_1\ldots\alpha_n} X_{\alpha_1\ldots\alpha_n}\,\mathsf{sgn}\,, \tag{3.68}$$

где $X^*(x)$ – некоторое (псевдо-)скалярное поле.

Аналогичное представление имеет место для произвольного контравариантного тензора максимального ранга. Это означает, что полностью антисимметричные тензоры максимального ранга имеют ровно одну независимую компоненту.

3.9 Отображения многообразий

Рассмотрим отображение h многообразия $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = m$, в многообразие $\mathbb{N}$, $\mathsf{dim}\,\mathbb{N} = n$,

$$h: \quad \mathbb{M} \ni x \mapsto y \in \mathbb{N}. \tag{3.69}$$

Многообразие $\mathbb{N}$ мы будем называть *пространством-мишенью*. Пусть при этом отображении карта $(\mathbb{U},\varphi)$ многообразия $\mathbb{M}$ отображается в некоторую карту $(\mathbb{V},\phi)$ многообразия $\mathbb{N}$, $h(\mathbb{U})\subset\mathbb{V}$. Обозначим координаты на $\mathbb{U}$ и $\mathbb{V}$ через x^α, $\alpha=1,\dots,m$, и y^μ, $\mu=1,\dots,n$. Тогда отображение

$$\phi\circ h\circ\varphi^{-1}:\quad \mathbb{R}^m\supset\varphi(\mathbb{U})\ \to\ \phi(\mathbb{V})\subset\mathbb{R}^n \tag{3.70}$$

двух областей евклидова пространства $\varphi(\mathbb{U})\subset\mathbb{R}^m$ и $\phi(\mathbb{V})\subset\mathbb{R}^n$ задается n функциями от m переменных $y^\mu(x)$. При этом размерность многообразия $\mathbb{M}$ может быть меньше, равна или больше размерности $\mathbb{N}$.

Определение. Отображение гладких многообразий называется *гладким (дифференцируемым)*, если задается гладкими (дифференцируемыми) функциями (3.70) в полных атласах на $\mathbb{M}$ и $\mathbb{N}$. □

В дальнейшем, если не оговорено противное, мы будем рассматривать гладкие отображения.

Определение. Дифференцируемое отображение (3.69) индуцирует линейное отображение касательных пространств

$$h_*:\quad \mathbb{T}_x(\mathbb{M})\ni X=X^\alpha\partial_\alpha\ \mapsto\ Y=Y^\mu\partial_\mu\in\mathbb{T}_{h(x)}(\mathbb{N}) \tag{3.71}$$

следующим образом. Рассмотрим кривую $\gamma_p\subset\mathbb{M}$, проходящую через произвольную точку $p\in\mathbb{M}$ в направлении произвольного вектора $X_{(\gamma)}(p)$. Эта кривая отобразится в некоторую кривую $h(\gamma_p)$ на $\mathbb{N}$. По определению, вектор $X_{(\gamma)}(p)\in\mathbb{T}_p(\mathbb{M})$ отображается в тот вектор $Y_{(h\circ\gamma)}\in\mathbb{T}_{h(p)}(\mathbb{N})$, который касается кривой $h(\gamma_p)$ в точке $h(p)$. Поскольку вектор в точке p – это класс эквивалентности кривых, то это условие записывается в виде

$$h_*\Big(X_{(\gamma)}(p)\Big)=[h\circ\gamma_p]=Y_{(h\circ\gamma)}\Big(f(p)\Big),\qquad X_{(\gamma)}(p)=[\gamma_p].$$

Теперь мы упростим обозначения, опустив индекс кривой у вектора и обозначение точки p. Поскольку по правилу дифференцирования сложной функции

$$\dot y^\mu\Big(\phi\circ h\circ\gamma(t)\Big)=\frac{\partial y^\mu}{\partial x^\alpha}\dot x^\alpha\Big(\gamma(t)\Big),$$

то вектор Y единственен и не зависит от представителя класса эквивалентности кривых, определяющих вектор X в точке $x\in\mathbb{M}$. Это отображение касательных пространств (3.71) называется *дифференциалом отображения*. □

В компонентах дифференциал отображения задается матрицей Якоби:

$$Y^\mu=X^\alpha\frac{\partial y^\mu}{\partial x^\alpha}. \tag{3.72}$$

Дифференциал отображения h_* является линейным отображением:

$$\begin{aligned} h_*(X+Y)&=h_*(X)+h_*(Y),\qquad X,Y\in\mathcal{X}(\mathbb{M}),\\ h_*(aX)&=ah_*(X),\qquad a\in\mathbb{R}. \end{aligned}$$

множества всех векторных полей $\mathcal{X}(\mathbb{M})$ на $\mathbb{M}$ в множество векторных полей $\mathcal{X}\Big(h(\mathbb{M})\Big)$ на образе $h(\mathbb{M})\subset\mathbb{N}$, который может не совпадать со всем $\mathbb{N}$.

Предложение 3.9.1. *Дифференциал отображения согласован со структурами алгебры Ли в пространствах векторных полей, т.е.*

$$h_*[X,Y] = [h_*X, h_*Y]. \tag{3.73}$$

Доказательство. Простая проверка. □

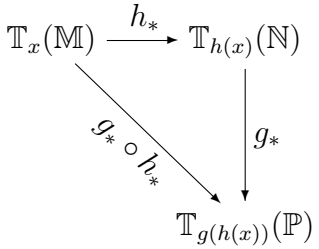
Рассмотрим два отображения $\mathbb{M} \xrightarrow{h} \mathbb{N} \xrightarrow{g} \mathbb{P}$. Если обозначить координаты на многообразиях $\mathbb{M}$, $\mathbb{N}$ и $\mathbb{P}$ соответственно через x, y и z, то по правилу дифференцирования сложной функции в соответствующих областях определения справедливо равенство

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y}\frac{\partial y}{\partial x}.$$

Здесь, для краткости, мы опустили индексы. Отсюда следует, что дифференциал произведения равен произведению дифференциалов каждого отображения

$$(g \circ h)_* = g_* \circ h_*. \tag{3.74}$$

В координатах мы имеем обычное произведение матриц Якоби. Это уравнение говорит о том, что прямое отображение $g \circ h :\ \mathbb{M} \to \mathbb{P}$ не зависит от выбора промежуточного многообразия $\mathbb{N}$: $\mathbb{M} \to \mathbb{N} \to \mathbb{P}$, т.е. диаграмма



коммутативна для всех точек $x \in \mathbb{M}$.

Замечание. Поскольку сами многообразия в общем случае не являются векторными пространствами, то дифференциал отображения не имеет смысла производной по Фреше отображения h. □

Дифференциал отображения естественным образом обобщается на случай произвольных тензорных полей типа $(r,0)$, имеющих только контравариантные индексы, и обозначается $(h_*)^r$. При этом каждый контравариантный индекс суммируется с матрицей Якоби отображения.

Пример 3.9.1. Для компонент контравариантных тензоров второго ранга имеем следующий закон преобразования

$$Y^{\mu\nu} = \frac{\partial y^\mu}{\partial x^\alpha}\frac{\partial y^\nu}{\partial x^\beta}X^{\alpha\beta},$$

где $X = X^{\alpha\beta}(x)\partial_\alpha \otimes \partial_\beta \in \mathcal{T}^2(\mathbb{M})$ и $Y = Y^{\mu\nu}(y)\partial_\mu \otimes \partial_\nu \in \mathcal{T}^2(\mathbb{N})$ □

Замечание. Если векторное поле рассматривается, как дифференцирование в алгебре функций, то дифференциал отображения определяется следующей формулой

$$(h_*X)f = X(f \circ h), \quad \forall\ f \in \mathcal{C}^1(\mathbb{N}),\ X \in \mathcal{X}(\mathbb{M}). \tag{3.75}$$

При таком определении дифференциала отображения, его линейность очевидна. Это еще раз говорит о том, что алгебраические определения удобнее в тех случаях, когда исследуются алгебраические свойства. □

Определение. Инъективное отображение многообразий $h : \mathbb{M} \to \mathbb{N}$ индуцирует *возврат отображения* в кокасательных пространствах, который мы обозначим тем же символом, но со звездочкой вверху:

$$h^* : \quad \mathbb{T}^*_{h(x)}(\mathbb{N}) \ni B \ \mapsto A \in \mathbb{T}^*_x(\mathbb{M}). \tag{3.76}$$

Возврат отображения h^* дуален к дифференциалу отображения h_* и определяется следующим равенством

$$A(X) = B(Y) \quad \text{или} \quad h^*B(X) = B(h_*X), \tag{3.77}$$

где $Y = h_*X$. □

Замечание. Возврат отображения действует в сторону, обратную самому отображению h. □

При определении возврата отображения h^* мы требуем, чтобы исходное отображение h было инъективным. В противном случае прообраз $h^{-1}\Big(h(x)\Big)$ для некоторого $x \in \mathbb{M}$ состоит не из одного элемента, и возврат отображения не определен. Возврат отображения определен не на всех формах из $\mathcal{T}_1(\mathbb{N})$, а только на формах из образа инъективного отображения $\mathcal{T}_1\Big(h(\mathbb{M})\Big)$. Возврат отображения h^* будет определен на множестве всех форм $\mathcal{T}_1(\mathbb{N})$ тогда и только тогда, когда отображение h биективно.

В компонентах возврат отображения записывается в виде

$$A_\alpha = \frac{\partial y^\mu}{\partial x^\alpha} B_\mu, \tag{3.78}$$

т.е. так же, как и дифференциал отображения, определяется матрицей Якоби.

Возврат отображения естественным образом обобщается на тензорные поля из $\mathcal{T}_s(\mathbb{N})$ типа $(0, s)$, имеющие только ковариантные индексы, и обозначается $(h^*)^s$.

Пример 3.9.2. Для компонент ковариантных тензоров второго ранга имеем следующий закон преобразования

$$X_{\alpha\beta} = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial y^\nu}{\partial x^\beta} Y_{\mu\nu},$$

где $X = X_{\alpha\beta}(x)dx^\alpha \otimes dx^\beta \in \mathcal{T}_2(\mathbb{M})$ и $Y = Y_{\mu\nu}(y)dy^\mu \otimes dy^\nu \in \mathcal{T}_2(\mathbb{N})$ □

Из правила дифференцирования сложных функций следует, что возврат произведения отображений равен произведению возвратов отображений, но взятых в обратном порядке

$$(g \circ f)^* = f^* \circ g^*.$$

Для тензоров смешанного типа в общем случае не существует индуцированного отображения, поскольку дифференциал и возврат отображения действуют в разные стороны.

Определение. Назовем *рангом отображения* $h : \mathbb{M} \to \mathbb{N}$ ранг соответствующей матрицы Якоби

$$\mathsf{rank}\, h = \mathsf{rank}\, \frac{\partial y^\mu}{\partial x^\alpha}.$$

Если $\mathsf{rank}\, h = m$ во всех точках многообразия $\mathbb{M}$, то отображение h называется *невырожденным*. Для этого необходимо, чтобы $m \leq n$.

Рассмотрим отображение (3.69) двух многообразий одинаковой размерности $\mathsf{dim}\, \mathbb{M} = \mathsf{dim}\, \mathbb{N}$ и одного класса дифференцируемости $\mathcal{C}^k$. Если отображение h биективно, и оба отображения h и h^{-1} в координатах задаются функциями класса $\mathcal{C}^k$, то отображение h называется *диффеоморфизмом*. □

Пример 3.9.3. Пусть $(\mathbb{U}, \varphi)$ карта на дифференцируемом многообразии $\mathbb{M}$. Тогда отображение $\varphi : \mathbb{U} \to \varphi(\mathbb{U}) \subset \mathbb{R}^n$ является диффеоморфизмом. □

Пример 3.9.4. Преобразование координат $\mathbb{U}_0 \to \mathbb{U}'_0$ в теореме 2.5.1 является диффеоморфизмом. □

Композиция двух диффеоморфизмов снова будет диффеоморфизмом. Таким образом, диффеоморфизм является отношением эквивалентности в категории дифференцируемых многообразий. С точки зрения дифференциальной геометрии два диффеоморфных между собой многообразия можно рассматривать, как одно многообразие, заданное в различных координатах, поэтому мы пишем $\mathbb{M} \approx \mathbb{N}$[1]. В этом случае индуцированные отображения касательных пространств представляют собой не что иное, как правила преобразования тензорных полей при преобразовании координат.

Замечание. Не имеет смысла рассматривать степень гладкости отображения, которая превышает гладкость самих многообразий. □

Замечание. Любой диффеоморфизм представляет собой гомеоморфизм многообразий, рассматриваемых, как топологические пространства, т.к. функции h и h^{-1} непрерывны. Обратное утверждение неверно. Как отмечено в разделе 3.1, на семимерной сфере можно задать несколько различных дифференцируемых структур. То есть гомеоморфные многообразия могут быть недиффеоморфны. □

Множество невырожденных гладких отображений $h : \mathbb{M} \to \mathbb{M}$ многообразия $\mathbb{M}$ в себя образует группу преобразований многообразия, которая обозначается $\mathsf{diff}\,(\mathbb{M})$. Если многообразие тривиально (т.е. покрывается одной картой), то это просто группа преобразований координат $\mathsf{diff}\,(\mathbb{R}^n)$.

Определение. Пусть для отображения (3.69) $m < n$. Если отображение касательного пространства $\mathbb{T}_x(\mathbb{M})$ для всех точек $x \in \mathbb{M}$ на его образ в касательном пространстве $\mathbb{T}_{h(x)}(\mathbb{N})$ является взаимно однозначным, т.е. $\mathsf{rank}\, h = \mathsf{dim}\, \mathbb{M}$, то отображение h называется *погружением*. По-определению, любое невырожденное отображение задает погружение. При погружении само отображение h может не быть взаимно однозначным. Размерность многообразия $\mathbb{M}$ не может превосходить размерности многообразия $\mathbb{N}$, т.к. в этом случае не может быть взаимной однозначности дифференциала отображения. Если само отображение h на его образ $h(\mathbb{M})$ и его дифференциал являются взаимно однозначными, то отображение h называется *вложением*. В дальнейшем вложение многообразий мы будем обозначать специальным символом: $\mathbb{M} \hookrightarrow \mathbb{N}$. □

Конечно, каждое вложение одновременно является и погружением.

Пример 3.9.5. Отображение окружности $\mathbb{S}^1$ на плоскость $\mathbb{R}^2$ в виде восьмерки является погружением, но не вложением, см. рис. 3.8,*a*. В то же время отображение окружности в гладкую замкнутую кривую без самопересечений представляет собой вложение, рис. 3.8,*b*. □

Если $m = n$, отображение h биективно и дифференцируемо, а образ $h(\mathbb{M})$ совпадает со всем $\mathbb{N}$, то точки многообразий $\mathbb{M}$ и $\mathbb{N}$ можно отождествить и рассматривать вложение $\mathbb{M} \overset{h}{\hookrightarrow} \mathbb{N} \approx \mathbb{M}$ как диффеоморфизм. Нетривиальная ситуация может возникнуть, если многообразие $\mathbb{M}$ отображается только на часть $\mathbb{N}$. Тогда мы имеем диффеоморфизм между многообразием $\mathbb{M}$ и его образом $h(\mathbb{M})$.

[1]Мы не используем обычный знак равенства, т.к. множества точек $\mathbb{M}$ и $\mathbb{N}$ могут отличаться по другим признакам, например, по наличию групповой структуры.

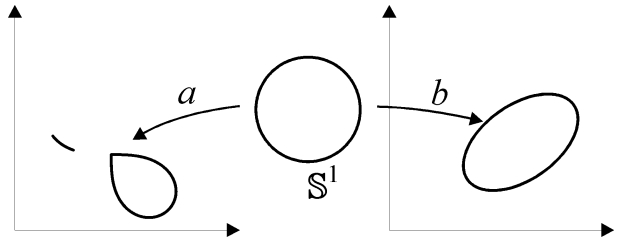


Рис. 3.8: Отображение окружности на плоскость в виде восьмерки является погружением, но не вложением, *a*. Вложение окружности в плоскость в виде гладкой замкнутой кривой без самопересечений, *b*.

3.10 Подмногообразия

Важным классом отображений многообразий являются подмногообразия.

Определение. Пусть f – вложение многообразия $\mathbb{M}$ в $\mathbb{N}$, размерностей m и n, при этом $m \leq n$, тогда пара $(f, \mathbb{M})$ называется *подмногообразием* в $\mathbb{N}$.

В дальнейшем, если не оговорено противное, под вложением мы понимаем гладкое вложение, когда отображение f определяется гладкими функциями.

Пример 3.10.1 (Лемниската Бернулли). Кривая четвертого порядка, заданная в декартовых координатах на евклидовой плоскости $\mathbb{R}^2$ уравнением

$$(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0, \qquad a \in \mathbb{R}, \tag{3.79}$$

называется *лемнискатой Бернулли*, рис.3.9. Эта кривая обладает следующим свойством. Произведение расстояний от произвольной точки лемнискаты до двух заданных точек (фокусов) постоянно и равно квадрату половины расстояния между фокусами. Верно и обратное утверждение. На плоскости можно выбрать такую систему декартовых координат, что произвольная кривая с данным свойством будет задана уравнением (3.79).

В полярных координатах уравнение лемнискаты имеет вид:

$$r^2 = 2a^2 \cos(2\varphi).$$

Лемниската Бернулли не является одномерным многообразием, т.к. имеет точку самопересечения – начало координат. Точка самопересечения имеет касательные $y = \pm x$ и является точкой перегиба. Площадь каждой петли равна a^2.

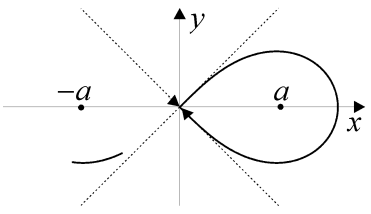


Рис. 3.9: Лемниската Бернулли.

Рассмотрим вложение прямой в евклидову плоскость

$$\gamma : \ \mathbb{R} \ni t \ \mapsto \ (x, y) \in \mathbb{R}^2,$$

гладко отобразив прямую в точки лемнискаты, как показано на рис.3.9. При этом мы считаем, что точка $t = 0$ отображается в начало координат $(0, 0) \in \mathbb{R}^2$. Стрелки показывают, что концы прямой $t = \pm\infty$ при вложении стремятся к началу координат $(0, 0) \in \mathbb{R}^2$. Это – гладкое вложение. Однако топология прямой не совпадает с топологией, которая индуцирована вложением. Действительно, любая последовательность $\{t_k\}$ такая, что $\lim_{k\to\infty} x_k = \pm\infty$ в естественной топологии на прямой, будет сходится к точке $t = 0$ в индуцированной топологии. Конечно, вместо лемнискаты Бернулли можно было бы выбрать произвольную гладкую кривую в виде восьмерки. □

Пример 3.10.2 (Всюду плотная обмотка тора). Двумерный тор $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ можно рассматривать, как двумерное многообразие, полученное путем отождествления противоположных сторон единичного квадрата на евклидовой плоскости. Тем самым точка тора задается упорядоченной парой чисел (x, y), каждое из которых определено по модулю один, $x \sim x + 1$ и $\sim y + 1$. Зафиксируем пару чисел a и $b \neq 0$ и рассмотрим отображение вещественной прямой в тор

$$f : \quad \mathbb{R} \ni t \mapsto (at \bmod 1, bt \bmod 1) \in \mathbb{T}^2.$$

Если отношение a/b иррационально, то отображение f является вложением, причем образ прямой $f(\mathbb{R})$ всюду плотен в $\mathbb{T}^2$ (*всюду плотная обмотка тора*). Если отношение a/b рационально, то f представляет собой погружение, и его образ диффеоморфен окружности. Согласно данному выше определению всюду плотная обмотка тора является одномерным подмногообразием на торе. Для всюду плотной обмотки тора индуцированная топология на образе прямой не совпадает с естественной топологией на $\mathbb{R}$. □

При вложении дифференцируемая структура и топология, заданные на $\mathbb{M}$, естественным образом переносятся на образ $f(\mathbb{M})$, поскольку отображение f инъективно и дифференцируемо. Тем самым многообразие $\mathbb{M}$ диффеоморфно своему образу $f(\mathbb{M})$. С другой стороны, поскольку образ $f(\mathbb{M})$ является некоторым подмножеством в пространстве-мишени $\mathbb{N}$, то на нем определяется индуцированная из $\mathbb{N}$ топология. Рассмотренные выше примеры вложения прямой в плоскость и всюду плотной обмотки тора показывает, что топология на $f(\mathbb{M})$, наследуемая из $\mathbb{M}$ при вложении, совсем не обязательно совпадает с топологией, индуцированной из $\mathbb{N}$. В общем случае топология из $\mathbb{M}$ является более тонкой, чем топология, индуцированная из $\mathbb{N}$. Это наблюдение мотивирует следующее

Определение. Пусть вложение $f : \mathbb{M} \hookrightarrow \mathbb{N}$ является также и гомеоморфизмом. Тогда пара $(f, \mathbb{M})$ называется *регулярным подмногообразием*, а f – *регулярным вложением* $\mathbb{M}$ в $\mathbb{N}$. □

Доказательство следующих трех теорем приведено в [?].

Теорема 3.10.1. *Пусть $(f, \mathbb{M})$ является m-мерным подмногообразием в n-мерном многообразии $\mathbb{N}$, причем $m < n$. Для того, чтобы пара $(f, \mathbb{M})$ была регулярным подмногообразием в $\mathbb{N}$, необходимо и достаточно, чтобы оно было замкнутым подмногообразием некоторого открытого подмногообразия в $\mathbb{N}$.*

Теорема 3.10.2. *Пара $(f, \mathbb{M})$ является регулярным подмногообразием в $\mathbb{N}$ тогда и только тогда, когда для каждой точки $x \in \mathbb{M}$ существует система координат y^α на $\mathbb{U}_{f(x)} \subset \mathbb{N}$, $y^\alpha\big(f(x)\big) = 0$, такая, что пересечение $f(\mathbb{M}) \cap \mathbb{U}_{f(x)}$ определяется системой, состоящей из $n - m$ уравнений:*

$$y^{m+1} = y^{m+2} = \ldots = y^n = 0.$$

Теорема 3.10.3. *Пусть пара* $(f, \mathbb{M})$ *является подмногообразием в* $\mathbb{N}$. *Если многообразие* $\mathbb{M}$ *компактно, тогда вложение* $f: \mathbb{M} \to \mathbb{N}$ *регулярно.*

В дальнейшем под вложением мы будем понимать регулярное вложение, для краткости не оговаривая этого явно. В этом случае мы будем отождествлять подмногообразие $\mathbb{M}$ и его образ $f(\mathbb{M})$ и будем писать $\mathbb{M} \subset \mathbb{N}$.

Определение. Пусть $\mathbb{U}$ является открытым подмножеством $\mathbb{N}$. Если ограничить гладкую структуру $\mathbb{N}$ на $\mathbb{U}$, то $\mathbb{U}$ становится многообразием той же размерности $\dim \mathbb{U} = \dim \mathbb{N}$. Тогда вложение $f: \mathbb{U} \hookrightarrow \mathbb{N}$ осуществляется тождественным отображением, а подмногообразие $\mathbb{U} \subset \mathbb{N}$ называется *открытым подмногообразием* в $\mathbb{N}$.

Пусть $f(\mathbb{M})$ является подмногообразием в $\mathbb{N}$ таким, что выполнены два условия:
1) $f(M)$ – замкнутое подмножество в $\mathbb{N}$;
2) для каждой точки $y \in f(\mathbb{M})$ существует такая система координат y^α на $\mathbb{U}_y \subset \mathbb{N}$, что пересечение $f(\mathbb{M}) \cap \mathbb{U}_y$ задается уравнениями $y^{m+1} = y^{m+2} = \ldots = y^n = 0$, где $m = \dim \mathbb{M}$.

Тогда подмногообразие $f(\mathbb{M})$ называется *замкнутым подмногообразием* в $\mathbb{N}$. Размерность замкнутого подмногообразия $f(\mathbb{M})$ всегда меньше размерности $\mathbb{N}$. □

Замечание. Если подмногообразие (без края) $\mathbb{M} \subset \mathbb{N}$ имеет ту же размерность, что и $\mathbb{N}$, то оно будет обязательно открытым подмногообразием, т.к. в противном случае оно имело бы край. Теорема 3.10.2 утверждает, что любое подмногообразие $\mathbb{M} \subset \mathbb{N}$ меньшей размерности $1 < m < n$ является замкнутым подмногообразием. □

Мы видим, что подмногообразия меньшей размерности $1 < m < n$ в евклидовом пространстве $\mathbb{M} \subset \mathbb{R}^n$ можно задавать с помощью системы алгебраических уравнений на координаты, покрывающие все $\mathbb{R}^n$. Вернее, алгебраические уравнения задают только множества точек. Чтобы превратить это множество в многообразие мы предполагаем, что топология на $\mathbb{M}$ задается вложением $f: \mathbb{M} \hookrightarrow \mathbb{R}^n$. Тогда подмногообразие определяется системой уравнений по крайней мере локально. Обратное утверждение неверно. Не каждая система алгебраических уравнений на координаты в евклидовом пространстве $\mathbb{R}^n$ определяет некоторое подмногообразие.

Пример 3.10.3. Рассмотрим уравнение $x^2 = y^2$ на евклидовой плоскости $\mathbb{R}^2$. Оно определяет две пересекающиеся прямые, которые не являются многообразием. □

Приведем критерий, который часто используется в приложениях.

Теорема 3.10.4. *Для того, чтобы подмножество* $\mathbb{M} \subset \mathbb{R}^n$, $\dim \mathbb{M} = m \leq n$, *было подмногообразием класса* $\mathcal{C}^k$, *необходимо и достаточно, чтобы для каждой точки* $x \in \mathbb{M}$ *существовала окрестность* $\mathbb{U}$ *этой точки и система* $n-m$ *функций* $f^{\text{м}}(x) \in \mathcal{C}^k(\mathbb{U})$, $\text{м} = m+1, \ldots, n$, *определенных на* $\mathbb{U}$, *и обладающих следующими свойствами:*
1) функции $f^{\text{м}}$ *– функционально независимы;*
2) пересечение $\mathbb{U} \bigcap \mathbb{M}$ *в точности определяется системой уравнений*

$$f^{\text{м}} = 0, \qquad \text{м} = m+1, \ldots, n. \tag{3.80}$$

Доказательство. См., например, [**?**]. □

Заметим, что если $m = n$, то уравнений не будет, а $\mathbb{M}$ будет некоторым открытым подмножеством в $\mathbb{M}$. При $m = 0$ число уравнений равно размерности многообразия.

Тогда, если в точке $x \in \mathbb{R}^n$ система уравнений (3.80) удовлетворена, то в некоторой окрестности точки x других решений кроме самой точки не будет. В этом случае подмногообразие $\mathbb{M}$ имеет размерность 0 и представляет собой множество корней системы уравнений $f^{\text{м}} = 0$. Подмногообразие $(f, \mathbb{M})$ размерности $n-1$ в $\mathbb{N}$ называется *гиперповерхностью* в $\mathbb{N}$.

Напомним, что функции $f^{\text{м}}$ функционально независимы в области $\mathbb{U}$ тогда и только тогда, когда $n \times m$ матрица, составленная из частных производных

$$(\partial_\alpha f^{\text{м}}) = \begin{pmatrix} \partial_1 f^{m+1} & \partial_1 f^{m+2} & \dots & \partial_1 f^n \\ \partial_2 f^{m+1} & \partial_2 f^{m+2} & \dots & \partial_2 f^n \\ \vdots & \vdots & \vdots & \vdots \\ \partial_n f^{m+1} & \partial_n f^{m+2} & \dots & \partial_n f^n \end{pmatrix}, \tag{3.81}$$

имеет постоянный ранг m во всех точках $\mathbb{U}$.

Замечание. Забегая вперед, переформулируем условие 1) в теореме 3.10.4 в терминах 1-форм. По-определению, 1-формы $df^{\text{м}} = dx^\alpha \partial_\alpha f^{\text{м}}$, заданные на области $\mathbb{U}$, называются *линейно независимыми* в области $\mathbb{U}$, если в каждой точке этой области ранг матрицы (3.81) равен m. Поэтому условие функциональной независимости функций можно заменить условием линейной независимости 1-форм $df^{\text{м}}$. □

Если подмногообразие $\mathbb{M} \subset \mathbb{R}^n$ задано с помощью $n - m$ алгебраических уравнений (3.80), то множество функций $f^{\text{м}}$, в окрестности произвольной точки $x \in \mathbb{M}$ всегда можно дополнить m функциями $g^1(x), \dots, g^m(x)$ такими, что весь набор функций $g^1, \dots, g^m, f^{m+1}, \dots, f^n$ будет функционально независим. В этом случае можно перейти в новую систему координат

$$\begin{gathered} \{y^\alpha\} = \{y^{\text{а}}, y^{\text{м}}\} = \{g^{\text{а}}, f^{\text{м}}\} = (g^1, \dots, g^n, f^{m+1}, \dots, f^n), \\ {\text{а}} = 1, \dots, m, \quad {\text{м}} = m+1, \dots, n, \end{gathered} \tag{3.82}$$

в которой подмногообразие $\mathbb{M}$ будет задаваться особенно просто $f^{\text{м}} = 0$. При этом функции $y^{\text{а}} = g^{\text{а}}$ образуют систему координат на подмногообразии $\mathbb{M} \subset \mathbb{R}^n$. Конечно, выбор координат $g^{\text{а}}$ на $\mathbb{M}$ является неоднозначным. Отметим, что $\{y^\alpha\}$ – это те координаты, которые фигурируют в теореме 3.10.2.

Пример 3.10.4. Двумерная сфера $\mathbb{S}^2_r$ радиуса r с центром в начале координат является замкнутым подмногообразием трехмерного евклидова пространства $\mathbb{R}^3$, которое задается одним уравнением (25.1). Она является гиперповерхностью в $\mathbb{R}^3$. В сферической системе координат уравнение сферы задается уравнением $r = \mathsf{const}$, а угловые координаты образуют систему координат на сфере. □

Функции $f^{\text{м}}(x)$ определяют $n - m$ точных 1-форм. В силу их линейной независимости, в каждой точке $x \in \mathbb{M}$ они определяют $(n - m)$-мерное подпространство в кокасательном пространстве $\mathbb{T}^*_x(\mathbb{R}^n)$. Ортогональное дополнение этого подпространства является касательным к подмногообразию $\mathbb{M}$ и образовано всеми векторами $X^\alpha \partial_\alpha \in \mathbb{T}_x(\mathbb{R}^n)$, для которых выполнено условие $X^\alpha \partial_\alpha f^{\text{м}} = 0$. Это подпространство касательного пространства имеет размерность m, такую же, как и само подмногообразие. Нетрудно проверить, что, если два векторных поля X_1 и X_2 касательны к подмногообразию $\mathbb{M}$, то их коммутатор $[X_1, X_2]$ также касателен.

Ответ на обратный вопрос, в каком случае распределение векторных полей на многообразии определяет касательные подпространства к некоторому подмногообразию, дает теорема Фробениуса (см. следующий раздел).

В координатах $\{g^{\text{A}}, f^{\text{M}}\}$ дополнение касательного пространства к подмногообразию $\mathbb{M}$ до полного касательного пространства к $\mathbb{R}^n$ имеет простой геометрический смысл. Пусть в $\mathbb{R}^n$ задано векторное поле $X = X^\alpha \partial_\alpha$. Тогда в координатах, связанных с подмногообразием, оно будет иметь компоненты

$$X = (X^\alpha \partial_\alpha g^{\text{A}})\partial_{\text{A}} + (X^\alpha \partial_\alpha f^{\text{M}})\partial_{\text{M}}.$$

При этом первое слагаемое лежит в касательном пространстве $\mathbb{T}(\mathbb{M})$.

В дальнейшем нам понадобится следующее

Определение. Пусть $f \in \mathcal{C}^\infty(\mathbb{N})$ и $\mathbb{M} \subset \mathbb{N}$ – подмногообразие, тогда отображение

$$f|_{\mathbb{M}} : \quad \mathbb{M} \to \mathbb{R}.$$

называется *сужением* функции на подмногообразие. Функция f является *продолжением* функции g с некоторого подмногообразия $\mathbb{M}$ на все $\mathbb{N}$, если ее сужение на $\mathbb{M}$ совпадает с g, $f|_{\mathbb{M}} = g$. Аналогично, сужением произвольного тензорного поля $T \in \mathcal{T}(\mathbb{N})$ на подмногообразие называется тензорное поле $T|_{\mathbb{M}}$, рассматриваемое только в точках $x \in \mathbb{M}$. Обратно, тензорное поле $T \in \mathcal{T}(\mathbb{N})$ является продолжением тензорного поля $K \in \mathcal{T}(\mathbb{N})$, заданного на подмногообразии $\mathbb{M}$, если его сужение совпадает с K, $T|_{\mathbb{M}} = K$. □

Определение. Два подмногообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ многообразия $\mathbb{N}$ называются *трансверсальными* в точке $x \in \mathbb{M}_1 \cap \mathbb{M}_2$, если касательные подпространства $\mathbb{T}_x(\mathbb{M}_1)$ и $\mathbb{T}_x(\mathbb{M}_2)$ порождают все касательное пространство $\mathbb{T}_x(\mathbb{N})$, т.е. касательные подпространства $\mathbb{T}_x(\mathbb{M}_1)$ и $\mathbb{T}_x(\mathbb{M}_2)$ трансверсальны. □

Трансверсальность подмногообразий означает, что в некоторой окрестности $\mathbb{U}$ точки $x \in \mathbb{M}_1 \cap \mathbb{M}_2$ существует такая система координат x^α, $\alpha = 1, \ldots, n$, что подмногообразия задаются условиями:

$$\mathbb{M}_1 \cap \mathbb{U} = \{x \in \mathbb{U} : \quad x^{\dim \mathbb{M}_1 + 1} = 0, \ldots, x^n = 0\},$$
$$\mathbb{M}_2 \cap \mathbb{U} = \{x \in \mathbb{U} : \quad x^1 = 0, \ldots, x^{n - \dim \mathbb{M}_2} = 0\}.$$

Пример 3.10.5. Пусть $\mathbb{N} := \mathbb{M}_1 \times \mathbb{M}_2$. Тогда подмногообразия $(\mathbb{M}_1, x_2) \subset \mathbb{N}$, где $x_2 \in \mathbb{M}_2$, и $(x_1, \mathbb{M}_2) \subset \mathbb{N}$, где $x_1 \in \mathbb{M}_1$, трансверсальны в точке $x = (x_1, x_2)$. Если в окрестностях $\mathbb{U}_1 \subset \mathbb{M}_1$ и $\mathbb{U}_2 \subset \mathbb{M}_2$, где $x_1 \in \mathbb{U}_1$ и $x_2 \in \mathbb{U}_2$, заданы системы координат, $x_1 = \{x_1^\mu\}$ и $x_2 = \{x_2^{\text{A}}\}$, то подмногообразия задаются уравнениями

$$(\mathbb{U}_1, x_2) = \{x \in \mathbb{U}_1 \times \mathbb{U}_2 : \quad x_2^{\text{A}} = \mathsf{const},\ \forall \text{A}\},$$
$$(x_1, \mathbb{U}_2) = \{x \in \mathbb{U}_1 \times \mathbb{U}_2 : \quad x_1^\mu = \mathsf{const},\ \forall \mu\}.$$

Конечно, системы координат можно выбрать таким образом, что точки x_1 и x_2 будут находиться в началах систем отсчета. □

Определение. Отображение $f : \mathbb{M} \to \mathbb{N}$ *трансверсально* к подмногообразию $\mathbb{L} \subset \mathbb{N}$ в точке $x \in f^{-1}\Big(\mathbb{L} \cap f(\mathbb{M})\Big)$, если образ $f_*\Big(\mathbb{T}_x(\mathbb{M})\Big)$ трансверсален к $\mathbb{T}_{f(x)}(\mathbb{L})$. Два отображения $f_1 : \mathbb{M}_1 \to \mathbb{N}$ и $f_2 : \mathbb{M}_2 \to \mathbb{N}$ *трансверсальны друг к другу* в точке $(x_1, x_2) \in \mathbb{M}_1 \times \mathbb{M}_2$, где $f_1(x_1) = f_2(x_2)$, если образы $f_{1*}\Big(\mathbb{T}_{x_1}(\mathbb{M}_1)\Big)$ и $f_{2*}\Big(\mathbb{T}_{x_2}(\mathbb{M}_2)\Big)$ порождают все касательное пространство $\mathbb{T}_{f_1(x_1)}(\mathbb{N})$. □

Понятия трансверсальных подмногообразий и трансверсальных отображений естественным образом сводятся друг к другу. Если $f_{1,2} : \mathbb{M}_{1,2} \hookrightarrow \mathbb{N}$ – вложения, то трансверсальность вложений эквивалентна трансверсальности подмногообразий $f_1(\mathbb{M}_1) \subset \mathbb{N}$ и $f_2(\mathbb{M}_2) \subset \mathbb{N}$.

3.11 Теорема Фробениуса

В разделе 3.6.5 было показано, что у любого векторного поля X существует интегральная кривая. Если векторное поле нигде не обращается в нуль на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, то интегральная кривая, проходящая через некоторую точку $x \in \mathbb{M}$, представляет собой одномерное подмногообразие, и в окрестности точки x существует такая система координат, что $X = \partial_1$. Поставим более общую задачу. Пусть в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ задано k векторных полей $(X_1, \ldots, X_k) = \{X_\text{а}\}$, $\text{а} = 1, \ldots, k < n$, которые линейно независимы в каждой точке $x \in \mathbb{U}$. Возникает вопрос о том, существуют ли такие подмногообразия в $\mathbb{M}$, касательные пространства к которым в каждой точке $x \in \mathbb{U}$ совпадают с подпространством, натянутым на векторы $\{X_\text{а}\}$? Ответ на этот вопрос дает теорема Фробениуса. Чтобы ее сформулировать введем несколько новых понятий.

Определение. *Распределением* $\mathcal{L}_k(\mathbb{M})$ размерности k на многообразии $\mathbb{M}$ называется сопоставление каждой точке $x \in \mathbb{M}$ k-мерного подпространства в касательном пространстве $\mathbb{L}_x(\mathbb{M}) \subset \mathbb{T}_x(\mathbb{M})$. Распределение называется *дифференцируемым*, если каждая точка $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}$ и k дифференцируемых векторных полей $(X_1, \ldots, X_k)$, которые образуют базис в $\mathbb{L}_x(\mathbb{M})$ для всех $x \in \mathbb{U}$. Множество векторных полей $(X_1, \ldots, X_k)$ называется *локальным базисом* распределения $\mathcal{L}_k(\mathbb{M})$ на $\mathbb{U}$. Векторное поле X принадлежит распределению, если $X(x) \in \mathbb{L}_x(\mathbb{M})$ для всех $x \in \mathbb{M}$. Распределение называется *инволютивным* или *вполне интегрируемым*, если для любых векторных полей из распределения $X, Y \in \mathcal{L}_k(\mathbb{M})$ их коммутатор также принадлежит распределению $[X, Y] \in \mathcal{L}_k(\mathbb{M})$. □

Замечание. Ни одно из векторных полей $X_\text{а}$ не может обращаться в нуль, т.к. в соответствующей точке векторные поля были бы линейно зависимы. □

Замечание. Инволютивность векторных полей, принадлежащих распределению, означает, что эти векторные поля образуют подалгебру Ли алгебры Ли векторных полей $\mathcal{X}(\mathbb{M})$. □

Определение. Связное подмногообразие $f : \mathbb{N} \hookrightarrow \mathbb{M}$ называется *интегральным многообразием* распределения $\mathcal{L}_k(\mathbb{M})$, если $f_*\Big(\mathbb{T}_x(\mathbb{N})\Big) = \mathbb{L}_{f(x)}(\mathbb{M})$, для всех $x \in \mathbb{N}$. Если не существует других интегральных многообразий, содержащих $\mathbb{N}$, то $\mathbb{N}$ называется *максимальным интегральным многообразием* для распределения $\mathcal{L}_k(\mathbb{M})$. □

Замечание. Поскольку интегральное многообразие – это подмногообразие и, следовательно задается парой $(f, \mathbb{N})$, где f – вложение $\mathbb{N}$ в $\mathbb{M}$, то определение максимального интегрального многообразия требует уточнения. А именно, мы говорим, что интегральное многообразие $(f_1, \mathbb{N}_1)$ содержит интегральное многообразие $(f_2, \mathbb{N}_2)$, если $f_2(\mathbb{N}_2)$ есть подмножество в $f_1(\mathbb{N}_1)$: $f_2(\mathbb{N}_2) \subset f_1(\mathbb{N}_1)$. Из определения максимального интегрального многообразия сразу следует его единственность. □

Предложение 3.11.1. *Пусть $\mathcal{L}_k(\mathbb{M})$ – гладкое распределение на многообразии $\mathbb{M}$ такое, что через каждую точку $x \in \mathbb{M}$ проходит интегральное многообразие, тогда распределение $\mathcal{L}_k(\mathbb{M})$ инволютивно.*

Доказательство. Пусть X, Y – два произвольных векторных поля, принадлежащих распределению $\mathcal{L}_k(\mathbb{M})$. Пусть $(f, \mathbb{N})$ – интегральное многообразие, проходящее через точку $f(x) \in \mathbb{M}$, $x \in \mathbb{N}$. Поскольку отображение f в каждой точке интегрального многообразия $x \in \mathbb{N}$ задает изоморфизм $\mathbb{T}_x(\mathbb{N}) \simeq \mathbb{L}_{f(x)}(\mathbb{M})$, то на подмногообразии

$\mathbb{N}$ существуют векторные поля $\tilde{X}(x) = f_*^{-1}X\Big(f(x)\Big)$ и $\tilde{Y} = f_*^{-1}Y\Big(f(x)\Big)$. Поскольку коммутатор касательных векторов к $\mathbb{N}$ также касателен к $\mathbb{N}$, то

$$[X,Y]\Big(f(x)\Big) = f_*[\tilde{X},\tilde{Y}](x) \in \mathbb{L}_x(\mathbb{M}).$$

Что и требовалось доказать. □

Теорема 3.11.1 (Фробениус). *Пусть $\mathcal{L}_k(\mathbb{M})$ – k-мерное гладкое инволютивное распределение на многообразии $\mathbb{M}$. Тогда для каждой точки $x \in \mathbb{M}$ существует интегральное многообразие распределения $\mathcal{L}_k(\mathbb{M})$, проходящее через точку x. Кроме того, существует такая система координат в некоторой окрестности точки x, что базис распределения имеет вид $(\partial_1, \dots, \partial_k)$.*

Доказательство. См., например, [**?**]. □

Замечание. Предложение 3.11.1 дает необходимое условие для существования интегральных многообразий, а теорема 3.11.1 – достаточное. Обе теоремы локальны, т.к. в них говорится об интегральных многообразиях, проходящих через точку многообразия. □

Если распределение $\mathcal{L}_k(\mathbb{U})$ задано набором гладких линейно независимых векторных полей $(X_1, \dots, X_k) = \{X_\text{A}\}$, $\text{A} = 1, \dots, k < n$, в каждой точке $x \in \mathbb{U} \subset \mathbb{M}$, то будем писать $\mathcal{L}_k(\mathbb{U}) = \{X_\text{A}\}$. В этом случае будем говорить, что распределение задано *распределением векторных полей*. Тогда теорему Фробениуса вместе с предложением 3.11.1 можно переформулировать.

Теорема 3.11.2 (Фробениус). *Пусть $\mathcal{L}_k(\mathbb{U}) = \{X_\text{A}\}$, $\text{A} = 1, \dots, k$, является k-мерным гладким распределением векторных полей в области $\mathbb{U} \subset \mathbb{M}$. Тогда для существования такой системы координат в некоторой подобласти $\mathbb{V} \subset \mathbb{U}$, что распределение имеет вид*

$$\mathcal{L}_k(\mathbb{V}) = (\partial_1, \dots, \partial_k) \tag{3.83}$$

необходимо и достаточно, чтобы коммутатор векторных полей также принадлежал распределению,

$$[X_\text{A}, X_\text{B}] = f_{\text{AB}}{}^\text{C} X_\text{C} \in \mathbb{L}_x(\mathbb{M}), \qquad \forall x \in \mathbb{U}, \tag{3.84}$$

где $f_{\text{AB}}{}^\text{C}(x)$ – некоторые гладкие функции на $\mathbb{U}$.

Доказательство. См., например, [**?**]. □

Замечание. Конечно, функции $f_{\text{AB}}{}^\text{C}$ в теореме Фробениуса не могут быть произвольными. Из определения (3.84) следует, что они антисимметричны по нижним индексам и удовлетворяют тождествам, вытекающим из тождеств Якоби для коммутатора векторных полей. □

Допустим, что на многообразии $\mathbb{M}$ существует гладкое распределение $\mathcal{L}_k(\mathbb{M})$. Тогда существование системы координат (3.83) в окрестности каждой точки $x \in \mathbb{M}$ означает следующее. Уравнения

$$x^\alpha = \mathsf{const}, \qquad \alpha = k+1, \dots, n$$

определяют k-мерные подмногообразия, которые являются интегральными многообразиями распределения $\mathcal{L}_k(\mathbb{M})$, причем каждое векторное поле X_A касается одного

из подмногообразий. Интегральные многообразия не пересекаются, т.к. в точке пересечения у нас не было бы подмногообразия. Кроме того, через каждую точку $\mathbb{M}$ проходит одно подмногообразие. Тем самым k-мерное распределение $\mathcal{L}_k(\mathbb{M})$ расслаивает область $\mathbb{V} \subset \mathbb{M}$, представляя ее в виде объединения несчетного числа k-мерных подмногообразий. Очевидно, что интегральная кривая для любого векторного поля из заданного инволютивного распределения, проходящая через точку $x \in \mathbb{M}$ целиком лежит в интегральном подмногообразии, содержащем эту точку.

Пример 3.11.1. Произвольное векторное поле $X \in \mathcal{X}(\mathbb{M})$, не равное нулю ни в одной точке, задает одномерное распределение на $\mathbb{M}$. Это распределение инволютивно. Интегральными многообразиями одномерного распределения являются интегральные кривые векторного поля X. □

Пример 3.11.2. Рассмотрим трехмерное евклидово пространство с выколотым началом $\mathbb{R}^3 \setminus \{0\}$ в сферической системе координат r, θ, φ. Векторные поля ∂_θ и ∂_φ коммутируют и, значит, находятся в инволюции. Интегральными подмногообразиями этого распределения являются сферы $r = \mathsf{const}$ с центром в начале координат. Угловые координаты θ, φ и есть те координаты, которые фигурируют в теореме Фробениуса.

Векторные поля

$$J_1 = \frac{1}{2}(y\partial_z - z\partial_y), \qquad J_2 = \frac{1}{2}(z\partial_x - x\partial_z), \qquad J_3 = \frac{1}{2}(x\partial_y - y\partial_x), \tag{3.85}$$

заданные в декартовых координатах, являются генераторами алгебры $\mathfrak{so}(3)$ и находятся в инволюции. Интегральными подмногообразиями для них также являются сферы. Однако поля (3.85) не задают базис распределения, т.к. линейно зависимы

$$xJ_1 + yJ_2 + zJ_3 = 0.$$ □

Сформулируем глобальный вариант теоремы.

Теорема 3.11.3 (Фробениус). *Пусть $\mathcal{L}_k(\mathbb{M})$ – гладкое инволютивное k-мерное распределение. Тогда через каждую точку $x \in \mathbb{M}$ проходит единственное максимальное k-мерное интегральное многообразие $f : \mathbb{N} \hookrightarrow f\mathbb{M}$ распределения $\mathcal{L}_k(\mathbb{M})$. Любое другое связное интегральное многообразие этого распределения, проходящее через точку x, содержится в $(f, \mathbb{N})$.*

Доказательство. См., например, [**?**]). □

Замечание. В общем случае интегральное многообразие распределения $\mathcal{L}_k(\mathbb{M})$ может не быть регулярным подмногообразием размерности k в $\mathbb{M}$. Примером является иррациональная обмотка тора в примере 3.10.2. □

Эта теорема означает, что, если на всем многообразии $\mathbb{M}$ существует гладкое инволютивное k-мерное распределение, то оно представляет собой слоение, т.е. объединение несчетного числа k-мерных максимальных интегральных многообразий (листов) данного распределения (см. следующий раздел).

Переформулируем теорему Фробениуса на языке дифференциальных форм. Дополним поля $\{X_{\text{а}}\} = (X_1, \ldots, X_k)$ в окрестности $\mathbb{U} \subset \mathbb{M}$ $n-k$ векторными полями $\{X_{\text{м}}\} = (X_{k+1}, \ldots, X_n)$ так, чтобы вся совокупность векторных полей $\{X_{\text{а}}, X_{\text{м}}\} = (X_1, \ldots, X_n)$ была линейно независима в каждой точке $x \in \mathbb{U}$. Пусть $\{A^{\text{а}}, A^{\text{м}}\} =$

$(A^1,\dots,A^n)$ – соответствующие дуальные 1-формы. Тогда в каждой точке $x\in\mathbb{U}$ 1-формы $A^{\text{м}}$ определяют ортогональное дополнение $(\mathcal{L}_k)^\perp$. Это значит, что локально задание распределения векторных полей эквивалентно нахождению решения системы $n-k$ уравнений на дифференциалы dx^α,

$$A^{\text{м}}=dx^\alpha A_\alpha{}^{\text{м}}=0,\qquad \text{м}=k+1,\dots,n. \tag{3.86}$$

Эта система называется *пфаффовой системой уравнений*. Из тождества (4.36) следует

$$2dA^{\text{м}}(X_{\text{а}},X_{\text{в}})=X_{\text{а}}\Big(A^{\text{м}}(X_{\text{в}})\Big)-X_{\text{в}}\Big(A^{\text{м}}(X_{\text{а}})\Big)-A^{\text{м}}\Big([X_{\text{а}},X_{\text{в}}]\Big)=0, \tag{3.87}$$

т.к. распределение векторных полей удовлетворяет условию теоремы Фробениуса (3.84). Это значит, что распределение векторных полей $\mathcal{L}_k$ удовлетворяет условию теоремы Фробениуса тогда и только тогда, когда

$$dA^{\text{м}}(X_{\text{а}},X_{\text{в}})=0,\qquad \text{а},\text{в}=1,\dots,k,\quad \text{м}=k+1,\dots,n.$$

Для любой 2-формы, в том числе и для $dA^{\text{м}}$, справедливо представление

$$dA^{\text{м}}=B^{\text{м}}{}_{\text{н}}\wedge A^{\text{н}}+g^{\text{м}}{}_{\text{ва}}A^{\text{а}}\wedge A^{\text{в}},$$

где $\{B^{\text{м}}{}_{\text{н}}\}$ и $\{g^{\text{м}}{}_{\text{ва}}\}$ – некоторые наборы, соответственно, 1-форм и функций. Поскольку $A^{\text{а}}(X_{\text{в}})=\delta^{\text{а}}_{\text{в}}$, то из равенства (3.87) следует, что $g^{\text{м}}{}_{\text{ва}}=0$. Таким образом, справедливо разложение

$$dA^{\text{м}}=B^{\text{м}}{}_{\text{н}}\wedge A^{\text{н}},$$

которое эквивалентно условию коммутативности векторных полей (3.84).

Определение. Система уравнений Пфаффа называется *вполне интегрируемой*, если существует такая система координат $y^\alpha(x)$, что подмногообразия, определяемые системой уравнений

$$y^{\text{м}}=\mathsf{const}\quad\Leftrightarrow\quad dy^{\text{м}}=dx^\alpha\partial_\alpha y^{\text{м}}=0,\qquad \text{м}=k+1,\dots,n, \tag{3.88}$$

удовлетворяют системе уравнений Пфаффа (3.86). □

Для вполне интегрируемой системы уравнений Пфаффа распределение векторных полей $\mathcal{L}_k$ задается в точности первыми k координатными полями $\partial_1,\dots,\partial_k$. Обратное утверждение также справедливо. Таким образом теорему Фробениуса можно переформулировать полностью в терминах 1-форм.

Теорема 3.11.4 (Фробениус). *Для того, чтобы система уравнений Пфаффа*

$$A^{\text{м}}=0,\qquad \text{м}=k+1,\dots,n, \tag{3.89}$$

была вполне интегрируемой, необходимо и достаточно, чтобы существовал такой набор 1-форм $B^{\text{м}}{}_{\text{н}}$, что

$$dA^{\text{м}}=B^{\text{м}}{}_{\text{н}}\wedge A^{\text{н}}. \tag{3.90}$$

Следствие. Если все 1-формы $A^{\text{м}}$ замкнуты, $dA^{\text{м}}=0$, то система уравнений Пфаффа (3.89) вполне интегрируема. □

Пример 3.11.3. Рассмотрим уравнение Пфаффа на 1-форму заданную на евклидовой плоскости $\mathbb{R}^2$ в декартовой системе координат

$$A = dxP + dyQ = 0, \tag{3.91}$$

где $P(x,y), Q(x,y) \in \mathcal{C}^1(\mathbb{R}^2)$. Для определенности будем считать, что $P \neq 0$. Тогда 1-формы

$$dxP + dyQ \quad \text{и} \quad dy$$

можно выбрать в качестве базиса 1-форм на плоскости. Соответствующий ему дуальный базис векторных полей имеет вид

$$X_1 = \left(\frac{1}{P}, 0\right), \quad X_2 = \left(-\frac{Q}{P}, 1\right).$$

Векторное поле X_2 определяет ортогональное дополнение к A и задает уравнения, определяющие интегральные кривые в параметрическом виде

$$\begin{aligned} \dot{x} &= -\frac{Q}{P}, \\ \dot{y} &= 1. \end{aligned}$$

Отсюда следует уравнение на форму интегральной кривой

$$\frac{dx}{dy} = -\frac{Q}{P},$$

которое эквивалентно уравнению Пфаффа (3.91). Это уравнение для достаточно гладких функций P и Q всегда интегрируемо. Нетрудно проверить, что в двумерном случае 1-формы B в условии теоремы Фробениуса (3.90) всегда существуют.

Отметим, что интегрируемость уравнения Пфаффа (3.91) совсем не означает, что 1-форма A является точной, т.е. $A = df$ для некоторой функции $f \in \mathcal{C}^\infty(\mathbb{R}^2)$. Критерием точности 1-формы на плоскости являются нетривиальные условия интегрируемости $\partial_x Q = \partial_y P$, которые не имеют отношения к интегрируемости уравнений Пфаффа. Полная интегрируемость уравнений Пфаффа $A = 0$ означает существование такой функции $f \in \mathcal{C}^\infty(\mathbb{R}^2)$, что $df = 0$ для любой траектории, определяемой уравнением Пфаффа (3.91), а не то, что $A = df$. □

Пример 3.11.4. Покажем, что в общем случае, когда распределение не является инволютивным, интегральные многообразия отсутствуют. Пусть в трехмерном евклидовом пространстве $\mathbb{R}^3$ задана 1-форма

$$A = dx + ydz.$$

Ее внешняя производная равна

$$dA = dy \wedge dz.$$

Поскольку

$$dA \wedge A = dx \wedge dy \wedge dz \neq 0,$$

то условие теоремы Фробениуса не выполняются, которое в трехмерном случае имеет вид (3.93). Это значит, что уравнение Пфаффа $A = 0$ в данном случае неинтегрируемо.

Посмотрим на эту задачу с точки зрения векторных полей. Ортогональное дополнение к форме A натянуто, например, на векторные поля

$$X = -y\partial_x + \partial_z \quad \text{и} \quad Y = \partial_y,$$

поскольку они линейно независимы и

$$A(X) = 0, \quad A(Y) = 0.$$

Эти векторные поля не находятся в инволюции, т.к. их коммутатор нельзя представить в виде линейной комбинации

$$[X, Y] = \partial_x \neq fX + gY, \quad f, g \in \mathcal{C}^\infty(\mathbb{R}^3).$$

Это значит, что условие теоремы Фробениуса нарушено.

Докажем от противного, что у векторных полей X, Y интегральных поверхностей не существует. Допустим, что через начало координат проходит интегральная поверхность. Это значит, что, двигая начало координат вдоль интегральных кривых, задаваемых векторными полями, мы никогда не покинем эту поверхность. Векторные поля X, Y задают потоки $s_X(t), s_Y(t)$, проходящие через точку (x_0, y_0, z_0) (см. раздел 3.6.5),

$$\begin{aligned} s_X: &\quad t \times (x_0, y_0, z_0) \to (-y_0 t + x_0, y_0, t + z_0), \\ s_Y: &\quad t \times (x_0, y_0, z_0) \to (x_0, t + y_0, z_0). \end{aligned}$$

Коммутатору векторных полей $[Y, X]$ соответствует отображение

$$s_Y^{-1}(t_2) s_X^{-1}(t_1) s_Y(t_2) s_X(t_1)$$

При этом начало координат перейдет в точку

$$(0,0,0) \xrightarrow{s_X(t_1)} (0,0,t_1) \xrightarrow{s_Y(t_2)} (0,t_2,t_1) \xrightarrow{s_X^{-1}(t_1)} (t_1 t_2, t_2, 0) \xrightarrow{s_Y^{-1}(t_2)} (t_1 t_2, 0, 0).$$

Ясно, что не существует такой линейной комбинации векторных полей $fX + gY$, чей поток переводил бы начало координат в точку $(t_1 t_2, 0, 0)$. Тем самым мы пришли к противоречию, что и доказывает отсутствие интегральной поверхности, проходящей через начало координат. Аналогичное построение можно провести для произвольной точки $\mathbb{R}^3$. По сути дела доказательство теоремы Фробениуса сводится к такому же построению в общем случае. □

Пример 3.11.5. Рассмотрим уравнение Пфаффа на 1-форму в трехмерном евклидовом пространстве $\mathbb{R}^3$ в декартовой системе координат

$$A = dxP + dyQ + dzR = 0, \tag{3.92}$$

где $P, Q, R \in \mathcal{C}^1(\mathbb{R}^3)$. Полная интегрируемость этого уравнения означает существование функции $f \in \mathcal{C}^2(\mathbb{R}^3)$ такой, что равенство $f = \mathsf{const}$ является первым интегралом уравнения Пфаффа (3.92). Из теоремы Фробениуса следует критерий интегрируемости, который можно (только для трехмерного многообразия) записать в виде

$$dA = B \wedge A \quad \Leftrightarrow \quad dA \wedge A = 0. \tag{3.93}$$

Поскольку

$$dA = dx \wedge dy(\partial_x Q - \partial_y P) + dy \wedge dz(\partial_y R - \partial_z Q) + dz \wedge dx(\partial_z P - \partial_x R),$$

то необходимым и достаточным условием интегрируемости уравнения (3.92) является равенство

$$dA \wedge A = dx \wedge dy \wedge dz\Big[P(\partial_y R - \partial_z Q) + Q(\partial_z P - \partial_x R) + R(\partial_x Q - \partial_y P)\Big] = 0. \qquad \square$$

3.12 Слоения

Слоения обобщают понятие расслоения, введенное в разделе 3.4. Они часто возникают и играют большую роль в теории динамических систем. Мы начнем с простейшего примера слоения, который поможет понять данное ниже общее определение.

Пример 3.12.1 (Тривиальное слоение). Представим евклидово пространство $\mathbb{R}^n$ с естественной топологией и декартовой системой координат в виде прямого произведения:

$$\mathbb{R}^n \ni x = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}, \quad 1 \leq k < n.$$

Построим многообразие $\mathbb{M}^k$ меньшей размерности k, которое как множество совпадает со всем $\mathbb{R}^n$, т.е. $\mathbb{M}^k = \mathbb{R}^n$, и задает на нем структуру слоения. Базу топологии $\mathbb{M}^k$ зададим открытыми множествами вида

$$\mathbb{B}_z := \{(y, z) \in \mathbb{R}^n : \ y \in \mathbb{O}, \},$$

где $\mathbb{O}$ – открытое множество в $\mathbb{R}^k$, для каждой фиксированной точки $z \in \mathbb{R}^{n-k}$. Эта топология тоньше исходной топологии в $\mathbb{R}^n$. Назовем листом тривиального слоения сечение евклидова пространства $\mathbb{R}^n$ гиперплоскостью $z = \mathsf{const}$:

$$\mathbb{V}_z := \{(y, z) \in \mathbb{R}^n : \ y \in \mathbb{R}^k, \ z = \mathsf{const}\}.$$

Тогда многообразие $\mathbb{M}^k$ представляет собой несвязное объединение всех листов

$$\mathbb{M}^k = \bigcup_{z \in \mathbb{R}^{n-k}} \mathbb{V}_z$$

и покрывается картами

$$\psi_z : \quad \mathbb{V}_z \ni (y, z) \mapsto y \in \mathbb{R}^k.$$

Рассмотренную конструкцию можно представить по-другому. Рассмотрим два евклидова пространства: $\mathbb{R}^k$ с обычной топологией и $\mathbb{R}^{n-k}_{\text{D}}$ с дискретной топологией, что отмечено индексом D. Тогда многообразие $\mathbb{M}^k$ является топологическим произведением $\mathbb{M}^k = \mathbb{R}^k \times \mathbb{R}^{n-k}_{\text{D}}$ и как множество совпадает с $\mathbb{R}^n$.

В этом примере многообразие $\mathbb{M}^k$ несвязно и состоит из несчетного множества связных компонент (листов) $\mathbb{V}_z$, которые параметризуются точками $z \in \mathbb{R}^{n-k}$. Эти свойства лежат в основе общего определения. □

Определение. Пусть на множестве $\mathbb{M}$ заданы две структуры многообразия разных размерностей. То есть у нас есть два многообразия: $\mathbb{M}^n$, $\dim \mathbb{M}^n = n$, и $\mathbb{M}^k$, $\dim \mathbb{M}^k = k$, где $1 \leq k < n$, которые совпадают как множества: $\mathbb{M}^n = \mathbb{M}^k = \mathbb{M}$. Многообразие $\mathbb{M}^n$ называется *слоением*, если на нем дополнительно определена другая более тонкая топология и вторая соответствующая ей структура многообразия меньшей размерности $\mathbb{M}^k$. Поскольку размерность многообразия $\mathbb{M}^k$ меньше размерности многообразия $\mathbb{M}^n$, то многообразие $\mathbb{M}^k$ не может быть связным, и $\mathbb{M}^n$ представляет собой несчетное объединение связных компонент $\mathbb{M}^k$. Связная компонента многообразия $\mathbb{M}^k$, содержащая точку $p \in \mathbb{M}^n$, называется *листом слоения*, проходящим через эту точку. Мы говорим, что многообразие $\mathbb{M}^k$ *слоит* многообразие $\mathbb{M}^n$. Числа k и $n - k$ называются соответственно *размерностью* и *коразмерностью* слоения. Слоение мы будем обозначать парой $(\mathbb{M}^n, \mathbb{M}^k)$. □

По-определению, через каждую точку $p \in \mathbb{M}^n$ проходит один и только один лист. В общем случае листы устроены довольно сложно. Как покажет дальнейший пример слоения тора, листы могут не быть даже регулярными подмногообразиями в $\mathbb{M}^n$. Многообразие $\mathbb{M}^n$ является объединением своих листов, число которых бесконечно и не может быть счетным из-за разных размерностей. Многообразие $\mathbb{M}^k$ также является объединением листов и не может быть связным, т.к. каждый лист является открытым подмножеством и представляет собой компоненту связности. В этом смысле многообразие $\mathbb{M}^n$ слоится на листы.

Можно доказать, что если $(\mathbb{M}^n, \mathbb{M}^k)$ – слоение, то на многообразии $\mathbb{M}^n$ существует атлас $\{\mathbb{U}_i, \varphi_i\}$, $i \in I$, со следующими свойствами. Пусть

$$\varphi: \quad \mathbb{M}^n \supset \mathbb{U} \to \varphi(\mathbb{U}) \subset \mathbb{R}^n$$

– карта исходной дифференцируемой структуры многообразия $\mathbb{M}^n$. Представим евклидово пространство $\mathbb{R}^n$ в виде прямого произведения $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$, тогда для всех $p \in \mathbb{U}$

$$\varphi(p) = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}.$$

Для каждой фиксированной точки $z \in \mathbb{R}^{n-k}$ обозначим

$$\varphi(\mathbb{U})|_z = \{(y, z) \in \varphi(\mathbb{U}): \ z = \mathsf{const}\}$$

сечение образа $\varphi(\mathbb{U})$ гиперплоскостью $z = \mathsf{const}$. Соответствующий прообраз

$$\mathbb{V}_z := \varphi^{-1}\Big(\varphi(\mathbb{U})|_z\Big)$$

называется *локальным листом*. Локальный лист является частью листа, проходящего через точку $\varphi^{-1}(y, z)$. Каждый локальный лист – это открытое подмногообразие (возможно пустое) в $\mathbb{M}^k$, а пара $(\mathbb{V}_z, \psi_z)$ для непустого $\mathbb{V}_z$ является картой на многообразии $\mathbb{M}^k$:

$$\psi_z: \quad \mathbb{M}^k \supset \mathbb{V}_z \to \psi_z(\mathbb{V}_z) \subset \mathbb{R}^k,$$

где отображение $\psi_z := \pi_k \circ \varphi|_{\mathbb{V}_z}$ определено естественной проекцией

$$\pi_k: \quad \mathbb{R}^k \times \mathbb{R}^{n-k} \ni (y, z) \mapsto y \in \mathbb{R}^k$$

и сужением $\varphi|_{\mathbb{V}_z}$ отображения φ на $\mathbb{V}_z$. Описанная конструкция изображена на рис.3.10.

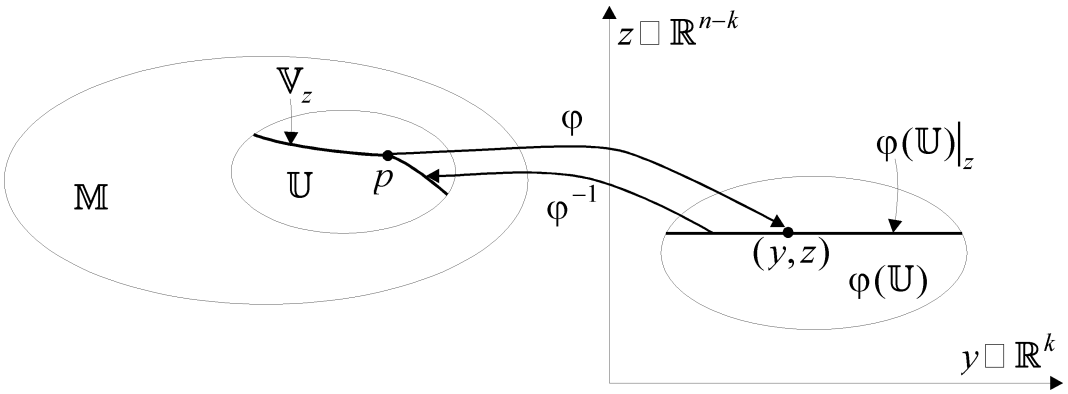


Рис. 3.10: Карта и локальный лист слоения

Функции склейки данного атласа для любых пересекающихся карт имеют вид

$$\varphi_j \circ \varphi_i^{-1} : \quad \mathbb{R}^k \times \mathbb{R}^{n-k} \supset \varphi_i(\mathbb{U}_i \cap \mathbb{U}_j) \ni y, z \mapsto y', z' \in \varphi_j(\mathbb{U}_i \cap \mathbb{U}_j) \subset \mathbb{R}^k \times \mathbb{R}^{n-k},$$

где $y' = y'(y, z)$, но $z' = z'(z)$. В этом случае локальные листы для этих карт принадлежат одному листу слоения тогда и только тогда, когда они содержат хотя бы одну точку из пересечения $p \in \mathbb{U}_i \cup \mathbb{U}_j$. Условие того, что функции перехода $z' = z'(z)$ зависят только от z необходимо и достаточно, для того, чтобы склеивание локальных листов не зависело от выбора точки p из пересечения. Действительно, если $q, p \in \mathbb{U}_i \cup \mathbb{U}_j$ – две различные точки из пересечения координатных окрестностей, то в этих картах они имеют координаты (y_p, z_p), (y'_p, z'_p) и (y_q, z_q), (y'_q, z'_q). Если $z_p = z_q$, то $z'_p = z'_q$ тогда и только тогда, когда $z' = z'(z)$.

Теперь рассмотрим несколько типичных примеров слоений.

Пример 3.12.2. Пусть $dy/dx = f(x, y)$ – обыкновенное дифференциальное уравнение, где $f(x, y)$ – функция класса $\mathcal{C}^r$ в некоторой области $\mathbb{U} \subset \mathbb{R}^2$. Из теории обыкновенных дифференциальных уравнений известно, что через каждую точку области $\mathbb{U}$ проходит одна и только одна интегральная кривая (решение) данного уравнения класса $\mathcal{C}^{r+1}$. Поскольку решение является функцией класса $\mathcal{C}^r$ от начальных данных, то объединение всех интегральных кривых образует одномерное многообразие $\mathbb{V}$, которое слоит $\mathbb{U}$.

Аналогичный результат справедлив и для систем обыкновенных дифференциальных уравнений. Для уравнений в частных производных ситуация существенно отличается. □

Пример 3.12.3 (Распределение). Пусть на многообразии $\mathbb{M}^n$ задано k-мерное распределение векторных полей $\mathcal{L}_k(\mathbb{M}^n)$, где $1 \le k < n$, которое находится в инволюции (см. раздел 3.11). Напомним, что k-мерное распределение векторных полей ставит каждой точке многообразия некоторое подпространство размерности k в касательном пространстве к данной точке: $p \mapsto \mathbb{L}_p(\mathbb{M}^n) \subset \mathbb{T}_p(\mathbb{M}^n)$. Тогда согласно теореме Фробениуса через каждую точку $p \in \mathbb{M}^n$ проходит интегральное подмногообразие данного распределения размерности k. Эти интегральные подмногообразия можно рассматривать как листы слоения $(\mathbb{M}^n, \mathbb{M}^k)$. □

Рассмотрим слоение $\mathbb{M}^n$, которое слоится многообразием $\mathbb{M}^k$. Тогда определено касательное пространство к листу $\mathbb{T}_p(\mathbb{M}^k)$ в произвольной точке $p \in \mathbb{M}^n$. Размерность этого пространства равна k, и оно является k-мерным подпространством касательного пространства $\mathbb{T}_p(\mathbb{M}^n)$. Следовательно, определено k-мерное распределение векторных полей $\mathcal{L}_k(\mathbb{M}^n)$ на многообразии $\mathbb{M}^n$. Вместе с предыдущим примером это доказывает следующее утверждение.

Теорема 3.12.1. *Любое гладкое слоение $(\mathbb{M}^n, \mathbb{M}^k)$, $1 \le k < n$, находится во взаимно однозначном соответствии с k-мерным гладким инволютивным распределением векторных полей $\mathcal{L}_k(\mathbb{M}^n)$, которое касается $\mathbb{M}^k$, т.е. $\mathbb{L}_p(\mathbb{M}^n) = \mathbb{T}_p(\mathbb{M}^k)$. При этом лист слоения, проходящий через точку $p \in \mathbb{M}^n$, совпадает с интегральным подмногообразием данного распределения, проходящим через ту же точку.*

Замечание. В силу данной теоремы слоение часто определяют с помощью распределений и их интегральных подмногообразий. □

Заметим, что не каждое многообразие допускает структуру слоения.

Пример 3.12.4. Теорема 11.2.1 утверждает, что на четномерной сфере $\mathbb{S}^{2n}$, $n = 1, 2, \ldots$, не существует непрерывного векторного поля, которое всюду отлично от нуля. Это означает, что на $\mathbb{S}^{2n}$ невозможно задать одномерное распределение векторных полей $\mathcal{L}_1(\mathbb{S}^{2n})$, и, следовательно одномерных слоений на четномерной сфере не существует. В частности, на двумерной сфере $\mathbb{S}^2$ нельзя задать структуру слоения. □

Далее мы рассмотрим более подробно связь слоений и расслоений.

Пример 3.12.5 (Расслоение). Рассмотрим дифференцируемое расслоение $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ (см., раздел 3.4) с n-мерным пространством расслоения $\mathbb{E}$, k-мерным типичным слоем $\mathbb{F}$ и $(n-k)$-мерной базой $\mathbb{M}$. Через каждую точку расслоения $p \in \mathbb{E}$ проходит слой $\mathbb{V}_{\pi(p)} := \pi^{-1}\Big(\pi(p)\Big)$, диффеоморфный типичному слою $\mathbb{F}$, и, значит, имеющий структуру дифференцируемого k-мерного многообразия. Известно, что слои расслоения либо совпадают: $\mathbb{V}_{\pi(p)} = \mathbb{V}_{\pi(q)}$, если $\pi(p) = \pi(q)$, либо не пересекаются: $\mathbb{V}_{\pi(p)} \cap \mathbb{V}_{\pi(q)} = \emptyset$, если $\pi(p) \neq \pi(q)$. Рассмотрим множество, которое состоит из объединения всех слоев

$$\mathbb{E}^k = \bigcup_{x \in \mathbb{M}} \mathbb{V}_x.$$

Как множество оно совпадает с пространством расслоения $\mathbb{E}$, но мы введем на нем другую топологию и дифференцируемую структуру. А именно, будем считать, что в этой топологии и дифференцируемой структуре каждый слой является открытым подмножеством и открытым подмногообразием. Для того, чтобы проверить, что многообразие $\mathbb{E}^k$ слоит расслоение $\mathbb{E}$ следует воспользоваться картами на $\mathbb{E}$, которые тривиализируют расслоение. Как и в предыдущем примере многообразие $\mathbb{E}^k$ несвязно и состоит из несчетного числа листов, которые совпадают со слоями расслоения $\mathbb{V}_x$ и параметризуются точками базы $x \in \mathbb{M}$. В этом слоении все листы диффеоморфны между собой. □

Рассмотренный пример показывает, что на каждом расслоении определена естественная структура слоения. Обратное утверждение неверно: не каждое слоение является расслоением. Чтобы построить соответствующий пример, нам понадобится

Определение. Дифференцируемое отображение двух многообразий $f : \mathbb{M} \to \mathbb{N}$ называется *субмерсией*, если дифференциал отображения $f_* : \mathbb{T}_p(\mathbb{M}) \to \mathbb{T}_{f(p)}(\mathbb{N})$ является сюръективным отображением для всех точек $p \in \mathbb{M}$. □

Пример 3.12.6 (Субмерсия). Рассмотрим субмерсию двух многообразий $f : \mathbb{M} \to \mathbb{N}$. Из определения следует, что $\dim \mathbb{M} \geq \dim \mathbb{N}$. Мы положим $\dim \mathbb{M} = n$ и $\dim \mathbb{N} = n - k$, где $1 \leq k < n$. Каждая точка $p \in \mathbb{M}$ порождает лист или несвязное объединение листов $\mathbb{V}_p := f^{-1}\Big(f(p)\Big)$. При этом прообразы либо совпадают: $\mathbb{V}_p = \mathbb{V}_q$, если $f(p) = f(q)$, либо не пересекаются: $\mathbb{V}_p \cap \mathbb{V}_q = \emptyset$, если $f(p) \neq f(q)$. Предположим, что отображение f сюръективно. Тогда слои можно параметризовать точками $x \in \mathbb{N}$. Из условия сюръективности $f_*\Big(\mathbb{T}_p(\mathbb{M})\Big) = \mathbb{T}_{f(p)}(\mathbb{N})$ для всех $p \in \mathbb{M}$ следует, что для каждой точки $p \in \mathbb{M}$ найдутся карты:

$$\begin{aligned} \varphi &: \ \mathbb{M} \supset \mathbb{U} \ni p \mapsto \varphi(p) \in \mathbb{R}^n, \\ \chi &: \ \mathbb{N} \supset \mathbb{W} \ni x = f(p) \mapsto \chi(x) \in \mathbb{R}^{n-k} \end{aligned}$$

со следующими свойствами:

1. $f(\mathbb{U}) \subset \mathbb{W}$,
2. $\varphi(q) = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k}, \ \ \forall q \in \mathbb{U}$,
3. $\chi(x) \in \mathbb{R}^{n-k}, \ \ \forall x \in \mathbb{W}$,
4. $(\chi \circ f \circ \varphi^{-1})(y, z) = z \in \chi(\mathbb{W}), \ \ \forall (y, z) \in \varphi(\mathbb{U})$,
5. $\varphi(\mathbb{V}_p) = \{(y, z) \in \varphi(\mathbb{U}) : \ z = \chi\Big(f(p)\}, \ \ \forall f(p) \in \mathbb{W}$.

Отсюда легко выводится, что на множестве $\mathbb{M}$ определена структура k-мерного слоения $\mathbb{M}^k$, в котором листы, содержащиеся в $f^{-1}(x)$, являются открытыми подмногообразиями. □

Ранее было показано, что любое расслоение является слоением. Приведем пример слоения, которое не является расслоением.

Пример 3.12.7. Рассмотрим субмерсию, изображенную на рис.3.11, где лента в $\mathbb{R}^2$ с вырезанной дыркой проектируется на ось x. Ленту с вырезанной дыркой мы рассматриваем, как открытое подмногообразие $\mathbb{M}$ в $\mathbb{R}^2$ с декартовыми координатами x, y. Субмерсия задается проекцией $f : \ (x, y) \mapsto x$ и определяет одномерное слоение $\mathbb{M}$. Соответствующее распределение векторных полей имеет вид ∂_y. Все интегральные кривые этого распределения, которые являются листами слоения, параллельны оси y. Если $x < a$ или $x > b$, то листами слоения $f^{-1}(x)$ являются интервалы. В то же время каждый прообраз $f^{-1}(x)$ для $a \leq x \leq b$ представляет несвязное объединение двух листов. Ясно, что не все прообразы $f^{-1}(x)$ данного слоения диффеоморфны между собой, и поэтому рассматриваемое слоение не может быть расслоением. □

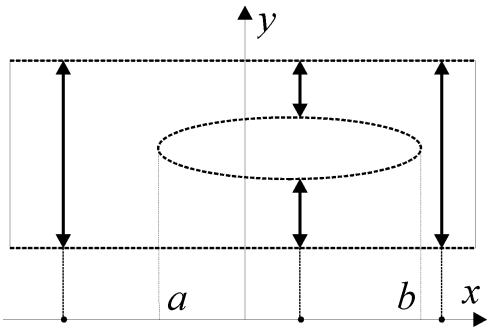


Рис. 3.11: Субмерсия ленты с вырезанной дыркой на ось x.

Следующий пример показывает, что листы слоения даже в простых случаях могут быть устроены достаточно сложно.

Пример 3.12.8 (Слоения тора). Рассмотрим двумерный тор, как квадрат на плоскости: $\mathbb{T}^2 = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ с отождествленными противоположными сторонами. Пусть на нем задано векторное поле $X^x = a$, $X^y = b$, где $a \neq 0, b$ – некоторые постоянные. Тем самым определено одномерное распределение векторного поля. Оно, конечно, находится в инволюции. Обозначим интегральные кривые этого векторного поля через γ_{ab}. Они же являются листами одномерного слоения тора. Если отношение b/a рационально, то каждый лист диффеоморфен окружности $\mathbb{S}^1$ и является регулярным замкнутым подмногообразием в $\mathbb{T}^2$. Если отношение b/a иррационально (всюду плотная обмотка тора в примере 3.10.2), то каждый лист диффеоморфен прямой $\mathbb{R}$. При этом отображение $\mathbb{R} \to \gamma_{ab} \in \mathbb{T}^2$ не является регулярным. В этом случае лист не является одномерным регулярным подмногообразием и замкнутым подмножеством в $\mathbb{T}^2$. □

3.13 Бесконечно малые преобразования координат

В настоящем разделе рассматриваются свойства полей на многообразии при бесконечно малых преобразованиях координат. Полученные ниже формулы важны в приложениях, в частности, они будут использованы в следующем разделе для вычисления производных Ли.

Поскольку рассмотрение носит локальных характер, то для простоты и наглядности мы рассмотрим евклидово пространство $\mathbb{R}^n$ с фиксированной декартовой системой координат. Будем рассматривать преобразования координат (2.55) с активной точки зрения. То есть будем считать, что в евклидовом пространстве точка с координатами x^α перемещается в новую точку с координатами $x^{\alpha'}(x^\alpha)$ в той же системе координат. Бесконечно малые преобразования координат

$$x'^\alpha = x^\alpha + u^\alpha(x) \tag{3.94}$$

определяются векторным полем

$$u = u^\alpha \partial_\alpha.$$

Обозначение $u^\alpha(x)$ заимствовано из теории упругости, где этот вектор называется *вектором смещения*. Будем считать, что малы сами смещения, $|u^\alpha| \ll 1$[2], и относительные смещения $|\partial_\alpha u^\beta| \ll 1$ для всех значений индексов α, β.

Рассмотрим, как меняются тензорные поля при инфинитезимальных преобразованиях координат (3.94). Начнем с простейшего случая скалярного поля (функции) $f \in \mathcal{C}^\infty(\mathbb{R}^n)$. Из закона преобразования скалярного поля (3.5) следует равенство

$$f'(x+u) = f(x). \tag{3.95}$$

Разлагая левую часть этого равенства в ряд Тейлора[3], в первом порядке по u получим следующее выражение

$$f' = f - u^\alpha \partial_\alpha f, \tag{3.96}$$

где все поля рассматриваются в точке x. Полученное соотношение интерпретируется следующим образом. При инфинитезимальном преобразовании координат (3.94) значение функции $f(x)$ в точке x меняется на бесконечно малую величину

$$\delta f(x) = f'(x) - f(x) = -u^\alpha \partial_\alpha f = -uf. \tag{3.97}$$

которая называется *приращением* или *вариацией формы* функции f в точке x. Правая часть этого равенства с обратным знаком представляет собой производную от функции f вдоль векторного поля u, определяющего преобразование координат.

Из закона преобразования векторных полей (3.11), который можно переписать в виде

$$X'^\alpha(x+u) = X^\beta(\delta^\alpha_\beta + \partial_\beta u^\alpha),$$

следует, что

$$\delta X^\alpha = X^\beta \partial_\beta u^\alpha - u^\beta \partial_\beta X^\alpha = [X, u]^\alpha. \tag{3.98}$$

То есть приращение векторного поля в точке x равно коммутатору этого векторного поля с инфинитезимальным векторным полем смещения u.

---

[2] В физике координаты являются размерными величинами, и условие малости смещений записывается в виде $|u^\alpha| \ll l$, где l – выбранная единица длины.

[3] Здесь и далее мы предполагаем, что разложение всех функций и компонент тензорных полей в ряд Тейлора имеет смысл в некоторой окрестности точки x.

Закон преобразования компонент 1-формы (3.12) приводит к следующему приращению компонент

$$\delta\omega_\alpha = -\partial_\alpha u^\beta \omega_\beta - u^\beta \partial_\beta \omega_\alpha \tag{3.99}$$

или в ковариантном виде

$$\begin{aligned}\delta\omega_\alpha &= -\nabla_\alpha u^\beta \omega_\beta - u^\beta \nabla_\beta \omega_\alpha + u^\beta T_{\alpha\beta}{}^\gamma \omega_\gamma \\ &= -\widetilde{\nabla}_\alpha u^\beta \omega_\beta - u^\beta \widetilde{\nabla}_\beta \omega_\alpha.\end{aligned} \tag{3.100}$$

При этом использовано явное выражение для аффинной связности (7.15). Напомним, что ∇ и $\widetilde{\nabla}$ обозначают ковариантные производные производные с аффинной связностью и символами Кристоффеля, соответственно.

Отметим, что приращения тензорных полей, а также связности при инфинитезимальных преобразованиях координат являются тензорными полями. Поскольку бесконечно малые преобразования координат параметризуются векторным полем u, то вариации полей всегда можно записать в явно ковариантном виде.

Аналогично находится явный вид бесконечно малых преобразований координат для тензоров произвольного ранга. Рассмотрим бесконечно малые преобразования основных геометрических объектов: метрики и аффинной связности, а также репера и линейной связности. В частности, для метрики справедливо равенство

$$\delta g_{\alpha\beta} = -\partial_\alpha u^\gamma g_{\gamma\beta} - \partial_\beta u^\gamma g_{\alpha\gamma} - u^\gamma \partial_\gamma g_{\alpha\beta} \tag{3.101}$$

или в ковариантном виде

$$\begin{aligned}\delta g_{\alpha\beta} &= -\nabla_\alpha u_\beta - \nabla_\beta u_\alpha + u^\gamma (T_{\alpha\gamma\beta} + T_{\beta\gamma\alpha} - Q_{\alpha\gamma\beta} - Q_{\beta\gamma\alpha} + Q_{\gamma\alpha\beta}) \\ &= -\widetilde{\nabla}_\alpha u_\beta - \widetilde{\nabla}_\beta u_\alpha.\end{aligned} \tag{3.102}$$

Отсюда следует, что, если метрика инвариантна относительно бесконечно малых преобразований координат, $\delta g_{\alpha\beta} = 0$, то соответствующее векторное поле должно удовлетворять уравнению Киллинга (16.4).

Из уравнения (3.101) следует правило преобразования определителя метрики

$$\delta g = -u^\alpha \partial_\alpha g - 2\partial_\alpha u^\alpha g. \tag{3.103}$$

Рассмотрим, как меняются компоненты аффинной связности при инфинитезимальных преобразованиях координат. Из закона преобразования (7.5) следует, что

$$\delta\Gamma_{\alpha\beta}{}^\gamma = -\partial_\alpha u^\delta \Gamma_{\delta\beta}{}^\gamma - \partial_\beta u^\delta \Gamma_{\alpha\delta}{}^\gamma + \Gamma_{\alpha\beta}{}^\delta \partial_\delta u^\gamma - \partial^2_{\alpha\beta} u^\gamma - u^\delta \partial_\delta \Gamma_{\alpha\beta}{}^\gamma. \tag{3.104}$$

В ковариантном виде это изменение имеет вид

$$\delta\Gamma_{\alpha\beta}{}^\gamma = -\nabla_\alpha \nabla_\beta u^\gamma + \nabla_\alpha u^\delta T_{\beta\delta}{}^\gamma + u^\delta (R_{\alpha\delta\beta}{}^\gamma + \nabla_\alpha T_{\beta\delta}{}^\gamma). \tag{3.105}$$

В частном случае для символов Кристоффеля справедливо равенство

$$\delta\widetilde{\Gamma}_{\alpha\beta}{}^\gamma = -\widetilde{\nabla}_\alpha \widetilde{\nabla}_\beta u^\gamma + u^\delta \widetilde{R}_{\alpha\delta\beta}{}^\gamma. \tag{3.106}$$

Эту вариацию можно переписать в виде

$$\delta\widetilde{\Gamma}_{\alpha\beta\gamma} = \frac{1}{2}(\widetilde{\nabla}_\alpha \delta g_{\beta\gamma} + \widetilde{\nabla}_\beta \delta g_{\alpha\gamma} - \widetilde{\nabla}_\gamma \delta g_{\alpha\beta}). \tag{3.107}$$

В римановой геометрии, где связность однозначно определяется метрикой, инвариантность метрики является достаточным условием инвариантности символов Кристоффеля. Необходимое и достаточное условие инвариантности символов Кристоффеля имеет вид (3.106). Антисимметризация уравнения (3.105) по индексам α, β дает приращение тензора кручения:

$$\begin{aligned}\delta T_{\alpha\beta}{}^{\gamma} &= \nabla_\alpha u^\delta T_{\beta\delta}{}^{\gamma} - \nabla_\beta u^\delta T_{\alpha\delta}{}^{\gamma} + T_{\alpha\beta}{}^{\delta}\nabla_\delta u^\gamma \\ &\quad + u^\delta(-R_{\alpha\beta\delta}{}^{\gamma} - R_{\delta\alpha\beta}{}^{\gamma} - R_{\beta\delta\alpha}{}^{\gamma} + \nabla_\alpha T_{\beta\delta}{}^{\gamma} - \nabla_\beta T_{\alpha\delta}{}^{\gamma})\end{aligned} \tag{3.108}$$

или, с учетом (7.68),

$$\begin{aligned}\delta T_{\alpha\beta}{}^{\gamma} &= \nabla_\alpha u^\delta T_{\beta\delta}{}^{\gamma} - \nabla_\beta u^\delta T_{\alpha\delta}{}^{\gamma} + T_{\alpha\beta}{}^{\delta}\nabla_\delta u^\gamma \\ &\quad - u^\delta(\nabla_\delta T_{\alpha\beta}{}^{\gamma} + T_{\alpha\beta}{}^{\epsilon}T_{\epsilon\delta}{}^{\gamma} + T_{\beta\delta}{}^{\epsilon}T_{\epsilon\alpha}{}^{\gamma} + T_{\delta\alpha}{}^{\epsilon}T_{\epsilon\beta}{}^{\gamma})\end{aligned} \tag{3.109}$$

Нетрудно получить также изменение тензора неметричности при инфинитезимальных преобразованиях координат:

$$\begin{aligned}\delta Q_{\alpha\beta\gamma} &= -\nabla_\alpha u^\delta Q_{\delta\beta\gamma} - \nabla_\beta u^\delta Q_{\alpha\delta\gamma} - \nabla_\gamma u^\delta Q_{\alpha\beta\delta} \\ &\quad + u^\delta(T_{\alpha\delta}{}^{\epsilon}Q_{\epsilon\beta\gamma} + T_{\beta\delta}{}^{\epsilon}Q_{\alpha\epsilon\gamma} + T_{\gamma\delta}{}^{\epsilon}Q_{\alpha\beta\epsilon}).\end{aligned} \tag{3.110}$$

В аффинной геометрии при заданной связности можно поставить задачу нахождения таких преобразований координат, которые не меняют связность, кручение или неметричность. Для этого необходимо решить, соответственно, уравнения (3.105), (3.108) или (3.110).

Репер и лоренцева связность являются ковекторами относительно преобразований координат. Поэтому их приращения такие же, как и приращения 1-форм (3.99). Приведем их явный вид, т.к. они часто встречаются в приложениях

$$\delta e_\alpha{}^a = -\partial_\alpha u^\beta e_\beta{}^a - u^\beta \partial_\beta e_\alpha{}^a = -\widetilde{\nabla}_\alpha u^a + u^\beta \widetilde{\omega}_{\beta b}{}^a e_\alpha{}^b, \tag{3.111}$$

$$\delta\sqrt{|g|} = -u^\alpha\partial_\alpha\sqrt{|g|} - \sqrt{|g|}\partial_\alpha u^\alpha = -\sqrt{|g|}\widetilde{\nabla}_\alpha u^\alpha, \tag{3.112}$$

$$\delta\omega_\alpha{}^{ab} = -\partial_\alpha u^\beta \omega_\beta{}^{ab} - u^\beta\partial_\beta\omega_\alpha{}^{ab} = -\widetilde{\nabla}_\alpha u^\beta \omega_\beta{}^{ab} - u^\beta(\partial_\beta\omega_\alpha{}^{ab} - \widetilde{\Gamma}_{\beta\alpha}{}^{\gamma}\omega_\gamma{}^{ab}). \tag{3.113}$$

Бесконечно малые приращения различных полей, рассмотренные в настоящем разделе, называются *вариацией формы*, т.к. показывают как меняется значение полей в фиксированной точке многообразия. Следующее свойство проверяется прямой проверкой и используется при доказательстве теоремы Нетер.

Предложение 3.13.1. *Вариация формы производной скалярного поля равно производной от вариации*

$$\delta(\partial_\alpha f) = \partial_\alpha(\delta f). \qquad \square$$

Рассмотрим коммутатор двух преобразований координат. Пусть эти преобразования задаются векторными полями $u = u^\alpha\partial_\alpha$ и $v = v^\alpha\partial_\alpha$. Для простоты ограничимся их последовательным действием на функцию $f(x)$. Поскольку преобразования координат евклидова пространства $\mathbb{R}^n$ образуют псевдогруппу, то коммутатор двух преобразований также является преобразованием координат с параметром, который квадратичен по u^α и v^α. Поэтому, в отличие от предыдущего рассмотрения, для вычисления коммутатора необходимо удерживать квадратичные слагаемые.

Сначала совершим преобразование координат, определяемое векторным полем u. Из закона преобразования функции (3.95) с учетом квадратичных слагаемых получаем

$$f' + u^\alpha \partial_\alpha f' + \frac{1}{2} u^\alpha u^\beta \partial^2_{\alpha\beta} f' = f.$$

Во втором слагаемом в левой части равенства f' можно выразить через f, воспользовавшись линейным приближением (3.96), а в третьем слагаемом заменим f' на f, потому что коэффициент перед ним уже квадратичен по u. В результате с точностью до квадратичных слагаемых получаем

$$f' := (1 + T_u) f = f - u^\alpha \partial_\alpha f + u^\alpha \partial_\alpha u^\beta \partial_\beta f + \frac{1}{2} u^\alpha u^\beta \partial^2_{\alpha\beta} f + \ldots, \tag{3.114}$$

где мы ввели генератор общих преобразований координат T_u и выписали его действие на функцию в квадратичном приближении.

Совершим теперь второе преобразование координат с параметром v

$$\begin{aligned}
(1 + T_v) f' &= (1 + T_v)(1 + T_u) f = \\
&= f' - v^\alpha \partial_\alpha f' + v^\alpha \partial_\alpha v^\beta \partial_\beta f' + \frac{1}{2} v^\alpha v^\beta \partial^2_{\alpha\beta} f' + \ldots = \\
&= f - (u^\alpha + v^\alpha - u^\beta \partial_\beta u^\alpha - v^\beta \partial_\beta v^\alpha - v^\beta \partial_\beta u^\alpha) \partial_\alpha f \\
&\quad + \frac{1}{2} (u^\alpha u^\beta + 2 u^\alpha v^\beta + v^\alpha v^\beta) \partial^2_{\alpha\beta} f + \ldots .
\end{aligned}$$

Вычитая из этого выражения результат тех же преобразований в обратном порядке, получим явное выражение для коммутатора двух преобразований координат

$$[T_v, T_u] f = T_{[v,u]} f = (v^\beta \partial_\beta u^\alpha - u^\beta \partial_\beta v^\alpha) \partial_\alpha f. \tag{3.115}$$

Таким образом, коммутатор двух преобразований координат в евклидовом пространстве $\mathbb{R}^n$, определяемых инфинитезимальными векторными полями u и v, является преобразованием координат, которое задается коммутатором векторных полей $[v, u]$.

Из групповых соображений следует, что коммутатор двух преобразований координат не зависит от представления. Другими словами, это же выражение для коммутатора имеет место не только для функций, но и для тензорных полей или плотностей более высокого ранга. В последнем случае вычисления являются более громоздкими.

Рассмотрим, как меняются при бесконечно малых преобразованиях тензорные плотности, заданные на многообразии $\mathbb{M}$. Мы уже получили правила преобразования определителей метрики (3.103) и репера (3.112), которые являются тензорными плотностями соответственно степеней -2 и -1. В общем случае, если на $\mathbb{M}$ задана скалярная плотность f степени $\mathsf{deg}\, f = p$, то из правила преобразования тензорных плотностей (3.7) следует правило

$$\delta f = -u^\alpha \partial_\alpha f + p \partial_\alpha u^\alpha f,$$

т.к. якобиан бесконечно малых преобразований координат (3.94) в первом порядке по вектору смещения имеет вид

$$J = \mathsf{det}\, (\delta^\beta_\alpha + \partial_\alpha u^\beta) \approx 1 + \partial_\alpha u^\alpha. \tag{3.116}$$

Если на многообразии $\mathbb{M}$ задана тензорная плотность T произвольного ранга и степени p, то ее вариация формы равна

$$\delta T = \ldots + p \partial_\alpha u^\alpha T,$$

где точки обозначают набор обычных тензорных слагаемых для всех ковариантных и контравариантных индексов.

3.14 Производная Ли

Понятие потока $\mathbb{R}\times\mathbb{M} \xrightarrow{s} \mathbb{M}$ или экспоненциального отображения $s^\alpha(t,x) = \mathsf{exp}\,(tX)x^\alpha$, генерируемого векторным полем $X \in \mathcal{X}(\mathbb{M})$ (раздел 3.6.5), позволяет определить производную Ли L_X от произвольного тензорного поля T вдоль векторного поля X. Для этого определения достаточно, чтобы поток существовал локально.

Рассмотрим произвольную точку $x \in \mathbb{M}$. Тогда у нее существует окрестность $\mathbb{U}_x$ такая, что отображение $s:\ \mathbb{U}_x \to s(\mathbb{U}_x)$ является диффеоморфизмом. Поэтому для него определены дифференциал отображения s_* и обратное отображение s^{*-1} к возврату отображения s^*. Пусть в некоторой окрестности точки $x \in \mathbb{M}$ задано тензорное поле $T \in \mathcal{T}^r_s(\mathbb{M})$ типа (r,s). В результате экспоненциального отображения $s(\epsilon,x)$ с малым параметром ϵ тензор $T(x)$ в точке x отобразится в тензор $\tilde{s}T(x)$ в точке $s(\epsilon,x)$, где $\tilde{s}T$ обозначает продолжение отображения $(s_*)^r(s^{*-1})^sT$, заданного в касательном и кокасательном пространствах, на всю тензорную алгебру (в компонентах: на каждый контравариантный индекс действует дифференциал отображения s_*, а на каждый ковариантный индекс – обратное отображение s^{*-1}). Это значит, что в точку x отобразится тензор из точки $s(-\epsilon,x)$:

$$T\Big(s(-\epsilon,x)\Big)\ \mapsto\ \tilde{s}\Big(\epsilon,s(-\epsilon,x)\Big)T\Big(s(-\epsilon,x)\Big).$$

Определение. *Производной Ли* от тензорного поля T вдоль векторного поля X в точке x называется предел

$$\mathsf{L}_XT = \lim_{\epsilon\to0}\frac{T(x)-\tilde{s}\Big(\epsilon,s(-\epsilon,x)\Big)T\Big(s(-\epsilon,x)\Big)}{\epsilon}. \tag{3.117}$$

В упрощенной записи мы пишем

$$\mathsf{L}_XT = \lim_{\epsilon\to0}\frac{T(x)-\tilde{s}(\epsilon)T}{\epsilon}, \tag{3.118}$$

где мы опустили аргумент $s(-\epsilon,x) \in \mathbb{M}$. □

Для дифференцируемых векторных и тензорных полей этот предел существует.

Замечание. При малых ϵ экспоненциальное отображение имеет вид $x^\alpha \mapsto s^\alpha(\epsilon,x) = x^\alpha + \epsilon X^\alpha + \ldots$, т.е. соответствует бесконечно малым преобразованиям координат, рассмотренным в предыдущем разделе, с вектором смещения $u^\alpha = \epsilon X^\alpha$. При этом выражение, стоящее в числителе производной Ли (3.117), совпадает с вариацией формы тензорного поля T, взятой с обратным знаком. Это следует непосредственно из определения вариации формы тензорного поля. Поэтому определение (3.117) для компонент тензорного поля типа (r,s) принимает вид

$$\mathsf{L}_XT_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x) = -\lim_{\epsilon\to0}\frac{\delta T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x)}{\epsilon}, \tag{3.119}$$

где $\delta T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}(x)$ – вариация формы компонент тензорного поля типа (r,s), рассмотренная в разделе 3.13. Несложные вычисления приводят к следующему выражению для производной Ли компонент тензорного поля

$$\begin{aligned}\mathsf{L}_XT_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} &= X^\gamma\partial_\gamma T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r}\\ &+ \partial_{\beta_1}X^\gamma T_{\gamma\beta_2\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} + \ldots + \partial_{\beta_s}X^\gamma T_{\beta_1\ldots\beta_{s-1}\gamma}{}^{\alpha_1\ldots\alpha_r}\\ &- T_{\beta_1\ldots\beta_s}{}^{\gamma\alpha_2\ldots\alpha_r}\partial_\gamma X^{\alpha_1} - \ldots - T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_{r-1}\gamma}\partial_\gamma X^{\alpha_1}.\end{aligned} \tag{3.120}$$

Первое слагаемое в правой части (3.120) соответствует смещению самой точки x, слагаемые во второй и третьей строках (3.120) возникают при действии отображений s^{*-1} и s_* на каждый ковариантный и контравариантный индекс, соответственно. □

Получим явные выражения для производных Ли различных тензорных полей в координатах.

Пример 3.14.1. В простейшем случае скалярного поля $f(x)$ производная Ли совпадает с производной функции вдоль векторного поля. Покажем это. Поскольку на скалярное поле ни дифференциал отображения, ни его возврат не действуют, то из определения (3.117) следует

$$\mathsf{L}_X f = \lim_{\epsilon\to 0} \frac{f(x) - f(x-\epsilon X)}{\epsilon} = X^\alpha \partial_\alpha f.$$

Таким образом, производная Ли от функции – это просто производная вдоль векторного поля. □

Пример 3.14.2. Из уравнения (3.120) следует выражение для производной Ли от векторного поля Y в координатах:

$$\mathsf{L}_X Y^\alpha = X^\beta \partial_\beta Y^\alpha - Y^\beta \partial_\beta X^\alpha,$$

что совпадает с коммутатором векторных полей (3.48)

$$\mathsf{L}_X Y = [X, Y]. \tag{3.121}$$

Определение производной Ли (3.117) для векторного поля принимает вид

$$\mathsf{L}_X Y = \lim_{\epsilon\to 0} \frac{Y - s_*(\epsilon)Y}{\epsilon}. \quad \square \tag{3.122}$$

Производная Ли от векторного поля обладает следующими свойствами:

$$\mathsf{L}_{fX} Y = f[X, Y] - (Yf)X,$$
$$\mathsf{L}_X (fY) = f[X, Y] + (Xf)Y,$$

где $f(x)$ – произвольное скалярное поле.

Отметим, что тождества Якоби (3.47) в алгебре Ли векторных полей можно переписать в эквивалентном виде, используя производную Ли,

$$\mathsf{L}_X [Y, Z] = [\mathsf{L}_X Y, Z] + [Y, \mathsf{L}_X Z].$$

В таком виде тождества Якоби аналогичны правилу Лейбница.

Можно также проверить следующее свойство производной Ли:

$$\mathsf{L}_{[X,Y]} = \mathsf{L}_X \circ \mathsf{L}_Y - \mathsf{L}_Y \circ \mathsf{L}_X. \tag{3.123}$$

Действие этого равенства на векторное поле сводится к тождеству Якоби.

Производная Ли отличается от ковариантной производной вдоль векторного поля (7.11). В частности, для векторного поля имеем

$$\mathsf{L}_X Y^\alpha - \nabla_X Y^\alpha = -Y^\beta (\nabla_\beta X^\alpha + X^\gamma T_{\beta\gamma}{}^\alpha), \tag{3.124}$$

где ∇_α – ковариантная производная и $T_{\beta\gamma}{}^\alpha$ – тензор кручения. Это равенство доказывается путем простой проверки.

Пример 3.14.3. Из формулы (3.99) следует, что производная Ли от 1-формы $A = dx^\alpha A_\alpha$ имеет вид

$$\mathsf{L}_X A_\alpha = X^\beta \partial_\beta A_\alpha + \partial_\alpha X^\beta A_\beta.$$

При этом разность с ковариантной производной вдоль векторного поля имеет другой знак по сравнению с (3.124)

$$\mathsf{L}_X A_\alpha - \nabla_X A_\alpha = (\nabla_\alpha X^\beta + X^\gamma T_{\alpha\gamma}{}^\beta) A_\beta. \qquad \square$$

Определение. *Дифференцированием* D тензорной алгебры $\mathcal{T}(\mathbb{M})$ называется линейный эндоморфизм $\mathcal{T}(\mathbb{M})$, удовлетворяющий следующим свойствам:

1) D сохраняет тип тензорных полей: $D\mathcal{T}^r_s(\mathbb{M}) \subset \mathcal{T}^r_s(\mathbb{M})$.
2) Дифференцирование удовлетворяет правилу Лейбница. Если Y и Z – два произвольных тензорных поля на $\mathbb{M}$, то

$$D(Y \otimes Z) = DY \otimes Z + Y \otimes DZ.$$

3) D перестановочен с каждым свертыванием: $DC = CD$. $\square$

Предложение 3.14.1. *Производная Ли* L_X *является дифференцированием в тензорной алгебре* $\mathcal{T}(\mathbb{M})$.

Доказательство. Прямая проверка. $\square$

Рассмотрим тензорное поле $S = S_\beta{}^\alpha(x) e_\alpha \otimes e^\beta$ типа $(1,1)$. В каждой точке $x \in \mathbb{M}$ оно задает линейный эндоморфизм касательного $\mathbb{T}_x(\mathbb{M})$ и кокасательного $\mathbb{T}^*_x(\mathbb{M})$ пространств, который естественным образом продолжается до линейного эндоморфизма D_S тензорной алгебры в этой точке. Если $T \in \mathcal{T}^r_s(\mathbb{M})$ – тензорное поле типа (r,s), то в компонентах

$$\begin{aligned}(D_S T)_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} = {} & S_{\beta_1}{}^\gamma T_{\gamma\beta_2\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} + \ldots + S_{\beta_s}{}^\gamma T_{\beta_1\ldots\beta_{s-1}\gamma}{}^{\alpha_1\ldots\alpha_r} - \\ & - T_{\beta_1\ldots\beta_s}{}^{\gamma\alpha_2\ldots\alpha_r} S_\gamma{}^{\alpha_1} - \ldots - T_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_{r-1}\gamma} S_\gamma{}^{\alpha_r}.\end{aligned}$$

Тогда D_S представляет собой дифференцирование тензорной алгебры $\mathcal{T}(\mathbb{M})$, индуцированное тензорным полем S типа $(1,1)$.

Теорема 3.14.1. *Каждое дифференцирование* D *тензорной алгебры* $\mathcal{T}(\mathbb{M})$ *допускает единственное разложение*

$$D = \mathsf{L}_X + D_S,$$

где X *есть векторное поле, а* S *– тензорное поле типа* $(1,1)$.

Доказательство. См., например, [**?**]. $\square$

Замечание. Поскольку любое дифференцирование D линейно отображает алгебру функций $\mathcal{C}^\infty(\mathbb{M})$ в себя, и выполняется правило Лейбница, то векторное поле X существует и определяется единственным образом из условия $Df = Xf$ для всех $f \in \mathcal{C}^\infty(\mathbb{M})$. $\square$

Следствие. Разность двух произвольных дифференцирований есть дифференцирование, индуцированное некоторым тензорным полем типа $(1,1)$. $\square$

Таким образом, у нас есть два дифференцирования в тензорной алгебре $\mathcal{T}(\mathbb{M})$: производная Ли вдоль векторного поля L_X и дифференцирование D_S, индуцированное тензорным полем S типа $(1,1)$. В разделе **??** будет введено еще одно дифференцирование в тензорной алгебре – ковариантное дифференцирование ∇. Согласно теореме 3.14.1, эти три вида дифференцирований связаны между собой. Нетрудно убедиться, что в общем случае производная Ли L_X и ковариантная производная вдоль векторного поля ∇_X связаны следующим соотношением

$$
\begin{aligned}
&(\mathsf{L}_X - \nabla_X) Y_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} = \\
&= (\nabla_{\beta_1} X^\gamma + T_{\beta_1\delta}{}^\gamma X^\delta) Y_{\gamma\beta_2\ldots\beta_s}{}^{\alpha_1\ldots\alpha_r} + \ldots + (\nabla_{\beta_s} X^\gamma + T_{\beta_s\delta}{}^\gamma X^\delta) Y_{\beta_1\ldots\beta_{s-1}\gamma}{}^{\alpha_1\ldots\alpha_r} \\
&- Y_{\beta_1\ldots\beta_s}{}^{\gamma\alpha_2\ldots\alpha_r} (\nabla_\gamma X^{\alpha_1} + T_{\gamma\delta}{}^{\alpha_1} X^\delta) - \ldots - Y_{\beta_1\ldots\beta_s}{}^{\alpha_1\ldots\alpha_{r-1}\gamma} (\nabla_\gamma X^{\alpha_r} + T_{\gamma\delta}{}^{\alpha_r} X^\delta).
\end{aligned}
$$

То есть разность ковариантного дифференцирования и производной Ли является дифференцированием, индуцированном тензорным полем с компонентами $S_\beta{}^\alpha := \nabla_\beta X^\alpha + T_{\beta\gamma}{}^\alpha X^\gamma$.

Как следствие этой формулы или непосредственной проверкой можно убедиться в том, что в правой части (3.120) все частные производные можно выразить через ковариантные. Это естественно, т.к. производная Ли является инвариантным оператором. В этом случае появляются дополнительные слагаемые, содержащие тензор кручения для каждого тензорного индекса. Если кручение аффинной связности равно нулю, то все частные производные в производной Ли (3.120) можно просто заменить на ковариантные.

Пример 3.14.4. Для иллюстрации приведем формулу производной Ли от компонент тензора $Y_\alpha{}^\beta$ типа (1,1)

$$
\begin{aligned}
\mathsf{L}_X Y_\alpha{}^\beta =& X^\gamma \partial_\gamma Y_\alpha{}^\beta + \partial_\alpha X^\gamma Y_\gamma{}^\beta - Y_\alpha{}^\gamma \partial_\gamma X^\beta \\
=& X^\gamma \nabla_\gamma Y_\alpha{}^\beta + \nabla_\alpha X^\gamma Y_\gamma{}^\beta - Y_\alpha{}^\gamma \nabla_\gamma X^\beta \\
&+ X^\gamma Y_\alpha{}^\delta T_{\delta\gamma}{}^\beta - X^\gamma Y_\delta{}^\beta T_{\alpha\gamma}{}^\delta.
\end{aligned}
$$

Аналогичные формулы справедливы для произвольных тензорных полей. □

Замечание. Производная Ли не зависит ни от метрики, ни от аффинной связности, которые могут быть заданы на многообразии совершенно независимо. □

Пример 3.14.5. Приведем также явное выражение для производной Ли от метрики

$$
\mathsf{L}_X g_{\alpha\beta} = X^\gamma \partial_\gamma g_{\alpha\beta} + \partial_\alpha X^\gamma g_{\gamma\beta} + \partial_\beta X^\gamma g_{\alpha\gamma}.
$$

Это выражение можно переписать в виде

$$
\mathsf{L}_X g_{\alpha\beta} = \widetilde{\nabla}_\alpha X_\beta + \widetilde{\nabla}_\beta X_\alpha, \tag{3.125}
$$

где $\widetilde{\nabla}_\alpha$ – ковариантная производная со связностью Леви–Чивиты (раздел 7.2). □

Пусть $T \in \mathcal{T}^r_s(\mathbb{M})$ и $X \in \mathcal{X}(\mathbb{M})$ – произвольное тензорное поле типа (r,s) и полное векторное поле на многообразии $\mathbb{M}$. Полное векторное поле порождает однопараметрическую группу преобразований $s(t,x)$. Если значение тензорного поля $T(s)$ в точке $s(t,x)$ равно $\tilde{s}(t,x)T(x)$, то мы говорим, что тензорное поле $T(x)$ *инвариантно* при действии однопараметрической группы преобразований. Из определения производной Ли следует

Предложение 3.14.2. *Тензорное поле $T \in \mathcal{T}^r_s(\mathbb{M})$ инвариантно относительно однопараметрической группы преобразований $s(t,x)$, порожденной векторным полем X, для всех t тогда и только тогда, когда производная Ли равна нулю, $\mathsf{L}_X T = 0$.*

Глава 4

Дифференциальные формы и интегрирование

В дифференциальной геометрии и приложениях важную роль играют дифференциальные формы. Они используются для определения связностей на расслоениях, интегрирования по многообразию, характеристических классов, когомологий де Рама, в гамильтоновом формализме и других областях. В настоящей главе будут даны определения и рассмотрены основные свойства дифференциальных форм. Будут также сформулированы две фундаментальные теоремы: теорема Дарбу и формула Стокса, играющие исключительно важную роль в приложениях.

4.1 Внешняя алгебра

Рассмотрим векторное и сопряженное к нему пространства $\mathbb{V}$ и $\mathbb{V}^*$. В разделе 2.3.3 над произвольным векторным пространством, например, над $\mathbb{V}^*$, была построена тензорная алгебра $\otimes\mathbb{V}^*$. В этой алгебре существуют двусторонние идеалы, и соответствующие фактор пространства также представляют собой алгебры. Наиболее важными из этих фактор алгебр являются внешняя алгебра и алгебры Клиффорда. В настоящем разделе мы рассмотрим внешнюю алгебру $\Lambda(\mathbb{V})$, которая лежит в основе теории дифференциальных форм.

Рассмотрим n-мерное векторное пространство $\mathbb{V}$ над полем вещественных чисел с базисом e_a, $a = 1, \ldots, n$. Дуальный базис сопряженного пространства $\mathbb{V}^*$ обозначим через e^a, $e^a(e_b) = \delta^a_b$. В следующем разделе мы отождествим пространства $\mathbb{V}$ и $\mathbb{V}^*$ с касательным $\mathbb{T}_x(\mathbb{M})$ и кокасательным $\mathbb{T}^*_x(\mathbb{M})$ пространствами в фиксированной точке $x \in \mathbb{M}$ многообразия $\mathbb{M}$.

Начнем с длинного, но конструктивного определения внешней алгебры над векторным пространством $\mathbb{V}$. Сначала введем вспомогательные понятия.

Определение. Множество полностью антисимметричных ковариантных тензоров типа $(0, r)$, $0 \leq r \leq n$, называется множеством *форм степени* r над векторным пространством $\mathbb{V}$ и обозначается $\Lambda_r(\mathbb{V})$.

В данном определении антисимметрия формы означает следующее. Пусть $A \in \Lambda_r(\mathbb{V})$. Тогда ее значение на произвольном наборе векторов $X_1, \ldots, X_r \in \mathbb{V}$ антисимметрично относительно перестановки любой пары векторов

$$A(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_r) = -A(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_r), \quad 1 \leq i < j \leq r.$$

В компонентах это условие означает антисимметрию относительно перестановки любой пары индексов.

Множество форм $\Lambda_r(\mathbb{V}) \subset \otimes\mathbb{V}^*$ степени r является векторным пространством над полем вещественных чисел размерности C_n^r.

По определению, вещественные числа (скаляры) образуют множество 0-форм, $\Lambda_0(\mathbb{V}) = \mathbb{R}$. Отметим, что 1-формы $A \in \Lambda_1(\mathbb{V}) = \mathbb{V}^*$ (ковекторы или линейные формы) являются элементами сопряженного пространства и представляются в виде $A = e^a A_a$. Максимальная степень формы совпадает с размерностью векторного пространства n, поскольку для более высоких степеней по крайней мере два индекса будут совпадать, а это невозможно для антисимметричных тензоров.

Из определения тензора следует, что r-форма $A \in \Lambda_r(\mathbb{V})$ представляет собой полилинейное отображение r экземпляров векторного пространства $\mathbb{V}$ в вещественную прямую

$$A: \quad \underbrace{\mathbb{V} \times \ldots \times \mathbb{V}}_{r} \quad \to \quad \mathbb{R}.$$

Полилинейность означает, что значение r-формы на r векторах $X_1, \ldots, X_r \in \mathbb{V}$ линейно по каждому аргументу при фиксированных остальных.

Ввиду антисимметрии r-форм относительно перестановки индексов, они имеют меньше независимых компонент, чем ковариантные тензоры типа $(0, r)$. Ведем обозначение

$$e^{a_1} \wedge \ldots \wedge e^{a_r} := \sum_\sigma e^{\sigma(a_1} \otimes \ldots \otimes e^{a_r)} \operatorname{sgn} \sigma, \tag{4.1}$$

где символ $\otimes$ обозначает тензорное произведение, сумма берется по всем перестановкам $\sigma(a_1 \ldots a_r)$ индексов, а $\operatorname{sgn} \sigma = \pm 1$ обозначает знак перестановки. Здесь мы ввели знак внешнего умножения $\wedge$, которое будет определено ниже. Пока он рассматривается просто, как некоторый символ.

По-построению, выражение (4.1) является ковектором типа $(0, r)$ и антисимметрично относительно перестановки любой пары индексов:

$$e^{\sigma(a_1} \wedge \ldots \wedge e^{a_r)} = e^{a_1} \wedge \ldots \wedge e^{a_r} \operatorname{sgn} \sigma.$$

Поэтому множество ковекторов с упорядоченным набором индексов

$$e^{a_1} \wedge \ldots \wedge e^{a_r}, \qquad 1 \le a_1 < \ldots < a_r \le n, \tag{4.2}$$

образует базис в пространстве r-форм. Тогда произвольная форма степени r имеет вид

$$A = \sum_{a_1 < \ldots < a_r} e^{a_1} \wedge \ldots \wedge e^{a_r} A_{a_1 \ldots a_r}, \tag{4.3}$$

$$= \frac{1}{r!} e^{a_1} \wedge \cdots \wedge e^{a_r} A_{a_1 \ldots a_r}. \tag{4.4}$$

В разложении (4.3) при фиксированной выборке индексов суммирование по индексам отсутствует, а сумма подразумевается только по различным выборкам. В последнем выражении (4.4) суммирование проводится по всем значениям индексов, и поэтому введен компенсирующий множитель $1/r!$. Это часто бывает удобнее, поскольку запись r-формы A в виде (4.4) имеет инвариантный вид в отличие от (4.3).

Пример 4.1.1. Рассмотрим 2-форму $A = e^a \otimes e^b A_{ab}$, где $A_{ab} = -A_{ba}$. Тогда

$$A = \frac{1}{2} e^a \otimes e^b (A_{ab} - A_{ba}) = \frac{1}{2} e^a \wedge e^b A_{ab},$$

где

$$e^a \wedge e^b = -e^b \wedge e^a = e^a \otimes e^b - e^b \otimes e^a. \quad \square \tag{4.5}$$

В компонентах значение r-формы A на векторных полях $\{X_1, \ldots, X_r\}$ определяется сверткой компонент:

$$A(X_1, \ldots, X_r) = X_1^{a_1} \ldots X_r^{a_r} A_{a_1 \ldots a_r}.$$

Построим 2^n-мерное векторное пространство, равное прямой сумме форм всех степеней

$$\Lambda(\mathbb{V}) = \bigoplus_{r=0}^{n} \Lambda_r(\mathbb{V}). \tag{4.6}$$

Базис этого пространства имеет вид

$$1, \quad e^a, \quad e^{a_1} \wedge e^{a_2}, \quad \ldots, \quad e^{a_1} \wedge \ldots \wedge e^{a_r}, \ldots, \quad e^1 \wedge \ldots \wedge e^n, \tag{4.7}$$

где последовательность $a_1 \ldots a_r$ является упорядоченным набором различных индексов, расположенных в порядке возрастания $1 \le a_1 < \cdots < a_r \le n$. Первый базисный вектор является единицей, поскольку поле вещественных чисел $\mathbb{R}$ является подпространством рассматриваемого векторного пространства. Последний базисный вектор единственен, т.к. содержит ровно n индексов. Элементы из $\Lambda(\mathbb{V})$ представимы в виде

$$A = \sum_{r=0}^{n} \frac{1}{r!} e^{a_1} \wedge \cdots \wedge e^{a_r} A_{a_1 \ldots a_r}, \tag{4.8}$$

где подразумевается суммирование по всем значениям индексов.

Введем на множестве форм *внешнее умножение*, отображающее $\Lambda(\mathbb{V}) \to \Lambda(\mathbb{V})$. Пусть A и B – формы фиксированных степеней r и s, соответственно. Их тензорное произведение $A \otimes B$ является ковариантным тензором типа $(0, r+s)$, однако не будет формой, т.к. не будет антисимметрично относительно всех перестановок индексов. Чтобы исправить ситуацию после тензорного умножения форм необходимо произвести полную антисимметризацию. Тогда внешнее умножение является отображением $\Lambda_r \times \Lambda_s \to \Lambda_{r+s}$.

Определение. Внешним произведением двух форм $A \in \Lambda_r(\mathbb{V})$ и $B \in \Lambda_s(\mathbb{V})$ называется форма $C \in \Lambda_{r+s}(\mathbb{V})$, построенная по правилу:

$$A \wedge B = \frac{1}{(r+s)!} e^{a_1} \wedge \cdots \wedge e^{a_{r+s}} C_{a_1 \ldots a_{r+s}}, \tag{4.9}$$

где коэффициенты внешнего произведения равны

$$\begin{aligned} C_{a_1 \ldots a_{r+s}} &= \frac{1}{r!s!} \sum_{\sigma} A_{\sigma(a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s})} \operatorname{sgn} \sigma, \\ &= C^r_{r+s} A_{[a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s}]}. \end{aligned} \tag{4.10}$$

Здесь $\sigma(a_1 \ldots a_r a_{r+1} \ldots a_{r+s})$ обозначает перестановку σ всех индексов $(a_1 \ldots a_{r+s})$, $\operatorname{sgn} \sigma$ – знак перестановки, суммирование ведется по всем перестановкам, а квадратные скобки обозначают антисимметризацию всех индексов. $\square$

Формулу внешнего произведения (4.9) можно переписать в эквивалентном виде

$$A \wedge B = \frac{1}{r!s!} e^{a_1} \wedge \cdots \wedge e^{a_{r+s}} A_{a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s}}. \tag{4.11}$$

Замечание. Множитель $1/(r+s)!$ в (4.9) связан с тем, что суммирование проводится по всем значениям индексов. Необходимость введения множителя $1/r!s!$ в (4.10) будет ясна из дальнейшего. □

Пример 4.1.2. Внешнее умножение произвольной r-формы $A \in \Lambda_r(\mathbb{V})$ на 0-форму (число) $f \in \mathbb{R} = \Lambda_0(\mathbb{V})$ сводится к обычному умножению компонент формы A на число,

$$f \wedge A = \frac{1}{r!} e^{a_1} \wedge \cdots \wedge e^{a_r} (f A_{a_1 \ldots a_r}). \qquad \square$$

Пример 4.1.3. Внешнее произведение двух 1-форм $A = e^a A_a$ и $B = e^a B_a$ равно

$$A \wedge B = \frac{1}{2} e^a \wedge e^b (A_a B_b - A_b B_a) = e^a \wedge e^b A_a B_b. \quad \square \tag{4.12}$$

Пример 4.1.4. Внешнее произведение 1-формы $A = e^a A_a$ на 2-форму $B = \frac{1}{2} e^a \wedge e^b B_{ab}$ равно

$$\begin{aligned} A \wedge B &= \frac{1}{6} e^a \wedge e^b \wedge e^c (A_a B_{bc} + A_b B_{ca} + A_c B_{ab}) \\ &= \frac{1}{2} e^a \wedge e^b \wedge e^c A_a B_{bc}. \quad \square \end{aligned} \tag{4.13}$$

Если суммарная степень форм превосходит размерность векторного пространства $r + s > n$, то внешнее произведение этих форм дает нуль.

Можно проверить, что внешнее умножение обладает следующими свойствами:

$$A \wedge B \ \text{ линейно по } A \text{ и } B, \tag{4.14}$$

$$A \wedge B = (-1)^{rs} B \wedge A, \tag{4.15}$$

$$(A \wedge B) \wedge C = A \wedge (B \wedge C). \tag{4.16}$$

Первые два свойства очевидны. Из второго свойства (4.15) следует, что, если A – форма нечетной степени, то $A \wedge A = 0$.

Третье свойство доказывается прямым вычислением. Для трех форм, степеней r, s и t, соответственно, имеем

$$\begin{aligned} &(A \wedge B) \wedge C = A \wedge (B \wedge C) = A \wedge B \wedge C \\ &= \frac{1}{(r+s+t)!} e^{a_1} \wedge \ldots \wedge e^{a_{r+s+t}} \frac{1}{r!s!t!} \sum_\sigma \mathsf{sgn}\,(\sigma) A_{\sigma(a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s}} C_{a_{r+s+1} \ldots a_{r+s+t})} \\ &= \frac{1}{r!s!t!} e^{a_1} \wedge \cdots \wedge e^{a_{r+s+t}} A_{a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s}} C_{a_{r+s+1} \ldots a_{r+s+t}}. \end{aligned}$$

Последнее выражение симметрично относительно перестановок r, s и t, что соответствует ассоциативности внешнего умножения и определяют выбор коэффициента в (4.10).

Определение. Форма фиксированной степени называется *однородной*. Прямая сумма форм разных степеней называется *неоднородной*. □

Внешнее умножение (4.9), определенное для однородных форм, продолжается на неоднородные формы общего вида (4.8) по линейности. Тем самым множество форм $\Lambda(\mathbb{V})$ над произвольным векторным пространством $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, с операцией внешнего умножения представляет собой ассоциативную алгебру над полем вещественных чисел.

Пример 4.1.5. Пусть A, B, C, D – четыре однородные формы различных степеней, тогда

$$(A \oplus B) \wedge (C \oplus D) := (A \wedge C) \oplus (B \wedge C) \oplus (A \wedge D) \oplus (B \wedge D). \qquad \square$$

Определение. Множество форм вида (4.8) с внешним умножением (4.9), продолженным на формы общего вида по линейности, называется *внешней алгеброй* $\Lambda(\mathbb{V})$ над векторным пространством $\mathbb{V}$, $\dim \mathbb{V} = n$. $\square$

Эта алгебра ассоциативна, градуирована, антикоммутативна и содержит единичный элемент $1 \in \Lambda_0(\mathbb{V})$. Внешняя алгебра является алгеброй Грассмана с образующими e^a (частным случаем алгебр Клиффорда).

В обозначении базисных векторов внешней алгебры (4.2) был использован знак внешнего умножения, которое было определено позже. Покажем, что это обозначение обосновано. Рассмотрим две базисные 1-формы $A^a = e^a$ и $B^b = e^b$. Тогда их внешнее произведение равно

$$A^a \wedge B^b = e^a \wedge e^b. \tag{4.17}$$

Более общо, для двух базисных форм $A^{a_1 \ldots a_r} = e^{a_1} \wedge \ldots \wedge e^{a_r}$ и $B^{b_1 \ldots b_s} = e^{b_1} \wedge \ldots \wedge e^{b_s}$ имеем

$$A^{a_1 \ldots a_r} \wedge B^{b_1 \ldots b_s} = e^{a_1} \wedge \ldots \wedge e^{a_r} \wedge e^{b_1} \wedge \ldots \wedge e^{b_s}.$$

Это оправдывает обозначение для базиса (4.1) и выбор коэффициентов для внешнего умножения (4.10).

Замечание. Для определения ассоциативного умножения в пространстве форм $\Lambda(\mathbb{V})$ выбор коэффициента в определении внешнего умножения не является единственным. При определении внешнего умножения (4.9), (4.10) можно было бы выбрать другой коэффициент в (4.10):

$$C_{a_1 \ldots a_{r+s}} = \frac{1}{(r+s)!} \sum_\sigma \operatorname{sgn} \sigma A_{\sigma(a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s})}. \tag{4.18}$$

В этом случае произведение трех форм

$$\begin{aligned}(A \wedge B) \wedge C &= A \wedge (B \wedge C) = A \wedge B \wedge C \\ &= \frac{1}{(r+s+t)!} e^{a_1} \wedge \cdots \wedge e^{a_{r+s+t}} A_{a_1 \ldots a_r} B_{a_{r+1} \ldots a_{r+s}} C_{a_{r+s+1} \ldots a_{r+s+t}}\end{aligned}$$

также симметрично относительно перестановок r, s, t и, значит, ассоциативно. Однако в произведении базисных векторов (4.17) появился бы дополнительный множитель. $\square$

К определению внешней алгебры можно подойти с другой, более абстрактной, точки зрения. Рассмотрим две произвольные 1-формы $A, B \in \Lambda_1(\mathbb{V}) = \mathbb{V}^*$. Тогда для их тензорного произведения справедливо тождество

$$A \otimes B = A \wedge B + \frac{1}{2}\Big[(A+B) \otimes (A+B) - A \otimes A - B \otimes B\Big]. \tag{4.19}$$

Выражение в квадратных скобках является элементом тензорной алгебры, и каждое слагаемое представляет собой тензорное произведение одинаковых 1-форм. Рассмотрим элементы более общего вида $X \otimes A \otimes A \otimes Y$, где $A \in \mathbb{V}^*$ – линейная форма, а $X, Y \in \otimes\mathbb{V}^*$ – произвольные элементы тензорной алгебры. Нетрудно проверить, что

элементы такого вида образуют двусторонний идеал $\mathbb{I}$ в тензорной алгебре $\otimes\mathbb{V}^*$, которая рассматривается, как кольцо по отношению к сложению и тензорному умножению. Тогда из представления (4.19) следует, что внешняя алгебра является факторпространством

$$\Lambda(\mathbb{V}) = \otimes\mathbb{V}^*/\mathbb{I}.$$

То есть два элемента $A, B \in \otimes\mathbb{V}^*$ являются эквивалентными $A \sim B$, если $A = B + C$, где $C \in \mathbb{I}$. Выражение в квадратных скобках в (4.19) принадлежит идеалу, и, значит,

$$A \otimes B \sim A \wedge B.$$

Таким образом каждый элемент внешней алгебры взаимно однозначно определяет класс эквивалентности в тензорной алгебре.

Сформулируем четыре утверждения, доказательства которых сводятся к алгебраическим выкладкам.

Предложение 4.1.1. *Для того, чтобы набор 1-форм $\{A_i\}$, $i = 1, \ldots, \mathrm{N}$, был линейно зависим, необходимо и достаточно, чтобы выполнялось равенство*

$$A_1 \wedge \ldots \wedge A_{\mathrm{N}} = 0. \tag{4.20}$$

Предложение 4.1.2. *Пусть $\{A_i\}$ и $\{B_i\}$, $i = 1, \ldots, \mathrm{N}$, – два набора 1-форм таких, что выполнено равенство*

$$\sum_{i=1}^{\mathrm{N}} A_i \wedge B_i = 0.$$

Если 1-формы $\{A_i\}$ линейно независимы, то 1-формы $\{B_i\}$ можно выразить в виде линейной комбинации 1-форм $\{A_i\}$

$$B_i = \sum_{j=1}^{\mathrm{N}} C_{ij} A_j,$$

причем $C_{ij} = C_{ji}$.

Предложение 4.1.3. *Пусть $\{A_i\}$, $i = 1, \ldots, \mathrm{N}$, – набор N линейно независимых 1-форм и $B \in \Lambda_s$. Для того, чтобы s-форму B можно было представить в виде*

$$B = A_1 \wedge C_1 + \ldots + A_{\mathrm{N}} \wedge C_{\mathrm{N}},$$

где $\{C_i\}$ – некоторый набор $(s-1)$-форм, необходимо и достаточно, чтобы

$$A_1 \wedge \ldots \wedge A_{\mathrm{N}} \wedge B = 0.$$

Замечание. При $\mathrm{N} + s > n$ эта теорема тривиально выполняется. □

Предложение 4.1.4. *Пусть $\{A_i, B_i\}$ и $\{A'_i, B'_i\}$, $i = 1, \ldots, \mathrm{N}$, – два набора по $2\mathrm{N}$ 1-форм. Если 1-формы $\{A_i, B_i\}$ линейно независимы и выполнено равенство*

$$\sum_{i=1}^{\mathrm{N}} A_i \wedge B_i = \sum_{i=1}^{\mathrm{N}} A'_i \wedge B'_i,$$

то набор 1-форм $\{A'_i, B'_i\}$ также линейно независим и представляется в виде линейных комбинаций 1-форм $\{A_i, B_i\}$.

Пусть некоторое семейство 1-форм $\{A_i\}$, $i = 1, \ldots, \mathrm{N}$, выражается в виде линейной комбинации 1-форм $\{B_i\}$,

$$A_i = M_i{}^j B_j,$$

где $M_i{}^j$ – некоторая квадратная $\mathrm{N}\times\mathrm{N}$ матрица. Тогда нетрудно проверить следующую формулу

$$A_1 \wedge \ldots \wedge A_{\mathrm{N}} = \det(M_i{}^j) B_1 \wedge \ldots \wedge B_{\mathrm{N}}.$$

Определение. Пусть $A \in \Lambda_r(\mathbb{V})$ – внешняя r-форма, тогда внешнее произведение

$$A^s := \underbrace{A \wedge \ldots \wedge A}_{s}$$

называется *внешней степенью* r-формы A.

На множестве 2-форм $\Lambda_2(\mathbb{V})$ можно определить *внешнюю экспоненту* $\widehat{\exp}$, которая каждой 2-форме $A \in \Lambda_2(\mathbb{V})$ ставит в соответствие прямую сумму форм четной степени по правилу

$$\widehat{\exp} A = 1 \oplus A \oplus \frac{1}{2} A \wedge A \oplus \frac{1}{3!} A \wedge A \wedge A \oplus \ldots \quad \square \tag{4.21}$$

Ряд (4.21) содержит конечное число слагаемых, т.к. максимальное число сомножителей в последнем отличном от нуля слагаемом не может превосходить $n/2$. Внешняя экспонента обладает следующими свойствами:

$$\widehat{\exp} A \wedge \widehat{\exp} B = \widehat{\exp}(A + B), \qquad \forall A, B \in \Lambda_2(\mathbb{V}).$$
$$\frac{d}{dt} \widehat{\exp}(tA) = A \wedge \widehat{\exp}(tA) = \widehat{\exp}(xA) \wedge A.$$

При доказательстве этих формул использовано равенство $A \wedge B = B \wedge A$, справедливое для форм четной степени.

Определение. В пространстве векторов и r-форм введем *внутреннее умножение* как билинейное отображение

$$\mathsf{i} : \quad \mathbb{V} \times \Lambda_r(\mathbb{V}) \ni (X, A) \;\mapsto\; \mathsf{i}_X A \in \Lambda^{r-1}(\mathbb{V}).$$

Оно определяется следующим образом. Пусть $X \in \mathbb{V}$ и $A \in \Lambda_r(\mathbb{V})$. Положим

$$\mathsf{i}_X A(X_1, \ldots, X_{r-1}) = A(X, X_1, \ldots, X_{r-1}). \tag{4.22}$$

Другими словами, компоненты вектора X нужно просто свернуть с первым индексом r-формы. Правая часть выражения (4.22) $(r-1)$-линейна и антисимметрична, как функция $X_1, \ldots, X_{r-1}$, а также билинейна по X и A. Удобно считать, что $\mathsf{i}_X A = 0$, если $A \in \Lambda_0$ – форма нулевой степени. Внутреннее умножение обозначают также $\mathsf{i}_X A = X \rfloor A$. Определение (4.22) можно переписать в эквивалентном виде

$$\mathsf{i}_X A = X \rfloor A = \frac{1}{(r-1)!} e^{a_1} \wedge \ldots \wedge e^{a_{r-1}} X^b A_{b a_1 \ldots a_{r-1}}. \quad \square \tag{4.23}$$

Внутреннее умножение линейно

$$\mathsf{i}_{X+Y} = \mathsf{i}_X + \mathsf{i}_Y,$$
$$\mathsf{i}_{aX} = a\, \mathsf{i}_X, \qquad a \in \mathbb{R},$$

и антисимметрично

$$\mathsf{i}_X \mathsf{i}_Y = -\mathsf{i}_Y \mathsf{i}_X \quad \Leftrightarrow \quad \mathsf{i}_X \mathsf{i}_X = 0,$$

для всех $X, Y \in \mathbb{V}$.

Нетрудно проверить, что внутреннее умножение удовлетворяет правилу Лейбница

$$\mathsf{i}_X(A \wedge B) = \mathsf{i}_X A \wedge B + (-1)^r A \wedge \mathsf{i}_X B, \qquad A \in \Lambda_r(\mathbb{V}).$$

Поэтому внутреннее умножение называют также *внутренним дифференцированием*.

Внутреннее умножение удобно использовать для характеристики внешней алгебры, построенной на фактор пространстве $\mathbb{V}/\mathbb{U}$, где $\mathbb{U} \subset \mathbb{V}$. А именно, внешняя алгебра $\Lambda(\mathbb{V}/\mathbb{U})$ состоит из тех форм $A \in \Lambda(\mathbb{V})$, для которых выполнено равенство $\mathsf{i}_X A = 0$ для всех $X \in \mathbb{U}$. Верно и обратное утверждение. Пусть задано подпространство $\mathbb{U} \subset \mathbb{V}$. Тогда множество векторов X, удовлетворяющих условию $\mathsf{i}_X A = 0$ для всех $A \in \Lambda(\mathbb{V}/\mathbb{U})$ образует подпространство, которое совпадает с $\mathbb{U}$.

Пусть f – линейное отображение векторных пространств

$$f: \ \mathbb{V} \ \to \ \mathbb{W},$$

и пусть на векторном пространстве-мишени $\mathbb{W}$ заданы две формы $A, B \in \Lambda(\mathbb{W})$, тогда для возврата отображения справедливо равенство

$$f^*(A \wedge B) = (f^* A) \wedge (f^* B) \ \in \Lambda(\mathbb{V}).$$

4.2 Дифференциальные формы

Рассмотрим многообразие $\mathbb{M}$ размерности $\mathsf{dim}\,\mathbb{M} = n$. В каждой точке многообразия мы имеем касательное и кокасательное векторные пространства. Построим над касательным векторным пространством $\mathbb{V} = \mathbb{T}_x(\mathbb{M})$ с координатным базисом ∂_α в каждой точке $x \in \mathbb{M}$ векторное пространство форм степени r: $\Lambda_{r,x}(\mathbb{M})$, как это было сделано в предыдущем разделе. Объединение форм степени r по всем точкам многообразия $\bigcup_{x\in\mathbb{M}} \Lambda_{r,x}(\mathbb{M})$ приведет к векторному расслоению, сечениями которого являются полностью антисимметричные ковариантные тензорные поля ранга r (дифференциальные формы). Очевидно, что множество дифференциальных форм фиксированной степени с поточечным сложением и умножением на числа образует бесконечномерное линейное пространство. Прямая сумма по r множества всех дифференциальных форм даст внешнюю алгебру дифференциальных форм $\Lambda(\mathbb{M})$ на многообразии $\mathbb{M}$, в которой все операции (умножение на числа, сложение и внешнее умножение) определены поточечно. Ясно, что все свойства внешнего умножения без изменений переносятся на внешнюю алгебру дифференциальных форм, и мы не будем их переписывать. Отметим изменение обозначений и терминологии.

Определение. Множество полностью антисимметричных ковариантных тензорных полей типа $(0, r)$, $0 \le r \le n$ называется множеством *дифференциальных форм степени* r на многообразии $\mathbb{M}$ и обозначается $\Lambda_r(\mathbb{M}) \subset \mathcal{T}_r(\mathbb{M})$. □

В дальнейшем дифференциальные формы на многообразии $\mathbb{M}$ степени r мы, для краткости, будем называть просто r-формами.

По определению, гладкие функции на многообразии образуют множество 0-форм, $\Lambda_0(\mathbb{M}) = \mathcal{C}^\infty(\mathbb{M})$. 1-формы $A \in \Lambda_1(\mathbb{M}) = \mathcal{T}_1(\mathbb{M})$ (ковариантные векторные поля = сечения кокасательных расслоений) представляются в виде

$$A(x) = dx^\alpha A_\alpha(x),$$

где дифференциалы dx^α представляют собой координатный базис кокасательного пространства, который дуален к ∂_α. Дифференциальные 1-формы называются также *формами Пфаффа*. Максимальная степень формы совпадает с размерностью многообразия n, поскольку для более высоких степеней по крайней мере два индекса будут совпадать, а это невозможно для антисимметричных тензоров.

Замечание. Напомним, что ∂_α это сокращенное обозначение набора векторных полей $(\partial_\alpha)_p \in \mathcal{X}(\mathbb{U})$, $p \in \mathbb{U}$, на карте $(\mathbb{U}, \varphi)$ многообразия $\mathbb{M}$ (см. раздел 3.6.4), а $dx^\alpha = (dx^\alpha)_p$ – дуальный базис 1-форм, $dx^\alpha(\partial_\beta) = \delta^\alpha_\beta$. □

Как и для векторных пространств, алгебра дифференциальных форм равна прямой сумме форм всех степеней

$$\Lambda(\mathbb{M}) = \bigoplus_{r=0}^{n} \Lambda_r(\mathbb{M}). \tag{4.24}$$

Координатный базис этого пространства имеет вид

$$1, \quad dx^\alpha, \quad dx^{\alpha_1} \wedge dx^{\alpha_2}, \quad \ldots, \quad dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r}, \ldots, \quad dx^1 \wedge \cdots \wedge dx^n, \tag{4.25}$$

где индексы упорядочены по возрастанию $1 \le \alpha_1 < \ldots \alpha_r \le r$. Элементы внешней алгебры $\Lambda(\mathbb{M})$, в соответствии с разложением (4.24), представимы в виде

$$A = \sum_{r=0}^{n} \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r}, \tag{4.26}$$

где подразумевается суммирование по всем значениям индексов. Форма A называется *дифференцируемой*, если все антисимметричные тензорные поля $A_{\alpha_1 \ldots \alpha_r}(x)$ (компоненты формы) дифференцируемы. Если не оговорено противное, мы будем предполагать, что все компоненты форм гладко зависят от точки многообразия.

Замечание. Запись дифференциальной формы в виде (4.26) имеет смысл только в определенной карте многообразия, где определен координатный базис кокасательного пространства dx^α. Однако она универсальна, т.к. инвариантна относительно преобразований координат, поскольку в точках пересечения двух карт с координатами x^α и $x^{\alpha'}$ имеем

$$dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r} = dx^{\alpha'_1} \wedge \cdots \wedge dx^{\alpha'_r} A_{\alpha'_1 \ldots \alpha'_r}.$$

Запись дифференциальных форм в виде (4.26) является общепринятой. □

Введем несколько новых понятий, которые являются полезными с точки зрения дифференциальной геометрии. Конечно, часть из них можно было бы ввести и для дифференциальных форм над векторным пространством еще в предыдущем разделе.

Можно рассматривать дифференциальные формы со значениями в произвольном векторном пространстве $\mathbb{W}$. Пусть e_i – базис векторного пространства $\mathbb{W}$. Тогда r-форма A со значениями в $\mathbb{W}$ имеет вид $A = A^i e_i$. При каждом значении индекса i коэффициенты этого разложения представляют собой r-формы на $A^i \in \Lambda_r(\mathbb{M})$. По индексу i эти формы преобразуются, как элементы векторного пространства $\mathbb{W}$.

Пример 4.2.1. Произвольную r-форму на многообразии $\mathbb{M}$

$$A = \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r} \in \Lambda_r(\mathbb{M}),$$

можно поточечно рассматривать, как внешнее произведение базисной r-формы $e_i = dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_r}$ на 0-форму $A_{\alpha_1 \ldots \alpha_r}$ со значениями в векторном пространстве $\mathbb{W} = \mathbb{T}^*_x(\mathbb{M}) \wedge \ldots \wedge \mathbb{T}^*_x(\mathbb{M})$. При этом индекс i соответствует набору индексов $\alpha_1 \ldots \alpha_r$. □

Пусть на многообразии $\mathbb{M}$ помимо форм заданы векторные поля. Тогда значение произвольной r-формы A на r векторных полях $X_1, \ldots, X_r \in \mathcal{X}(\mathbb{M})$ равно свертке компонент

$$A(X_1, \ldots, X_r) = X_1^{\alpha_1} \ldots X_r^{\alpha_r} A_{\alpha_1 \ldots \alpha_r}.$$

(Данное определение не зависит от выбора карты и потому корректно.) Отсюда следует, что произвольная r-форма в каждой точке $x \in \mathbb{M}$ определяет полилинейное отображение ($\mathcal{C}^\infty$-модуль): $\Big(\mathbb{T}_x(\mathbb{M})\Big)^r \to \mathbb{R}$, которое антисимметрично относительно перестановки аргументов.

Пример 4.2.2. Значение 1-формы $A = dx^\alpha A_\alpha$ на векторном поле $X = X^\alpha \partial_\alpha$ определяет функцию, которая равна сумме компонент: $A(X) = X^\alpha A_\alpha$. □

Пусть A_i и X_i, $i = 1, \ldots, r \le n$ – произвольный набор 1-форм и векторных полей на $\mathbb{M}$. Тогда из определения внешнего умножения следует, что значение внешнего произведения 1-форм на векторных полях равно определителю матрицы с элементами $A_i(X_j)$

$$(A_1 \wedge \cdots \wedge A_r)(X_1 \ldots X_r) = \det\left(A_i(X_j)\right).$$

Пусть две 1-формы A и B линейно зависимы, т.е. $A = fB$, где $f \in \Lambda_0(\mathbb{M})$, тогда их внешнее произведение, очевидно, равно нулю. Верно также и обратное утверждение, из которого следует, что, если $A \wedge B \neq 0$ в некоторой области $\mathbb{U} \subset \mathbb{M}$, то 1-формы линейно независимы в каждой точке $\mathbb{U}$. Более общо, справедливо следующее утверждение.

Предложение 4.2.1. *Для того, чтобы линейные формы $e^i = dx^\alpha e_\alpha{}^i$, $i = 1, \ldots, r \le n$ были линейно независимы в области $\mathbb{U} \subset \mathbb{M}$, необходимо и достаточно, чтобы в каждой точке этой области $e^1 \wedge \ldots \wedge e^r \neq 0$.*

Дифференциальные формы часто бывает удобно рассматривать в неголономном базисе. Выберем в качестве базиса кокасательного пространства 1-формы $e^a = dx^\alpha e_\alpha{}^a$, $a = 1, \ldots, n$, где $e_\alpha{}^a(x)$ – поле репера, тогда соответствующий базис r-форм имеет вид

$$e^{a_1} \wedge \cdots \wedge e^{a_r}, \qquad 1 \le a_1 < \cdots < a_r \le n. \tag{4.27}$$

Мы используем для неголономного базиса те же обозначения, что и для базиса форм над векторным пространством (4.2). Разница заключается в том, что теперь базис $e^a(x)$ зависит от точки многообразия, и это всегда ясно из контекста.

Для того, чтобы 1-формы e^a действительно образовывали базис, необходимо и достаточно, чтобы

$$\det e_\alpha{}^a \neq 0 \quad \Leftrightarrow \quad e^1 \wedge \ldots \wedge e^n \neq 0$$

во всех точках многообразия $\mathbb{M}$, что совпадает с определением репера. Компоненты разложения r-формы

$$A = \frac{1}{r!} e^{a_1} \wedge \cdots \wedge e^{a_r} A_{a_1 \ldots a_r}$$

по этому базису связаны с компонентами в голономном базисе простым соотношением

$$A_{a_1 \ldots a_r} = e^{\alpha_1}{}_{a_1} \ldots e^{\alpha_r}{}_{a_r} A_{\alpha_1 \ldots \alpha_r},$$

где $e^\alpha{}_a$ – матрица обратного репера: $e^\alpha{}_a e_\alpha{}^b = \delta_a^b$.

4.3 Внешнее дифференцирование

Во внешней алгебре дифференциальных форм $\Lambda(\mathbb{M})$ на многообразии $\mathbb{M}$ определим *внешнее дифференцирование* как линейное отображение

$$d: \quad \Lambda_r(\mathbb{M}) \ \to \ \Lambda_{r+1}(\mathbb{M}) \tag{4.28}$$

множества r-форм в множество $(r+1)$-форм. Начнем с локального определения. Пусть в некоторой карте $(\mathbb{U}, \varphi)$ задана r-форма

$$A = \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r} \in \Lambda_r(\mathbb{U}).$$

Положим

$$\begin{aligned} dA &= \frac{1}{r!} dA_{\alpha_1 \ldots \alpha_r} \wedge dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} \\ &= \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_{r+1}} \partial_{\alpha_1} A_{\alpha_2 \ldots \alpha_{r+1}}. \end{aligned} \tag{4.29}$$

Здесь под $dA_{\alpha_1 \ldots \alpha_r} = dx^\alpha \partial_\alpha A_{\alpha_1 \ldots \alpha_r}$ понимается обычный дифференциал. Внешнее дифференцирование обладает следующими свойствами:

$$\text{1)} \quad \text{если } \ f \ - \ \text{0-форма, то } \ df = dx^\alpha \partial_\alpha f, \tag{4.30}$$

$$\text{2)} \quad \text{если } \ A \ - \ r\text{-форма, то } \ d(A \wedge B) = dA \wedge B + (-1)^r A \wedge dB, \tag{4.31}$$

$$\text{3)} \quad d(dA) = 0 \ \text{ для любой формы } A. \tag{4.32}$$

Первые два свойства легко проверяются, исходя из определения (4.29). Доказательство третьего свойства сводится к тому, что вторая частная производная $\partial^2_{\alpha\beta}$ симметрична по индексам и дает нуль при свертке с антисимметричным базисным вектором.

Формула для внешнего дифференцирования (4.29) является прямым следствием свойств 1)–3). Действительно, представив r-форму как внешнее произведение 0-формы $A_{\alpha_1 \ldots \alpha_n}$ со значениями в векторном пространстве на базисную r-форму $dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_n}$ сразу приходим к (4.29). Поэтому она задает единственное дифференцирование во внешней алгебре, удовлетворяющее условиям 1)–3).

Теорема 4.3.1 (Внешнее дифференцирование). *Существует единственное линейное отображение (4.28) такое, что выполнены условия (4.30)–(4.32). Оно называется внешним дифференцированием.*

Доказательство. Выше мы доказали существование и единственность внешнего дифференцирования в определенной карте. Осталось доказать, что существование и единственность не зависят от выбора карты. Действительно, компоненты в разложении внешней производной от r-формы (4.29)

$$\partial_{[\alpha_1} A_{\alpha_2 \ldots \alpha_{r+1}]},$$

где квадратные скобки обозначают антисимметризацию всех индексов, являются компонентами полностью антисимметричного тензора типа $(0, r+1)$. Это утверждение нетривиально, потому что обычная частная производная от компонент тензора не дает тензора. Для 1-формы имеем следующий закон преобразования частной производной в области пересечения двух карт

$$\partial_{\alpha'} A_{\beta'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{\partial x^\beta}{\partial x^{\beta'}} \partial_\alpha A_\beta + \frac{\partial^2 x^\alpha}{\partial x^{\alpha'} \partial x^{\beta'}} A_\alpha.$$

Первое слагаемое соответствует тензорному закону преобразования, а второе – нет. Однако второе слагаемое исчезает при антисимметризации и мы получаем тензорный закон преобразования для внешней производной. Аналогично сокращаются дополнительные слагаемые для произвольной r-формы. Подробности доказательства приведены в [?]. □

Для практических вычислений внешней производной, которые проводятся в координатах, достаточно формулы (4.29).

Пример 4.3.1. Внешний дифференциал от 1-формы $A = dx^\alpha A_\alpha$ равен

$$dA = dA_\alpha \wedge dx^\alpha = \frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha A_\beta - \partial_\beta A_\alpha) = dx^\alpha \wedge dx^\beta \partial_\alpha A_\beta. \quad \square \tag{4.33}$$

Пример 4.3.2. Внешний дифференциал от 2-формы $B = \frac{1}{2} dx^\alpha \wedge dx^\beta B_{\alpha\beta}$ равен

$$\begin{aligned} dB &= \frac{1}{2} dB_{\alpha\beta} \wedge dx^\alpha \wedge dx^\beta = \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha B_{\beta\gamma} \\ &= \frac{1}{12} dx^\alpha \wedge dx^\beta \wedge dx^\gamma (\partial_\alpha B_{\beta\gamma} + \partial_\beta B_{\gamma\alpha} + \partial_\gamma B_{\alpha\beta} - \partial_\alpha B_{\gamma\beta} - \partial_\gamma B_{\beta\alpha} - \partial_\beta B_{\alpha\gamma}). \quad \square \end{aligned} \tag{4.34}$$

Формулу внешнего дифференцирования легко запомнить с помощью следующего мнемонического правила. Сравнивая последнее выражение в (4.29) с формулой для внешнего умножения (4.11), можно записать

$$dA = \partial \wedge A.$$

Из определения 4.29 следует, что внешний дифференциал можно записать с помощью ковариантной производной, определяемой символами Кристоффеля или другой произвольной *симметричной* аффинной связностью

$$dA = \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_{r+1}} \widetilde{\nabla}_{\alpha_1} A_{\alpha_2 \ldots \alpha_{r+1}}. \tag{4.35}$$

В противном случае возникнут слагаемые с тензором кручения.

Приведем некоторые свойства внешнего дифференцирования. С помощью прямых вычислений доказывается формула

$$2dA(X,Y) = X\Big(A(Y)\Big) - Y\Big(A(X)\Big) - A\Big([X,Y]\Big), \tag{4.36}$$

где A – произвольная 1-форма, а X, Y – любые векторные поля на $\mathbb{M}$. В общем случае справедлива

Теорема 4.3.2. *Пусть A – гладкая дифференциальная форма степени r и $X_1 \ldots X_{r+1}$ – гладкие векторные поля. Тогда значение формы dA на этих векторных полях может быть найдено по формуле*

$$\begin{aligned} dA(X_1, \ldots, X_{r+1}) = &\frac{1}{r+1} \sum_i (-1)^{i-1} X_i^\alpha \partial_\alpha A(X_1, \ldots, \breve{X}_i, \ldots, X_{r-1}) + \\ &+ \frac{1}{r+1} \sum_{i<j} (-1)^{i+j} A([X_i, X_j], X_1, \ldots, \breve{X}_i, \ldots, \breve{X}_j, \ldots, X_{r+1}), \end{aligned} \tag{4.37}$$

где символ $\breve{X}_i$ означает, что соответствующее векторное поле отсутствует.

Доказательство. Прямая проверка. □

Определение. Форма A степени $r \geq 1$ называется *точной*, если существует такая $(r-1)$-форма B, что

$$A = dB. \tag{4.38}$$

Форма A степени $r \geq 0$ называется *замкнутой*, если выполнено условие

$$dA = 0. \tag{4.39}$$

По-определению, точных 0-форм не существует. □

Предложение 4.3.1. *Множество всех замкнутых 0-форм (функций) на связном многообразии состоит из функций, которые постоянны на всем многообразии.*

Любая точная форма является замкнутой в силу свойства внешнего дифференцирования (4.32). Обратное утверждение в общем случае неверно. Для замкнутой формы представление (4.38) справедливо только локально, т.е. в некоторой окрестности произвольной точки (лемма Пуанкаре). При этом глобальное представление может не иметь места. В общем случае множество замкнутых форм больше множества точных форм. В дальнейшем будет показано, что отличие множества замкнутых и точных форм связано с топологическими свойствами многообразий.

Замечание. Если форма A точна, $A = dB$, и принадлежит классу $\mathcal{C}^k$, то это не значит, что форма B принадлежит классу $\mathcal{C}^{k+1}$, потому что в определении (4.29) содержится только часть частных производных. Например, если $B \in \Lambda_n(\mathbb{M})$ – форма максимальной степени и класса $\mathcal{C}^1$, то $A = dB = 0$, и, следовательно, принадлежит классу $\mathcal{C}^\infty$, а это ничего не говорит о степени гладкости формы B. В лемме Пуанкаре нам встретится ситуация, когда обе формы A и B одного класса гладкости. □

Пример 4.3.3. Пусть на компактном многообразии $\mathbb{M}$ задана гладкая функция $f \in \mathcal{C}^\infty(\mathbb{M})$. Тогда 1-форма $A = dx^\alpha \partial_\alpha f$ точна, замкнута и имеет по крайней мере два нуля. Наличие нулей следует из теоремы 2.3.1 о том, что на компактном многообразии всякая непрерывная функция принимает свое минимальное и максимальное значения, где все частные производные обращаются в нуль. Этот пример показывает, что нули замкнутых форм связаны с топологией многообразий. □

Предложение 4.3.2. *Внешнее произведение замкнутых форм замкнуто, а если одна из форм, кроме того, точна, то это произведение является точной формой.*

Доказательство. Прямая проверка. □

Приведем важную теорему, имеющую многочисленные применения в дифференциальной геометрии.

Теорема 4.3.3 (Лемма Пуанкаре). *Пусть $A \in \Lambda_r(\mathbb{R}^n)$, $r \geq 1$, – замкнутая r-форма, $dA = 0$, класса $\mathcal{C}^1$, задана в евклидовом пространстве $\mathbb{R}^n$. Тогда существуют $(r-1)$-формы $B \in \Lambda_{r-1}(\mathbb{R}^n)$, класса $\mathcal{C}^1$, такие, что $A = dB$. При этом все формы B отличаются друг от друга на точные формы $B_1 = B_2 + dC$, где $C \in \Lambda_{r-2}(\mathbb{R}^n)$, класса $\mathcal{C}^1$ при $r \geq 2$. При $r = 1$, 0-формы B определены с точностью до константы.*

Доказательство. Доказательство проводится по индукции по размерности евклидова пространства путем явного построения формы B с помощью интеграла вдоль кривой [**?**]. □

Эта теорема локальна, т.к. все евклидово пространство диффеоморфно конечной области в $\mathbb{R}^n$. В случае, когда многообразие $\mathbb{M}$ нетривиально и не покрывается одной картой, теорема справедлива для каждой карты, но не для всего $\mathbb{M}$. Отличие классов замкнутых и точных форм служит для глобальной характеристики многообразий и лежит в основе теории когомологий де Рама – одного из подходов к теории гомологий.

Замечание. Согласно лемме Пуанкаре, если форма A определена на области $\mathbb{U} \subset \mathbb{M}$, гомеоморфной евклидову пространству $\mathbb{R}^n$, то она всегда точна. Поэтому различие между точными и замкнутыми формами проявляется только для форм, заданных на всем многообразии $\mathbb{M}$. □

Пример 4.3.4. Поясним лемму Пуанкаре на примере 1-формы A и 0-формы f. В этом случае условие точности формы (4.38) сводится к уравнению

$$A_\alpha = \partial_\alpha f. \tag{4.40}$$

Из теории дифференциальных уравнений известно, что это уравнение локально разрешимо тогда и только тогда, когда

$$\partial_\alpha A_\beta - \partial_\beta A_\alpha = 0 \quad \text{или} \quad dA = 0. \tag{4.41}$$

Таким образом, замкнутость некоторой формы обеспечивает ее точность только локально, но не глобально.

Дифференциальная 1-форма $dx^\alpha A_\alpha$ на n-мерном многообразии задается n независимыми компонентами, а скалярное поле f – только одной. С другой стороны, если форма A точна, то она параметризуется только одной функцией, т.к. выполнено равенство (4.40). На первый взгляд, число уравнений на компоненты A_α в (4.41) равно C_n^2, что превышает разность в числе независимых компонент, которая равна $n-1$, при $n \geq 3$. Кажущееся противоречие заключается в наивном подсчете числа уравнений. Подсчитаем число независимых уравнений в (4.41). Общее число уравнений равно C_n^2, однако независимых уравнений меньше, поскольку имеется дифференциальное тождество $d^2A = 0$. Число этих тождеств равно C_n^3, однако среди них также есть зависимые в силу тождества $d^3A = 0$. Продолжая этот процесс до порядка $n-1$ (дальнейшая антисимметризация по индексам не имеет смысла) получим, что число независимых уравнений среди (4.41) равно

$$C_n^2 - C_n^3 + C_n^4 - \ldots (-1)^n C_n^n = C_n^1 - 1 = n - 1.$$

Это равенство следует из (28.3). То есть среди n компонент замкнутой 1-формы A (ковекторного поля) только одна является независимой, в качестве которой можно выбрать f. □

Пример 4.3.5. Продемонстрируем локальность леммы Пуанкаре и важность нетривиальных замкнутых форм для описания глобальных свойств многообразий. Рассмотрим евклидову плоскость $\mathbb{R}^2$ в декартовых координатах x, y. Тогда дифференциал полярного угла φ имеет вид (8.22)

$$A = \frac{xdy - xdy}{x^2 + y^2}.$$

1-форма A определена и замкнута, $dA = 0$, всюду, за исключением начала координат, и, значит, не удовлетворяет условиям леммы Пуанкаре. Отсюда следует, что

не существует дифференцируемой функции $\varphi(x,y)$, определенной на всей плоскости $\mathbb{R}^2$, такой, что $A = d\varphi$. Действительно, при обходе начала координат по замкнутому контуру один раз, функция φ получит приращение 2π. В то же время на любой односвязной области, не содержащей начала координат, такая функция существует. □

Предложение 4.3.3. *Если на многообразии задано векторное поле X, то производная Ли (см. раздел 3.14) от произвольной r-формы $A \in \Lambda_r(\mathbb{M})$ представляет собой симметричную комбинацию внешней и внутренней производной*

$$\mathsf{L}_X = d\,\mathsf{i}_X + \mathsf{i}_X d. \tag{4.42}$$

Это соотношение называется основной формулой гомотопии.

Доказательство. Достаточно показать, что правая часть этого выражения является дифференцированием, коммутирующим с внешней производной d, и что оно совпадает с производной Ли на функциях. Действительно, пусть $A \in \Lambda_r(\mathbb{M})$ и $B \in \Lambda_s(\mathbb{M})$, тогда нетрудно проверить равенство

$$(d\,\mathsf{i}_X + \mathsf{i}_X d)(A \wedge B) = (d\,\mathsf{i}_X + \mathsf{i}_X d)A \wedge B + A \wedge (d\,\mathsf{i}_X + \mathsf{i}_X d)B. \tag{4.43}$$

На функциях $f \in \Lambda_0$

$$(d\,\mathsf{i}_X + \mathsf{i}_X d)f = \mathsf{i}_X df = Xf = \mathsf{L}_X f.$$

Коммутативность $(d\,\mathsf{i}_X + \mathsf{i}_X d)$ с внешней производной легко проверяется. □

Формулу (4.43) можно переписать в виде

$$\mathsf{L}_X(A \wedge B) = \mathsf{L}_X A \wedge B + A \wedge \mathsf{L}_X B.$$

Предложение 4.3.4. *Пусть задан диффеоморфизм двух многообразий $h:\ \mathbb{M} \to \mathbb{N}$. Тогда возврат отображения*

$$h^*:\quad \mathbb{T}^*(\mathbb{N}) \to \mathbb{T}^*(\mathbb{M}),$$

действует на формы в обратную сторону и имеет следующие свойства.

1) Для произвольных форм $A, B \in \Lambda(\mathbb{N})$, заданных на многообразии $\mathbb{N}$, справедливо равенство

$$h^*(A \wedge B) = h^*A \wedge h^*B.$$

2) Возврат отображения коммутирует с внешней производной

$$h^* dA = d(h^*A), \tag{4.44}$$

где A – произвольная r-форма на многообразии $\mathbb{N}$.

3) Пусть задан набор векторных полей $X_1, \ldots, X_r$ на $\mathbb{M}$. Тогда

$$h^*A(X_1, \ldots, X_r) = A(h_*X_1, \ldots, h_*X_r),$$

где A – произвольная r-форма на $\mathbb{N}$.

Доказательство. Прямая проверка. □

Поскольку производная Ли от форм определяется возвратом отображения, то из свойства 1) вытекает, что производная Ли от произвольной r-формы коммутирует с внешним дифференцированием:

$$\mathsf{L}_X d = d\,\mathsf{L}_X.$$

Это свойство было использовано в доказательстве равенства (4.43).

4.4 Теорема Дарбу

Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана произвольная r-форма

$$A = \frac{1}{n!} dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r} \quad \in \Lambda_r(\mathbb{M}).$$

Зададим вопрос о том, каково минимальное число функций $f^1(x), \ldots, f^q(x)$ таких, что r-форма A представима в виде

$$A = \sum_{1 \le \alpha_1 < \ldots < \alpha_r \le q} df^{\alpha_1} \wedge \ldots \wedge df^{\alpha_r} A_{\alpha_1 \ldots \alpha_r}(f^1 \ldots f^q). \tag{4.45}$$

Ясно, что при $q = n$ любая форма уже имеет такой вид. С другой стороны, число функций не может быть меньше степени формы. Поэтому представляет интерес случай $r \le q < n$. Важность этого вопроса заключается в том, что функции $f^1, \ldots, f^q$ можно выбрать в качестве части координат некоторой новой координатной системы и тем самым привести форму к каноническому виду, в котором она будет зависеть не от всех координат, а только от их части. Ответ на поставленный вопрос дает теорема 4.46, сформулированная в настоящем разделе. Для 1- и 2-форм ответ дают теоремы Дарбу.

Для формулировки теорем мы введем новые понятия: ранг и класс дифференциальных форм. Начнем с форм над произвольным векторным пространством $\mathbb{V}$ с базисом e_α, $\alpha = 1, \ldots, n$. Дадим несколько определений.

Определение. Назовем r-форму $A \in \Lambda_r(\mathbb{V})$ *разложимой*, если существует r линейно независимых 1-форм $a^1, \ldots, a^r \in \Lambda_1(\mathbb{V})$ таких, что

$$A = a^1 \wedge \ldots \wedge a^r. \qquad \square$$

Ясно, что всякую r-форму можно представить в виде суммы разложимых r-форм, например, разложив по базису. Однако, не всякая r-форма разложима.

Пример 4.4.1. Рассмотрим четырехмерное векторное пространство $\mathbb{V}$ с базисом e_α, $\alpha = 1, 2, 3, 4$. Базис сопряженного пространства $\mathbb{V}^*$ обозначим через e^α. Рассмотрим 2-форму

$$A = e^1 \wedge e^2 + e^3 \wedge e^4 \quad \in \Lambda_2(\mathbb{V}).$$

Если форма A разложима, то существуют две 1-формы $a = a_\alpha e^\alpha$ и $b = b_\alpha e^\alpha$ такие, что выполнено условие

$$A = a \wedge b = e^\alpha \wedge e^\beta a_\alpha b_\beta = e^1 \wedge e^2 + e^3 \wedge e^4.$$

Отсюда вытекает следующая система уравнений на компоненты

$$\begin{aligned} a_1 b_2 &= 1, & a_2 b_3 &= 0, \\ a_1 b_3 &= 0, & a_2 b_4 &= 0, \\ a_1 b_4 &= 0, & a_3 b_4 &= 1. \end{aligned}$$

Не ограничивая общности, можно положить $a_1 = 1$. Тогда из трех уравнений в первой колонке следует $b_2 = 1$, $b_3 = b_4 = 0$. Подставив эти решения в остальные уравнения, приходим к противоречию. Следовательно, 2-форма A является неразложимой.

Рассмотрим другую 2-форму над тем же векторным пространством $\mathbb{V}$

$$A = e^1 \wedge e^2 + e^1 \wedge e^3 \quad \in \Lambda_2(\mathbb{V}).$$

Ее можно представить в виде $A = e^1 \wedge (e^2 + e^3)$. Отсюда следует, что эта форма разложима: $A = e^1 \wedge b$, где $b = e^2 + e^3 \in \Lambda_1(\mathbb{V})$. $\square$

Определение. Назовем подпространство $\mathbb{A}(A) \subset \mathbb{V}$ *ассоциированным* с формой A, если это наибольшее из тех подпространств $\mathbb{H} \subset \mathbb{V}$, для которых $A \in \Lambda_r(\mathbb{V}/\mathbb{H})$. Другими словами, пространство $\mathbb{A}(A)$ образовано теми векторами $X \in \mathbb{V}$, чье внутреннее произведение с формой равно нулю, $\mathsf{i}_X A = 0$. Это значит, что подпространство, ассоциированное с формой A, есть подпространство в $\mathbb{V}$, которое определяется системой уравнений $\mathsf{i}_X A = 0$.

Каждому ассоциированному пространству $\mathbb{A}(A)$ можно поставить в соответствие *ассоциированную с формой систему* $\mathbb{A}^*(A)$, как ортогональное дополнение сопряженного пространства

$$\mathbb{A}^*(A) := \mathbb{A}^\perp(A) \in \mathbb{V}^*.$$

Размерность ассоциированной системы называется *рангом* формы A

$$\mathsf{rank}\, A = \mathsf{dim}\, \mathbb{A}^*(A). \qquad \square$$

Ясно, что ранг r-формы A равен минимальному числу линейно независимых 1-форм таких, что форма A равна линейной комбинации их внешних произведений.

Пример 4.4.2. Для 2-форм $A = \frac{1}{2} e^\alpha \wedge e^\beta A_{\alpha\beta}$ понятие ранга особенно наглядно: это просто ранг матрицы компонент,

$$\mathsf{dim}\, \mathbb{A}^*(A) = \mathsf{rank}\, A_{\alpha\beta}. \qquad \square$$

Пример 4.4.3. Рассмотрим 0-формы $A \in \Lambda_0(\mathbb{V})$, т.е. $A \in \mathbb{R}$. Тогда ассоциированное подпространство $\mathbb{A}(A) = \mathbb{V}$, т.к. факторпространство $\mathbb{V}/\mathbb{V}$ состоит из одной точки – начала координат, и $\Lambda_0(\mathbb{V}/\mathbb{V}) = \mathbb{R}$. Соответствующая система $\mathbb{A}^*(A)$, ассоциированная с формой A, также состоит из одной точки – начала координат сопряженного пространства $\mathbb{A}^*(A) = \{0\}$. Таким образом, ранг произвольной 0-формы равен нулю. $\square$

Пример 4.4.4. Пусть $\mathbb{V} = \mathbb{R}^n$ с базисом $\{e_1, \ldots, e_n\}$ и $A = e^1 + e^2 \in \Lambda_1(\mathbb{R}^n)$, тогда максимальное подпространство $\mathbb{A}(A) = \mathbb{R}^{n-1}$ порождено базисными векторами $\{e_1 - e_2, e_3, \ldots, e_n\}$. Соответствующая ассоциированная система $\mathbb{A}^*(A)$ представляет собой одномерное подпространство сопряженного пространства $\mathbb{V}^* = \mathbb{R}^n$, порожденное формой A, т.е. натянутое на базисный вектор $e^1 + e^2$. Ранг 1-формы $e^1 + e^2$ равен единице. Ясно, что ранг произвольной 1-формы равен единице. $\square$

Пример 4.4.5. Пусть $\mathbb{V} = \mathbb{R}^n$ и $A = e^1 \wedge e^2 + e^3 \wedge e^4 \in \Lambda_2(\mathbb{R}^n)$. Эта 2-форма неразложима, поэтому ассоциированное с ней подпространство имеет размерность $n - 4$. Следовательно, ассоциированное подпространство имеет вид $\mathbb{A}(A) = \mathbb{R}^{n-4}$ и порождено векторами $e_5, \ldots, e_n$. Соответствующая ассоциированная система четырехмерна и порождена векторами e^1, e^2, e^3, e^4. Таким образом, ранг 2-формы $e^1 \wedge e^2 + e^3 \wedge e^4$ равен четырем. $\square$

В общем случае справедлива следующая теорема, дающая ограничения на возможный ранг формы степени r.

Теорема 4.4.1. *Если $A \in \Lambda_r(\mathbb{V})$ – ненулевая r-форма на $\mathbb{V}$, то ее ранг не меньше r и не больше n*

$$r \leq \mathsf{rank}\, A \leq n.$$

Ранг r-формы A равен r тогда и только тогда, когда форма A разложима.

Доказательство. См., например, [?]. □

Как следствие получаем, что, поскольку любая n-форма разложима, то ее ранг равен n. Можно доказать, что любая $(n-1)$-форма разложима и ее ранг, следовательно, равен $n-1$. А также, что ранг $(n-2)$-формы равен либо $n-2$, либо n.

Доказательства следующих трех предложений для внешних форм второй степени $A \in \Lambda_2(\mathbb{V})$, $\mathsf{dim}\,(\mathbb{V}) = n$, приведены в [?].

Предложение 4.4.1. *Пусть $A \in \Lambda_2(\mathbb{V})$ – внешняя 2-форма. Тогда существует целое четное число $2s$ такое, что $2 \le 2s \le n$, и $2s$ независимых линейных форм $a^1, \ldots, a^{2s}$ таких, что*

$$A = a^1 \wedge a^2 + \ldots + a^{2s-1} \wedge a^{2s}.$$

При этом 1-форму a^1 можно выбрать произвольно в ассоциированной системе $\mathbb{A}^*(A)$.

Предложение 4.4.2. *Внешняя 2-форма имеет четный ранг.*

Предложение 4.4.3. *Пусть $A \in \Lambda_2(\mathbb{V})$ – внешняя 2-форма. Для того, чтобы 2-форма A имела ранг $2s$, необходимо и достаточно, чтобы*

$$A^s \neq 0 \quad и \quad A^{s+1} = 0,$$

где в левой части соотношений стоят внешние степени формы A.

Приведенные выше определения и утверждения даны для внешней алгебры $\Lambda(\mathbb{V})$ над произвольным векторным пространством $\mathbb{V}$ и применимы в каждой точке многообразия $\mathbb{M}$ для внешней алгебры $\Lambda(\mathbb{M})$. При рассмотрении форм на многообразии эти понятия, в том числе ранг формы, могут меняться от точки к точке.

Введем несколько новых понятий, для которых уже важна зависимость компонент формы от точки многообразия.

Определение. *Характеристическим подпространством* r-формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется подпространство касательного пространства $\mathbb{C}_x(A) \subset \mathbb{T}_x(\mathbb{M})$, которое является пересечением ассоциированных подпространств для формы A и ее внешней производной dA,

$$\mathbb{C}_x(A) = \mathbb{A}(A) \cap \mathbb{A}(dA).$$

Характеристической системой формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется подпространство $\mathbb{C}^*_x(A) \subset \mathbb{T}^*_x(\mathbb{M})$, которое ортогонально характеристическому подпространству этой формы $\mathbb{C}_x(A)$. □

Из определения и формулы (2.41) вытекает, что характеристическая система $\mathbb{C}^*_x(A)$ строится по ассоциированным системам $\mathbb{A}^*(A)$ и $\mathbb{A}^*(dA)$ следующим образом

$$\begin{aligned}\mathbb{C}^*_x(A) &= \mathbb{A}^*(A) \oplus \Big[\mathbb{A}^*(dA) \setminus \Big(\mathbb{A}^*(A) \cap \mathbb{A}^*(dA)\Big)\Big] \\ &= \Big[\mathbb{A}^*(A) \setminus \Big(\mathbb{A}^*(A) \cap \mathbb{A}^*(dA)\Big)\Big] \oplus \mathbb{A}^*(dA).\end{aligned}$$

Определение. *Классом* r-формы $A \in \Lambda_r(\mathbb{M})$ в точке $x \in \mathbb{M}$ называется размерность характеристической системы: $\mathsf{class}\,(A) = \mathsf{dim}\,\mathbb{C}^*_x(A)$. □

Из определения следует, что класс ненулевой r-формы A не меньше ее ранга и, следовательно, не меньше степени r

$$r \le \mathsf{rank}\,(A) \le \mathsf{class}\,(A).$$

В общем случае класс формы может зависеть от точки многообразия.

Пример 4.4.6. Рассмотрим 1-форму $A = dy(x^2 + y^2)$ на евклидовой плоскости $\mathbb{R}^2$ с декартовыми координатами x, y. Ее внешний дифференциал равен $dA = dx \wedge dy\, x$. Существует три случая.

1) $x \neq 0$. В этом случае ассоциированное подпространство $\mathbb{A}(A)$ порождено вектором ∂_x, а ассоциированное подпространство внешнего дифференциала нульмерно, $\mathbb{A}(dA) = \{0\}$. Поэтому характеристическое подпространство состоит из одной точки $\mathbb{C}_x(A) = \{0\}$, а характеристическая система совпадает с кокасательным пространством, $\mathbb{C}_x(A) = \mathbb{T}^*_x(\mathbb{R}^2)$. Следовательно, класс формы A равен двум, $\mathsf{class}\,(A) = 2$.

2) $x = 0$, $y \neq 0$. Ассоциированное подпространство $\mathbb{A}(A)$ по-прежнему порождено вектором ∂_x. Внешний дифференциал формы A равен нулю, $dA = 0$, и ассоциированное подпространство внешнего дифференциала совпадает с касательным пространством, $\mathbb{A}(dA) = \mathbb{T}_x(\mathbb{R}^2)$. Следовательно, характеристическое подпространство $\mathbb{C}_x(A)$ порождено вектором ∂_x, а характеристическая система – дифференциалом dy. Поэтому класс формы A равен единице, $\mathsf{class}\,(A) = 1$.

3) $x = y = 0$. В этом случае и форма, и ее внешний дифференциал обращаются в нуль: $A = 0$, $dA = 0$. Поэтому ассоциированное подпространство совпадает с касательным пространством, $\mathbb{C}_x(A) = \mathbb{T}_x(\mathbb{R}^2)$, а ассоциированная система состоит из одной точки $\mathbb{C}^*_x = \{0\}$. Следовательно, класс формы A равен нулю, $\mathsf{class}\,(A) = 0$.

Таким образом, класс рассматриваемой формы зависит от точки плоскости и равен:

$\mathsf{class}\,(\mathrm{A})$=2, если $x \neq 0$;
$\mathsf{class}\,(\mathrm{A})$=1, если $x = 0$ и $y \neq 0$;
$\mathsf{class}\,(\mathrm{A})$=0, если $x = y = 0$. □

Продолжим общее рассмотрение. Нетрудно проверить, что если класс r-формы A равен ее степени r в некоторой точке $x \in \mathbb{M}$, то эта форма замкнута, $dA(x) = 0$. В обратную сторону справедливо другое утверждение. Если форма замкнута, $dA(x) = 0$, в точке x, то в этой точке ее класс равен рангу, а не степени. Отсюда, в частности, следует, что замкнутая форма второй степени имеет четный класс.

Теперь сформулируем основное утверждение настоящего раздела.

Теорема 4.4.2. *Пусть $A \in \Lambda_r(\mathbb{M})$ – дифференциальная r-форма на многообразии $\mathbb{M}$ постоянного класса, $\mathsf{class}\,(A) = q$, $r \leq q \leq n$. Тогда в окрестности любой точки $x \in \mathbb{M}$ существует такая система координат $y^1, \ldots, y^n$, что форма A представима в виде*

$$A = \sum_{1 \leq \alpha_1 < \ldots < \alpha_r \leq q} dy^{\alpha_1} \wedge \ldots \wedge dy^{\alpha_r} A_{\alpha_1 \ldots \alpha_r}(y^1 \ldots y^q). \tag{4.46}$$

Для форм непостоянного класса представления (4.46) не существует.

Доказательство. См., например, [**?**] □

Нетривиальность этой теоремы заключается в том, что форма записана в виде, содержащим только первые q координат. Эта теорема дает ответ на вопрос, поставленный в начале раздела. Другими словами, класс формы равен минимальному числу независимых функций, необходимых для явного выражения r-формы постоянного класса.

Следствие. Пусть $A \in \Lambda_r(\mathbb{M})$ – дифференциальная r-форма на многообразии $\mathbb{M}$ постоянного класса, который равен степени, $\mathsf{class}\,(A) = r$. Тогда в окрестности любой точки $x \in \mathbb{M}$ существует такая система координат $y^1, \ldots, y^n$, что форма A имеет вид

$$A = dy^{\alpha_1} \wedge \ldots \wedge dy^{\alpha_r} \tag{4.47}$$

Доказательство. См., например, [**?**] □

Для замкнутых 1- и 2-форм постоянного класса теорему 4.4.2 можно уточнить.

Теорема 4.4.3 (**Дарбу**). *Пусть A – произвольная 1-форма без нулей на многообразии $\mathbb{M}$ постоянного класса $2s+1$ (или $2s$). Тогда для любой точки $x \in \mathbb{M}$ существуют $2s+1$ (или $2s$) функций $y^1, \ldots, y^{2s+1}$ (или $y^1, \ldots, y^{2s}$), заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x, такие, что*

$$\begin{aligned} & y^1(x) = \ldots = y^{2s+1}(x) = 0, \\ \textit{или} \quad & y^1(x) = \ldots = y^{2s}(x) = 0, \end{aligned}$$

и

$$\begin{aligned} & A|_{\mathbb{U}} = dy^1 + y^2 dy^3 + \ldots + y^{2s} dy^{2s+1}, \\ \textit{или} \quad & A|_{\mathbb{U}} = (1+y^1)dy^2 + y^3 dy^4 + \ldots + y^{2s-1} dy^{2s}. \end{aligned}$$

Доказательство. См., например, [**?**]. □

Замечание. Если 1-форма имеет постоянный нечетный класс, то она не имеет нулей на многообразии. Напротив, если 1-форма имеет постоянный четный класс, то она может иметь нули. В этом случае невозможно привести общую локальную модель 1-формы. □

Теорема 4.4.4 (**Дарбу**). *Пусть ω – замкнутая дифференциальная 2-форма на многообразии $\mathbb{M}$ постоянного класса $\mathsf{class}\,\omega = 2s$. Тогда для любой точки $x \in \mathbb{M}$ существуют $2s$ дифференцируемых функций $y^1, \ldots, y^{2s}$, заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x, такие, что*

$$\begin{aligned} & y^1(x) = \ldots = y^{2s}(x) = 0, \\ & \omega|_{\mathbb{U}} = dy^1 \wedge dy^2 + \ldots + dy^{2s-1} \wedge dy^{2s}. \end{aligned}$$

Доказательство. Из замкнутости 2-формы ω по лемме Пуанкаре следует, что она локально представима в виде $\omega = dA$, где 1-форма A в точке x имеет класс $2s$ или $2s+1$. Тогда утверждение теоремы следует из теоремы 4.4.3. □

Пример 4.4.7. Рассмотрим симплектическое многообразие $(\mathbb{M}, \omega)$, $\mathsf{dim}\,\mathbb{M} = 2n$. По определению, симплектическая форма ω на многообразии $\mathbb{M}$ невырождена, и, следовательно, ее ранг и класс постоянны и равны размерности многообразия

$$\mathsf{rank}\,\omega = \mathsf{class}\,\omega = \mathsf{dim}\,\mathbb{M} = 2n.$$

Симплектическая форма ω также замкнута. Поэтому из теоремы Дарбу следует, что для любой точки $x \in \mathbb{M}$ существуют $2n$ дифференцируемых функций $y^1, \ldots, y^{2n}$, заданных в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$, содержащей точку x, таких, что

$$\begin{aligned} & y^1(x) = \ldots = y^{2n}(x) = 0, \\ & \omega|_{\mathbb{U}} = dy^1 \wedge dy^2 + \ldots + dy^{2n-1} \wedge dy^{2n}. \end{aligned} \tag{4.48}$$

Другими словами, для любой симплектической формы в окрестности произвольной точки существует такая система координат, в которой она имеет канонический вид. Именно это следствие теоремы Дарбу, которое часто тоже называют теоремой Дарбу, обычно используют в гамильтоновой динамике. □

Определение. Система координат, в которой симплектическая форма имеет канонический вид (4.48), называется *координатами Дарбу.* □

Замечание. В римановой геометрии $(\mathbb{M}, g)$ с метрикой g общего вида всегда можно выбрать такие координаты, чтобы метрика совпала с евклидовой метрикой (единичной матрицей) в любой заданной точке многообразия. В общем случае этого можно добиться вдоль экстремали, но не в окрестности. ??? На симплектическом многообразии симплектическую форму можно привести к каноническому виду в окрестности произвольной точки, что является значительно более сильным утверждением. □

4.5 Оператор Лапласа–Бельтрами

Определение. Пространства форм $\Lambda_r(\mathbb{M})$ и $\Lambda_{n-r}(\mathbb{M})$ имеют одинаковую размерность $C_n^r = C_n^{n-r}$, и при наличии метрики $g_{\alpha\beta}$ между их элементами можно установить взаимно однозначное соответствие (изоморфизм $\mathcal{C}^\infty(\mathbb{M})$-модулей), которое задается оператором

$$* : \quad \Lambda_r(\mathbb{M}) \;\leftrightarrow\; \Lambda_{n-r}(\mathbb{M}).$$

Рассмотрим дифференциальную r-форму $A = \frac{1}{r!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r}$ в (псевдо-)римановом пространстве, т.е. при наличии метрики. Поставим ей в соответствие $(n-r)$-форму по следующему правилу

$$* A = \frac{1}{(n-r)!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_{n-r}} (*A)_{\alpha_1 \ldots \alpha_{n-r}}, \tag{4.49}$$

где компоненты $(n-r)$-формы определены выражением

$$(*A)_{\alpha_{r+1} \ldots \alpha_n} = \frac{1}{r!} \varepsilon_{\alpha_1 \ldots \alpha_n} A^{\alpha_1 \ldots \alpha_r}, \qquad A^{\alpha_1 \ldots \alpha_r} = g^{\alpha_1 \beta_1} \ldots g^{\alpha_r \beta_r} A_{\beta_1 \ldots \beta_r}. \tag{4.50}$$

Действие оператора $*$ можно записать также в виде

$$* A = \frac{1}{(n-r)!} dx^{\alpha_{r+1}} \wedge \cdots \wedge dx^{\alpha_n} \frac{1}{r!} \varepsilon_{\alpha_1 \ldots \alpha_n} A^{\alpha_1 \ldots \alpha_r}. \tag{4.51}$$

Эти формулы определяют *оператор* $*$, который также называется *оператором Ходжа.* Форма $*A$ называется *дуальной* к форме A. □

Замечание. Метрика риманова пространства входит в определение оператора $*$ дважды. Во-первых, с ее помощью определяется полностью антисимметричный тензор $\varepsilon_{\alpha_1 \ldots \alpha_n} = \sqrt{|g|}\hat{\varepsilon}_{\alpha_1 \ldots \alpha_n}$. Во-вторых, метрика используется для подъема индексов в уравнении (4.50). Отметим, что для определения оператора $*$ необходимо не только наличие метрики на многообразии, но и его ориентируемость. В противном случае плотность $\sqrt{|g|}$ глобально не определена. □

Оператор Ходжа $*$ линеен:

$$*(fA + gB) = f(*A) + g(*B), \qquad \forall\ f, g \in \mathcal{C}^\infty(\mathbb{M}), \quad A, B \in \Lambda_r(\mathbb{M}).$$

Оператор $*$, действуя на r-форму, обладает следующим свойством

$$** A = (-1)^{r(n-r)} A \,\mathsf{sgn}\,, \tag{4.52}$$

где множитель sgn равен знаку определителя метрики (3.66). Вид этой формулы сохраняется при замене $r \mapsto n-r$. Отсюда следует, что обратный оператор, действующий на $(n-r)$-форму в порядке $*^{-1}*$ или на r-форму в порядке $**^{-1}$, имеет одинаковый вид

$$*^{-1} = (-1)^{r(n-r)} * \mathsf{sgn}\,.$$

Замечание. Отметим, что при проведении расчетов необходимо учитывать, что множитель $(-1)^{r(n-r)}$ зависит от степени формы, на которую действует оператор $*^{-1}$. Для форм четной степени он всегда равен единице. □

Замечание. Оператор $*$ можно определить в произвольной внешней алгебре, снабженной метрикой, поскольку он не содержит дифференцирования. □

Поскольку на $\mathbb{M}$ задана метрика, то определена свертка двух r-форм A и B:

$$(A,B) = g^{\alpha_1\beta_1}\dots g^{\alpha_r\beta_r}A_{\alpha_1\dots\alpha_r}B_{\beta_1\dots\beta_r},$$

которая дает скалярное поле.

Предложение 4.5.1. *Для двух r-форм $A, B \in \Lambda_2(\mathbb{M})$ справедлива формула*

$$A\wedge *B = dx^1\wedge\dots\wedge dx^n\sqrt{|g|}\frac{(A,B)}{r!}. \tag{4.53}$$

Доказательство. Прямая проверка. □

Определение. Введем *скалярное произведение*

$$\langle\cdot,\cdot\rangle:\quad \Lambda_r(\mathbb{M})\times\Lambda_r(\mathbb{M})\to\mathbb{R}$$

в пространстве r-форм $A, B \in \Lambda_r(\mathbb{M})$ с помощью интеграла

$$\langle A,B\rangle := \int_{\mathbb{M}} A\wedge *B = \int_{\mathbb{M}} B\wedge *A = \frac{1}{r!}\int_{\mathbb{M}} dx^1\dots dx^n\sqrt{|g|}(A,B)\,\mathsf{sgn}\,. \tag{4.54}$$

Мы предполагаем, что этот интеграл сходится. В противном случае скалярное произведение неопределено. □

Скалярное произведение симметрично, $\langle A,B\rangle = \langle B,A\rangle$. Если метрика риманова, т.е. положительно определена, то скалярное произведение $\langle A,B\rangle$ также положительно определено.

Предложение 4.5.2. *Оператор $*$ ортогонален:*

$$\langle *A, *B\rangle = \langle A,B\rangle\,\mathsf{sgn}\,.$$

Доказательство. Прямая проверка. □

Определение. Введем дифференциальный оператор *кограницы* $\delta\colon \Lambda_r(\mathbb{M})\to\Lambda_{r-1}(\mathbb{M})$

$$\delta := (-1)^{n(r+1)+1}\,\mathsf{sgn}\; * d *, \tag{4.55}$$

понижающий степень формы на единицу. Его можно записать также в виде

$$\delta = (-1)^r\, *^{-1} d *,$$

учитывая, что обратный оператор $*^{-1}$ действует на $(n-r+1)$-форму. □

В компонентах действие оператора δ на r-форму A имеет вид

$$\delta A = -\frac{1}{(r-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_r} \frac{1}{\sqrt{|g|}} (g_{\alpha_2\beta_2} \dots g_{\alpha_r\beta_r}) \partial_{\alpha_1} (\sqrt{|g|} A^{\alpha_1\beta_2\dots\beta_r})$$
$$= -\frac{1}{(r-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_r} \widetilde{\nabla}_{\alpha_1} A^{\alpha_1}{}_{\alpha_2\dots\alpha_r}. \tag{4.56}$$

Отсюда следует, что оператор δ можно интерпретировать, как ковариантную дивергенцию антисимметричного тензора с верхними индексами. По-определению, действие оператора когранцы на функцию дает нуль $\delta f = 0$.

Предложение 4.5.3. *Для замкнутых (компактных и без края) многообразий оператор δ сопряжен с оператором внешнего дифференцирования относительно скалярного произведения (4.54)*

$$\langle dA, B\rangle = \langle A, \delta B\rangle, \tag{4.57}$$

где $A \in \Lambda_{r-1}(\mathbb{M})$ и $B \in \Lambda_r(\mathbb{M})$.

Доказательство. Прямая проверка. Компактность многообразия достаточна для существования интеграла (4.54), а отсутствие края позволяет отбросить граничные слагаемые, возникающие при интегрировании по частям. □

Предложение 4.5.4. *Квадрат оператора δ равен нулю:*

$$\delta\delta = 0, \tag{4.58}$$

Доказательство. Прямое следствием определения (4.55) и нильпотентности внешнего дифференцирования (4.32). □

Определение. Оператором *Лапласа–Бельтрами* в римановом пространстве называется оператор

$$\triangle_{\text{F}} = -(d+\delta)^2 = -d\delta - \delta d, \tag{4.59}$$

который действует в пространстве форм, $\triangle_{\text{F}} : \Lambda_r(\mathbb{M}) \to \Lambda_r(\mathbb{M})$, $0 \le r \le n$, □

Предложение 4.5.5. *Для замкнутых многообразий оператор Лапласа–Бельтрами самосопряжен:*

$$\langle \triangle_{\text{F}} A, B\rangle = \langle A, \triangle_{\text{F}} B\rangle,$$

Доказательство. Из симметричности скалярного произведения и формулы (4.57) следует равенство

$$\langle d\delta A, B\rangle = \langle \delta A, \delta B\rangle = \langle A, d\delta B\rangle.$$

Отсюда вытекает самосопряженность оператора Лапласа–Бельтрами. □

Замечание. Оператор Лапласа–Бельтрами является дифференциальным оператором второго порядка эллиптического типа в римановом пространстве с положительно определенной метрикой. Если на многообразии задана метрика лоренцевой сигнатуры, то инвариантный оператор (4.59) также определен. В этом случае он будет гиперболического типа. Если для метрики лоренцевой сигнатуры скалярное произведение определено, то оператор (4.59) будет самосопряжен, поскольку в доказательстве сигнатура метрики не используется. □

Предложение 4.5.6. *Оператор Лапласа–Бельтрами удовлетворяет следующим перестановочным соотношениям:*

$$\triangle_{\text{F}} d = d\triangle_{\text{F}}, \qquad \triangle_{\text{F}}\delta = \delta\triangle_{\text{F}}, \qquad \triangle_{\text{F}} * = * \triangle_{\text{F}}, \tag{4.60}$$

Доказательство. Прямое следствие определения (4.59). □

С помощью прямых вычислений можно получить формулы действия оператора Лапласа–Бельтрами на 0-форму f, 1-форму A и произвольную r-форму B в компонентах:

$$\begin{aligned}
&\triangle_{\text{F}} f = \tilde{\triangle} f,\\
&\triangle_{\text{F}} A = dx^\alpha(\tilde{\triangle} A_\alpha + \tilde{R}_\alpha{}^\beta A_\beta),\\
&\triangle_{\text{F}} B = \frac{1}{r!} dx^{\alpha_1}\wedge\ldots\wedge dx^{\alpha_r}\left(\tilde{\triangle} A_{\alpha_1\ldots\alpha_r} + r\tilde{R}_{\alpha_1}{}^\beta A_{\beta\alpha_2\ldots\alpha_r} - \frac{r(r-1)}{2}\tilde{R}_{\alpha_1\alpha_2}{}^{\beta\gamma} A_{\beta\gamma\alpha_3\ldots\alpha_r}\right),
\end{aligned} \tag{4.61}$$

где $\tilde{\triangle} = g^{\alpha\beta}\widetilde{\nabla}_\alpha\widetilde{\nabla}_\beta$ – оператор Лапласа–Бельтрами, построенный с помощью символов Кристоффеля.

Замечание. В аффинной геометрии мы имеем три инвариантных дифференциальных оператора второго порядка $\triangle_{\text{F}}$, $\tilde{\triangle}$ и $\triangle$, действующих из $\Lambda_r(\mathbb{M}) \to \Lambda_r(\mathbb{M})$. Первый оператор определен только на формах и поэтому помечен индексом F. Два других определены для произвольных тензорных полей. В (псевдо-)римановой геометрии при нулевом кручении и неметричности $\tilde{\triangle} = \triangle$. Отметим также, что действие операторов $\triangle_{\text{F}}$ и $\tilde{\triangle}$ на функции совпадает. Как правило, под оператором Лапласа–Бельтрами мы понимаем оператор $\tilde{\triangle}$, действующий на произвольные тензорные поля и зависящий только от метрики. □

Замечание. В геодезических (римановых) координатах (см. раздел 17.9) символы Кристоффеля и все симметризованные частные производные от них обращаются в нуль в некоторой точке p, поэтому в этой точке оператор Лапласа–Бельтрами принимает вид $\tilde{\triangle}|_{x=p} = g^{\alpha\beta}\partial_\alpha\partial_\beta$ □

Пример 4.5.1. Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами $x^\alpha = \{x, y, z\}$. Приведем явные формулы действия операторов, введенных ранее, на формы. В трехмерном пространстве могут быть заданы только 0-, 1-, 2- и 3-формы:

$$\begin{aligned}
&f && - \text{ 0-форма},\\
&A = dx^\alpha A_\alpha = dxA_x + dyA_y + dzA_z && - \text{ 1-форма},\\
&B = \frac{1}{2} dx^{\alpha_1}\wedge dx^{\alpha_2} B_{\alpha_1\alpha_2} = dx\wedge dyB_{xy} + dy\wedge dzB_{yz} + dz\wedge dxB_{zx} && - \text{ 2-форма},\\
&C = \frac{1}{3!} dx^{\alpha_1}\wedge dx^{\alpha_2}\wedge dx^{\alpha_3} C_{\alpha_1\alpha_2\alpha_3} = dx\wedge dy\wedge dzC_{xyz} && - \text{ 3-форма}.
\end{aligned}$$

Формы более высокого порядка тождественно равны нулю в силу трехмерности пространства. Пространства 0- и 3-форм являются одномерными, а пространства 1- и 2-форм – трехмерны.

Оператор внешнего дифференцирования (4.29) повышает на единицу степень формы и действует следующим образом:

$$\begin{aligned}
df &= dx\partial_x f + dy\partial_y f + dz\partial_z f,\\
dA &= dx\wedge dy(\partial_x A_y - \partial_y A_x) + dy\wedge dz(\partial_y A_z - \partial_z A_y) + dz\wedge dx(\partial_z A_x - \partial_x A_z),\\
dB &= dx\wedge dy\wedge dz(\partial_x B_{yz} + \partial_y B_{zx} + \partial_z B_{xy}),\\
dC &= 0.
\end{aligned}$$

Нетрудно проверить, что оператор $*$, определенный соотношением (4.49), действует на формы по-правилам:

$$\begin{aligned}
*f &= dx\wedge dy\wedge dz f,\\
*A &= dx\wedge dy A_z + dy\wedge dz A_x + dz\wedge dx A_y,\\
*B &= dx B_{yz} + dy B_{zx} + dz B_{xy},\\
*C &= C_{xyz}.
\end{aligned}$$

Этот оператор устанавливает взаимно однозначное соответствие между пространствами $\Lambda_0(\mathbb{R}^3)$ и $\Lambda_3(\mathbb{R}^3)$, а также между $\Lambda_1(\mathbb{R}^3)$ и $\Lambda_2(\mathbb{R}^3)$. Заметим, что для любой формы в трехмерном евклидовом пространстве квадрат оператора $*$ является тождественным оператором:

$$** = 1.$$

Оператор кограницы δ, определенный соотношением (4.55), понижает степень форм на единицу:

$$\begin{aligned}
\delta f &= 0,\\
\delta A &= -(\partial_x A_x + \partial_y A_y + \partial_z A_z),\\
\delta B &= dx(\partial_y B_{xy} - \partial_z B_{zx}) + dy(\partial_z B_{yz} - \partial_x B_{xy}) + dz(\partial_x B_{zx} - \partial_y B_{yz}),\\
\delta C &= -(dx\wedge dy\partial_z C_{xyz} + dy\wedge dz\partial_x C_{xyz} + dz\wedge dx\partial_y C_{xyz}).
\end{aligned}$$

Теперь нетрудно найти явное действие оператора Лапласа–Бельтрами в пространстве форм:

$$\begin{aligned}
\triangle_{\text{F}} f &= \triangle f,\\
\triangle_{\text{F}} A &= dx\triangle A_x + dy\triangle A_y + dz\triangle A_z,\\
\triangle_{\text{F}} B &= dx\wedge dy\triangle B_{xy} + dy\wedge dz\triangle B_{yz} + dz\wedge dx\triangle B_{zx},\\
\triangle_{\text{F}} C &= dx\wedge dy\wedge dz\triangle C_{xyz},
\end{aligned}$$

где

$$\triangle = \partial_x^2 + \partial_y^2 + \partial_z^2$$

– обычный оператор Лапласа в трехмерном евклидовом пространстве. □

Определение. Действие оператора внешнего дифференцирования на 0-формы совпадает с определением градиента функции, поэтому назовем *градиентом* произвольной r-формы ее внешнюю производную

$$\mathsf{grad} = d. \tag{4.62}$$

Действие оператора δ на 1-формы с точностью до знака совпадает с дивергенцией 1-формы. Поэтому примем в качестве инвариантного определения *дивергенции* от произвольной r-формы соотношение

$$\mathsf{div} = -\delta. \tag{4.63}$$

□

В заключение рассмотрим еще один инвариантный оператор, действующий в пространстве форм произвольной степени,

$$\mathsf{rot} = *d: \quad \Lambda_r \to \Lambda_{n-r-1}. \tag{4.64}$$

В частном случае трехмерного евклидова пространства (пример 4.5.1) он отображает 1-формы в 1-формы:

$$*dA = dx(\partial_y A_z - \partial_z A_y) + dy(\partial_z A_x - \partial_x A_z) + dz(\partial_x A_y - \partial_y A_x),$$

что совпадает с определением ротора от 1-формы. Поэтому примем выражение (4.64) в качестве инвариантного определения *ротора*. Тогда определение оператора Лапласа–Бельтрами (4.59) сводится к хорошо известному тождеству из векторного анализа

$$\triangle_{\text{F}} = \mathsf{grad}\ \mathsf{div} - \mathsf{rot}\ \mathsf{rot}\,. \tag{4.65}$$

4.6 Интегрирование дифференциальных форм

Начнем с обсуждения интегрирования в евклидовом пространстве $\mathbb{R}^n$. Для наших целей достаточно знакомство с определением интеграла по Риману, т.к. мы будем интегрировать непрерывные функции по достаточно хорошим областям. Напомним правило замены переменных интегрирования. Пусть h – диффеоморфизм ограниченного открытого подмножества $\mathbb{U} \subset \mathbb{R}^n$ на ограниченное открытое подмножество $h(\mathbb{U}) \subset \mathbb{R}^n$. В координатах диффеоморфизм h задается n гладкими функциями $x^\alpha \mapsto y^\alpha(x)$, $\alpha = 1, \ldots, n$, где x^α – координаты точки $x \in \mathbb{U}$. Пусть f – ограниченная непрерывная функция на $\mathbb{U}$, включая границу. Тогда

$$\int\limits_{\mathbb{U}} dx^1 \ldots dx^n f(x) = \int\limits_{h(\mathbb{U})} dy^1 \ldots dy^n J^{-1} f\Big(x(y)\Big).$$

где $J = \mathsf{det}\,(\partial y^\beta / \partial x^\alpha)$ – якобиан преобразования координат.

Определение. Рассмотрим теперь произвольную n-форму на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$. На карте $(\mathbb{U}, \varphi)$ она имеет вид

$$A = dx^1 \wedge \ldots \wedge dx^n A_{1\ldots n} = \frac{1}{n!} dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_n} A_{\alpha_1 \ldots \alpha_n},$$

где в последнем выражении производится суммирование по всем значениям индексов. Положим

$$\int_{\mathbb{U}} A = \int_{\mathbb{U}} dx^1 \wedge \ldots \wedge dx^n A_{1\ldots n} := \int_{\varphi(\mathbb{U})} dx^1 \ldots dx^n A_{1\ldots n}(x) := \int_{\varphi(\mathbb{U})} d^n x A_{1\ldots n}(x). \tag{4.66}$$

Правая часть этого равенства определена, т.к. $\varphi(\mathbb{U}) \subset \mathbb{R}^n$, а запись $A_{1\ldots n}(x)$ обозначает значение тензорной плотности $A_{1\ldots n} \circ \varphi^{-1}$ в точке $x^\alpha \in \varphi(\mathbb{U})$. Формула (4.66) корректна, т.к. при преобразовании координат $x^\alpha \mapsto y^\alpha(x)$ (замене переменных интегрирования) единственная нетривиальная компонента n-формы $A_{1\ldots n}$ преобразуется, как тензорная плотность степени -1: $A_{1\ldots n}(x) \mapsto J^{-1} A_{1\ldots n}(y)$. Равенство (4.66) примем за определение *интеграла от n-формы* по области $\mathbb{U} \subset \mathbb{M}$. □

Замечание. Определение интеграла от n-формы по области $\mathbb{U} \subset \mathbb{M}$ не зависит от того задана на многообразии метрика или нет. □

Из равенства

$$dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_n} = dx^1 \wedge \ldots \wedge dx^n \hat{\varepsilon}^{\alpha_1 \ldots \alpha_n}, \tag{4.67}$$

где $\hat{\varepsilon}^{\alpha_1 \ldots \alpha_n} = \sqrt{|g|}\varepsilon^{\alpha_1 \ldots \alpha_n}$ – полностью антисимметричная тензорная плотность степени -1, следует, что интеграл (4.66) можно записать в виде

$$\int_{\mathbb{U}} A = \int_{\varphi(\mathbb{U})} d^n x \hat{\varepsilon}^{\alpha_1 \ldots \alpha_n} A_{\alpha_1 \ldots \alpha_n}.$$

Поскольку компоненты тензорной плотности $\hat{\varepsilon}^{\alpha_1 \ldots \alpha_n}$ во всех системах координат по модулю равны единице, то этот интеграл не зависит от метрики на $\mathbb{M}$.

Правило замены переменных интегрирования можно переписать иначе на языке возврата отображения. Пусть $h : \mathbb{U} \to h(\mathbb{U})$ диффеоморфизм двух областей многообразия $\mathbb{M}$, и пусть форма A задана на области $h(\mathbb{U})$. Тогда

$$\int_{h(\mathbb{U})} A = \int_{\mathbb{U}} h^* A,$$

где h^* – возврат отображения h.

Замечание. Запись интеграла $\int d^n x A_{1\ldots n}$ в виде $\int dx^1 \wedge \ldots \wedge dx^n A_{1\ldots n}$ удобнее, т.к. якобиан преобразования координат возникает автоматически при замене дифференциалов:

$$dy^1 \wedge \ldots \wedge dy^n = \left(\frac{\partial y^1}{\partial x^{\alpha_1}} dx^{\alpha_1}\right) \wedge \ldots \wedge \left(\frac{\partial y^n}{\partial x^{\alpha_n}} dx^{\alpha_n}\right) = J dx^1 \wedge \ldots \wedge dx^n. \quad \square \tag{4.68}$$

Теперь определим интеграл по произвольному ориентируемому многообразию с помощью разбиения единицы.

Определение. Пусть на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, заданы: ориентация, некоторый атлас $\{\mathbb{U}_i\}$, согласованный с ориентацией, и подчиненное ему разбиение единицы $\{f_i\}$. Рассмотрим произвольную непрерывную n-форму $A \in \Lambda_n(\mathbb{M})$ с компактным носителем. Тогда

$$A = \left(\sum_i f_i\right) A = \sum_i (f_i A).$$

Поскольку $\mathsf{supp}\,(f_i A) \subset \mathsf{supp}\, f_i$, то положим

$$\int_{\mathbb{M}} f_i A = \int_{\mathbb{U}_i} f_i A,$$

где интеграл в правой части был определен ранее (4.66). Этот интеграл определен корректно, т.к. не зависит от выбора системы координат на $\mathbb{U}_i$. Определим *интеграл по ориентируемому многообразию* от n-формы с компактным носителем следующей формулой

$$\int_{\mathbb{M}} A = \sum_i \int_{\mathbb{M}} f_i A. \tag{4.69}$$

Для любого заданного разбиения единицы правая часть равенства однозначно определена. $\square$

Нетрудно проверить, что данное определение не зависит от разбиения единицы. Требование компактности носителя формы A является достаточным для существования интеграла.

Замечание. При изменении ориентации системы координат в области $\mathbb{U}_i$ интеграл по этой области изменит знак. Поскольку на неориентируемом многообразии не существует возможности однозначно задать ориентацию, то правая часть равенства (4.69) будет зависеть от выбора ориентации областей и, следовательно, определена неоднозначно. Поэтому требование ориентируемости $\mathbb{M}$ является существенным. □

Для любых двух n-форм A и B с компактным носителем очевидны равенства

$$\int_{\mathbb{M}}(A+B)=\int_{\mathbb{M}}A+\int_{\mathbb{M}}B,$$
$$\int_{\mathbb{M}}aA=a\int_{\mathbb{M}}A,\qquad a\in\mathbb{R}.$$

Поэтому интеграл является линейным функционалом на множестве n-форм с компактным носителем.

Если $\mathbb{M}$ – ориентируемое многообразие с краем $\partial\mathbb{M}$, то положим

$$\int_{\mathbb{M}}A:=\int_{\mathbb{M}\setminus\partial\mathbb{M}}A.$$

Правая часть этого равенства хорошо выше, т.к. $\mathbb{M}\setminus\partial\mathbb{M}$ – ориентируемое многообразие без края.

Если ориентируемое многообразие $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M}=n$, можно представить в виде объединения

$$\mathbb{M}=\Big(\cup_i\mathbb{U}_i\Big)\cup\Big(\cup_{\mathrm{A}}\mathbb{V}_{\mathrm{A}}\Big)$$

непересекающихся открытых подмногообразий $\mathbb{U}_i$ (и, следовательно, многообразий той же размерности $\mathsf{dim}\,\mathbb{U}_i=n$, $\forall i$) и произвольного набора подмногообразий меньшей размерности $\mathbb{V}_{\mathrm{A}}$, $\mathsf{dim}\,\mathbb{V}_{\mathrm{A}}<n$, $\forall\mathrm{A}$, тогда

$$\int_{\mathbb{M}}A=\sum_i\int_{\mathbb{U}_i}A.$$

Это свойство можно использовать для вычисления интегралов, используя "глобальные" координаты, заданные на всем многообразии, за исключением подмногообразий меньшей размерности, которые не дают вклада в интеграл.

Пример 4.6.1. В задачах со сферической симметрией в $\mathbb{R}^3$ интегралы удобнее вычислять в сферических координатах, которые покрывают все многообразие, за исключением оси z. В этом случае евклидово пространство $\mathbb{R}^3$ представляется в виде

$$\mathbb{R}^3=\mathbb{U}\cup\mathbb{V},$$

где $\mathbb{V}=\mathbb{R}$ – ось z, и $\mathbb{U}=\mathbb{R}^3\setminus\mathbb{V}$ – неодносвязное открытое подмногообразие в $\mathbb{R}^3$, $\mathsf{dim}\,\mathbb{U}=3$. □

4.6.1 Форма объема

Пусть на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M}=n$, задана метрика $dx^\alpha\otimes dx^\beta g_{\alpha\beta}$[1]. Рассмотрим карту $(\mathbb{U},\varphi)$. В этой карте метрика имеет компоненты $g_{\alpha\beta}$, и ее определитель при преобразовании координат $x\mapsto y(x)$ (в областях пересечения двух карт с координатами x^α и y^α) преобразуется по правилу

$$\mathsf{det}\,g_{\alpha\beta}(y)=\mathsf{det}\,g_{\alpha\beta}(x)J^{-2},$$

[1]Согласно теореме 5.1.1 на любом гладком многообразии можно задать риманову (положительно определенную) метрику.

где $J = \det(\partial y^\beta/\partial x^\alpha)$ – якобиан преобразования координат. Поэтому определитель метрики является тензорной плотностью степени -2. Обозначим

$$g = \det g_{\alpha\beta}. \tag{4.70}$$

Тогда $\sqrt{|g|}$ является скалярной плотностью степени -1. Точнее, при преобразованиях координат $\sqrt{|g|}$ умножается на якобиан

$$\sqrt{|g|}(y) = \sqrt{|g|}(x)J^{-1} \tag{4.71}$$

и, следовательно, является тензорной плотностью степени -1. Если задано координатное покрытие многообразия $\mathbb{M} = \cup_i \mathbb{U}_i$, то в каждой карте определена плотность $\sqrt{|g|}$. Для того, чтобы плотность $\sqrt{|g|}$ была определена глобально, необходимо существование такого координатного покрытия, в которым все якобианы преобразований координат были положительны, а это влечет за собой ориентируемость многообразия. Таким образом плотность $\sqrt{|g|}$ определена только на ориентируемых многообразиях. Она всюду отлична от нуля, что следует из невырожденности метрики.

Определение. n-форма на ориентируемом многообразии $\mathbb{M}$ с заданной метрикой $g_{\alpha\beta}$

$$\upsilon = dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} = \frac{1}{n!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\alpha_1 \ldots \alpha_n}, \tag{4.72}$$

называется *формой объема* υ, задаваемой метрикой $g_{\alpha\beta}$. □

Замечание. Форма объема на ориентируемом многообразии $\mathbb{M}$ задается произвольной метрикой, независимо от ее сигнатуры. Ясно, что различные метрики могут определять одну и ту же форму объема из-за наличия знака модуля. Например, евклидова метрика и метрика Лоренца в евклидовом пространстве $\mathbb{R}^n$ задают одинаковую форму объема. □

Замечание. Для обозначения формы объема мы используем греческую букву эпсилон υ, которая по написанию очень похожа на латинскую букву v. □

Форма объема определена глобально на ориентируемых многообразиях и всюду отлична от нуля, поэтому она задает ориентацию многообразия и выбор знака в форме (4.72) соответствует выбору ориентации. При замене координат с отрицательным якобианом ориентация формы объема меняется на противоположную, а интеграл меняет знак.

Метрика евклидова пространства $\mathbb{R}^n$ в декартовых координатах равна единичной матрице (2.8). Поэтому форма объема евклидова пространства $\mathbb{R}^n$ в декартовых координатах имеет канонический вид

$$\upsilon_0 = dx^1 \wedge \ldots \wedge dx^n. \tag{4.73}$$

Будем говорить, что данная форма объема задает *каноническую ориентацию*.

Замечание. Покажем, что форма объема υ на многообразии $\mathbb{M}$ соответствует нашему обычному представлению об объеме. Для этого зафиксируем произвольную точку x риманова многообразия $\mathbb{M}$ с положительно определенной метрикой и диагонализируем в ней метрику, что всегда возможно в произвольной фиксированной точке. Тогда все базисные векторы в точке x будут ортогональны, а длина базисного вектора ∂_α будет равна $\sqrt{g_{\alpha\alpha}}$. В этом случае объем бесконечно малого параллелепипеда со сторонами ∂_α равен $\prod_\alpha \sqrt{g_{\alpha\alpha}} = \sqrt{g}$. □

Интеграл (4.66) от произвольной n-формы по многообразию можно переписать, используя форму объема,

$$\int_{\mathbb{U}} A = \frac{1}{n!}\int_{\mathbb{U}} dx^{\alpha_1}\wedge\cdots\wedge dx^{\alpha_n} A_{\alpha_1\ldots\alpha_n} = \frac{1}{n!}\int_{\mathbb{U}} d^n x\,\sqrt{|g|}\varepsilon^{\alpha_1\ldots\alpha_n} A_{\alpha_1\ldots\alpha_n}\,\mathsf{sgn}\,, \qquad (4.74)$$

где множитель sgn равен знаку определителя метрики (3.66).

Приведем два свойства формы объема. Нетрудно проверить, что форма объема дуальна к единице

$$\upsilon = \frac{1}{n!} dx^{\alpha_1}\wedge\cdots\wedge dx^{\alpha_n}\varepsilon_{\alpha_1\ldots\alpha_n} = *1.$$

Кроме того она является гармонической формой $\triangle_{\text{г}}\upsilon = 0$, т.к. форма объема точна $d\upsilon = 0$, как форма максимальной размерности, и коточна $\delta\upsilon = 0$, поскольку $(n+1)$-форм на $\mathbb{R}^n$ не существует и $\widetilde{\nabla}_\alpha\sqrt{|g|} = 0$.

Если на многообразии задан репер, $g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b\eta_{ab}$, где η_{ab} – евклидова или лоренцева метрика, то $\sqrt{|g|} = \mathsf{det}\, e_\alpha{}^a$. Тогда репер $e_\alpha{}^a$ однозначно определяет форму объема υ. В обратную сторону утверждение неверно, т.к. форма объема определяет только определитель репера. Даже если метрика и форма объема на ориентируемом многообразии $\mathbb{M}$ существуют, то репер может не существовать. Локально репер всегда существует, однако глобально это не так.

Пример 4.6.2. В теореме 11.2.1 о невозможности причесать ежа доказано, что на четномерной сфере не существует непрерывных векторных векторных полей, нигде не обращающихся в нуль. Поскольку репер представляет собой набор из n линейно независимых в каждой точке векторных полей на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, то на четномерной сфере не существует глобально определенного репера. В то же время метрика определена глобально. □

Наличие формы объема на (псевдо-)римановом ориентируемом многообразии позволяет определить инвариантным образом интеграл от произвольной функции $f(x)\in\mathcal{C}(\mathbb{M})$:

$$\int_{\mathbb{M}}\upsilon f = \int_{\mathbb{M}} dx^1\wedge\ldots\wedge dx^n\sqrt{|g|}f(x) = \int_{\mathbb{M}} d^n x\,\sqrt{|g|}f(x).$$

Используя определение скалярного произведения (4.54), этот интеграл можно переписать в виде

$$\int_{\mathbb{M}}\upsilon f = \langle f, 1\rangle.$$

Замечание. Именно это свойство чаще всего используется для построения моделей математической физики. А именно, мы фиксируем некоторый набор полей $\{\varphi^a(x)\}$, $a = 1, 2, \ldots$, на многообразии $\mathbb{M}$, среди которых содержится метрика пространства-времени, как в общей теории относительности. Затем строится лагранжева плотность $L(\varphi, \partial\varphi, \partial^2\varphi, \ldots)$, которая является скалярным полем (функцией), зависящим от данного набора полей и их частных производных. Определяем инвариантное действие на многообразии

$$S := \int_{\mathbb{M}}\upsilon L.$$

Уравнения Эйлера–Лагранжа для данного действия принимаются в качестве уравнений движения (уравнений равновесия и т.д.) для рассматриваемой модели. Попостроению, действие инвариантно относительно общих преобразований координат

и, возможно, других преобразований симметрии. Соответствующие уравнения Эйлера–Лагранжа ковариантны, т.е. преобразуются, как тензорные поля, и из первой теоремы Нетер следуют законы сохранения. В релятивистских моделях теории поля, не описывающих гравитационные взаимодействия, предполагается, что многообразие $\mathbb{M}$ является пространством-временем Минковского $\mathbb{R}^{1,3}$ с заданной лоренцевой метрикой, по которой варьирование не проводится. Соответствующие модели инвариантны относительно преобразований из группы Пуанкаре и, возможно, других групп, а инвариантность относительно общих преобразований координат в них отсутствует. □

Приведем без доказательства два свойства формы объема υ на ориентированном римановом многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$. Пусть $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_n$ – два набора векторных полей на $\mathbb{M}$. Тогда

$$\upsilon(X_1, \ldots, X_n) \cdot \upsilon(Y_1, \ldots, Y_n) = \det\{(X_i, Y_j)\}, \qquad i, j = 1, \ldots, n,$$
$$\upsilon(X_1, \ldots, X_n) \cdot \upsilon = \tilde{X}_1 \wedge \ldots \wedge \tilde{X}_n,$$

где $(X_i, Y_j) = X_i^\alpha Y_j^\beta g_{\alpha\beta}$ и $\tilde{X}_i = dx^\alpha X_{i\alpha}$ обозначает 1-форму, сопряженную в смысле римановой метрики к векторному полю $X_i = X_i^\alpha \partial_\alpha$: $X_{i\alpha} = X_i^\beta g_{\beta\alpha}$.

Определение. Интеграл от формы объема (4.72) по всему многообразию, если он существует, называется *объемом* ориентируемого многообразия $\mathbb{M}$

$$V_{\mathbb{U}} = \int_{\mathbb{M}} \upsilon = \int_{\mathbb{M}} d^n x \sqrt{|g|}. \quad \square \tag{4.75}$$

Объем многообразия определяется метрикой, входящей в определение формы объема, и всегда положителен, т.к. ориентация многообразия фиксирована.

4.6.2 Формула Стокса

Пусть $\mathbb{N}$, $\dim \mathbb{N} = n$ – ориентируемое многообразие, на котором задана произвольная r-форма $A \in \Lambda_r(\mathbb{N})$. Рассмотрим r-мерное ориентируемое подмногообразие $(f, \mathbb{M})$ в $\mathbb{N}$. Поскольку r-форма задана также на подмногообразии $\mathbb{M} \subset \mathbb{N}$, то ее можно по нему проинтегрировать. Пусть подмногообразие $\mathbb{M} \subset \mathbb{N}$ задано параметрически[2]

$$x^\alpha = x^\alpha(u^i), \qquad \alpha = 1, \ldots, n, \quad i = 1, \ldots, r \le n.$$

Тогда интеграл от произвольной r-формы A степени r по этому подмногообразию равен

$$\begin{aligned} \int_{f(\mathbb{M})} A &= \frac{1}{r!} \int_{f(\mathbb{M})} dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_r} A_{\alpha_1 \ldots \alpha_r} = \\ &= \frac{1}{r!} \int_{\mathbb{M}} du^{i_1} \wedge \ldots \wedge du^{i_r} \frac{dx^{\alpha_1}}{du^{i_1}} \ldots \frac{dx^{\alpha_r}}{du^{i_r}} A_{\alpha_1 \ldots \alpha_r}. \end{aligned} \tag{4.76}$$

Этот интеграл имеет явно ковариантный вид и не зависит ни от выбора системы координат x^α, ни от параметризации подмногообразия u^i.

[2]Чтобы не усложнять ситуацию склейкой карт, которую можно провести, будем считать, что каждое многообразие покрывается одной картой.

Пример 4.6.3. Рассмотрим интеграл от 1-формы $A = dx^\alpha A_\alpha$ вдоль кривой $x^\alpha(t)$, $t \in [a,b]$, соединяющей точки x_1 и x_2 в евклидовом пространстве $\mathbb{R}^n$. Тогда интеграл (4.76) принимает стандартный вид

$$\int_{x_1}^{x_2} dx^\alpha A_\alpha = \int_a^b dt \frac{dx^\alpha}{dt} A_\alpha.$$

Этот интеграл инвариантен относительно перепараметризации кривой и общего преобразования координат на многообразии. □

Пример 4.6.4. Интеграл от 2-формы $A = \frac{1}{2} dx^\alpha \wedge dx^\beta A_{\alpha\beta}$ по поверхности в евклидовом пространстве $\mathbb{R}^n$, заданной параметрически $x^\alpha = x^\alpha(u,v)$, $(u,v) \in \mathbb{U}$, равен

$$\frac{1}{2} \int_{f(\mathbb{U})} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \frac{1}{2} \int_{\mathbb{U}} du \wedge dv \left(\frac{dx^\alpha}{du} \frac{dx^\beta}{dv} - \frac{dx^\alpha}{dv} \frac{dx^\beta}{du} \right) A_{\alpha\beta}. \tag{4.77}$$

В трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^α этот интеграл можно преобразовать к виду, хорошо знакомому из курса математического анализа. Касательные векторы к поверхности имеют вид

$$\xi = \frac{dx^\alpha}{du} \boldsymbol{e}_\alpha, \qquad \eta = \frac{dx^\alpha}{dv} \boldsymbol{e}_\alpha,$$

где $\boldsymbol{e}_\alpha$ – ортонормированные базисные векторы декартовой системы координат. Их векторное произведение $[\xi,\eta]^\alpha = \varepsilon^{\alpha\beta\gamma} \xi_\beta \eta_\gamma$ нормально к поверхности. Длина нормального вектора равна $\sqrt{n^\alpha n_\alpha} = \sqrt{\xi^2\eta^2 - (\xi\eta)^2}$, что совпадает с $e = \sqrt{g}$, где $g = \det g_{ij}$ – определитель индуцированной метрики на поверхности

$$g_{ij} = \frac{dx^\alpha}{dw^i} \frac{dx^\beta}{dw^j} \delta_{\alpha\beta}, \qquad \{w^i\} = \{u,v\}, \quad i,j = 1,2.$$

Так как в трехмерном евклидовом пространстве 2-форма взаимно однозначно определяется векторным полем $A_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma} X^\gamma$, то интеграл (4.77) можно переписать в виде

$$\frac{1}{2} \int_{f(\mathbb{U})} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \int_{\mathbb{U}} du \wedge dv \sqrt{|g|} (X^\alpha n_\alpha), \tag{4.78}$$

где

$$n^\alpha = \frac{[\xi,\eta]^\alpha}{\sqrt{|g|}}$$

– единичный вектор нормали к поверхности. Таким образом интеграл от 2-формы по поверхности в трехмерном евклидовом пространстве $\mathbb{R}^3$ сводится к суммарному потоку дуального векторного поля через эту поверхность. □

Сформулируем общую теорему, связывающую интеграл от $(n-1)$-формы A по границе некоторой области $\partial\mathbb{U}$ с интегралом от n-формы dA по самой области $\mathbb{U}$.

Теорема 4.6.1 (Формула Стокса). *Пусть $\mathbb{U} \subset \mathbb{M}$ – открытое подмногообразие в $\mathbb{M}$ такое, что его замыкание $\overline{\mathbb{U}}$ компактно. Пусть $\partial\mathbb{U} = \overline{\mathbb{U}} \setminus \mathbb{U}$ – его край, который предполагается кусочно гладким. Тогда справедлива формула Стокса*

$$\int_{\partial\mathbb{U}} A = \int_{\mathbb{U}} dA. \tag{4.79}$$

где $A \in \Lambda_{n-1}(\mathbb{M})$ – произвольная гладкая дифференциальная $(n-1)$-форма на $\mathbb{M}$. При этом предполагается, что ориентация границы задана каноническим образом.

Доказательство. Доказательство формулы Стокса довольно громоздко и приведено, например, в [?]. Каноническая ориентация края определена в доказательстве теоремы 3.3.2. □

Формулу Стокса можно переписать в другом виде. Пусть на многообразии задано векторное поле $X = X^\alpha \partial_\alpha$. Тогда ему можно взаимно однозначно поставить в соответствие $(n-1)$-форму:

$$A = *X = \frac{1}{(n-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\beta\alpha_2\ldots\alpha_n} X^\beta.$$

Внешний дифференциал этой формы равен

$$\begin{aligned} dA &= \frac{1}{(n-1)!} dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_n} \partial_{\alpha_1} (\varepsilon_{\beta\alpha_2\ldots\alpha_n} X^\beta) \\ &= dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} \widetilde{\nabla}_\alpha X^\alpha. \end{aligned}$$

Отсюда следует, что формулу Стокса можно записать в виде

$$\int_{\partial\mathbb{U}} X^\alpha ds_\alpha = \int_{\mathbb{U}} dx^1 \wedge \cdots \wedge dx^n \sqrt{|g|} \widetilde{\nabla}_\alpha X^\alpha, \tag{4.80}$$

где

$$ds_\alpha = \frac{1}{(n-1)!} dx^{\alpha_2} \wedge \cdots \wedge dx^{\alpha_n} \varepsilon_{\alpha\alpha_2\ldots\alpha_n} \tag{4.81}$$

– ориентированный *элемент площади* гиперповерхности.

Рассмотрим частные случаи формулы Стокса.

Пример 4.6.5. Пусть многообразие $\mathbb{M}$ одномерно $n = 1$, а форма $A = f(x)$ имеет нулевую степень. Тогда граница области $\mathbb{U} = (a, b)$ состоит из двух точек $\partial\mathbb{U} = \{a, b\}$, и формула Стокса принимает вид

$$\int_{\partial\mathbb{U}} f = f(b) - f(a) = \int_a^b df = \int_a^b dx \partial_x f.$$

Здесь под интегралом от функции (0-формы) по подмногообразию нулевой размерности, состоящему из отдельных точек, понимается алгебраическая сумма значений функции в этих точках. Таким образом мы получили хорошо известную *формулу Ньютона–Лейбница.* □

Пример 4.6.6. Пусть задана 1-форма $dx^\alpha A_\alpha$ на евклидовой плоскости с декартовыми координатами $\{x^\alpha\} = \{x, y\}$. Рассмотрим область $\mathbb{U}$, ограниченную замкнутой кривой $x^\alpha(t)$. Тогда

$$\int_{\partial\mathbb{U}} dt \frac{dx^\alpha}{dt} A_\alpha = \int_{\partial\mathbb{U}} (dx A_x + dy A_y) = \int_{\mathbb{U}} dx \wedge dy \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right). \tag{4.82}$$

Таким образом мы получили *формулу Грина.* □

Пример 4.6.7. Пусть задана 1-форма $A = dx^\alpha A_\alpha$ в трехмерном евклидовом пространстве с декартовыми координатами x^α. Рассмотрим поверхность $\mathbb{U} \subset \mathbb{R}^3$, с кусочно гладкой границей $\partial\mathbb{U}$. Тогда общая формула Стокса (4.6.1) принимает вид

$$\oint_{\partial\mathbb{U}} dx^\alpha A_\alpha = \int_{\mathbb{U}} \Big[dx^1 \wedge dx^2 (\partial_1 A_2 - \partial_2 A_1) + dx^1 \wedge dx^3 (\partial_1 A_3 - \partial_3 A_1) + dx^2 \wedge dx^3 (\partial_2 A_3 - \partial_3 A_2) \Big].$$

Если поверхность задана параметрически $x^\alpha = x^\alpha(u,v)$, то последний интеграл можно преобразовать к виду

$$\oint_{\partial\mathbb{U}} dx^\alpha A_\alpha = \int_{\mathbb{U}} du \wedge dv \sqrt{|g|}(\,\mathsf{rot}\, A, n), \tag{4.83}$$

где введен ротор (4.64) 1-формы: $[\,\mathsf{rot}\, A]^\alpha = \varepsilon^{\alpha\beta\gamma}\partial_\beta A_\gamma$. Таким образом получена *формула Стокса.* □

Пример 4.6.8. Пусть в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^α задана 2-форма $A = \frac{1}{2} dx^\alpha \wedge dx^\beta A_{\alpha\beta}$. Рассмотрим двумерную компактную поверхность $\mathbb{U} \subset \mathbb{R}^3$ с кусочно дифференцируемой границей $\partial\mathbb{U}$. Будем считать, что поверхность задана параметрически $x^\alpha = x^\alpha(u,v)$. Тогда общая формула Стокса (4.79) принимает вид

$$\frac{1}{2}\int_{\partial\mathbb{U}} dx^\alpha \wedge dx^\beta A_{\alpha\beta} = \int_{\mathbb{U}} dx^1 \wedge dx^2 \wedge dx^3 (\partial_1 A_{23} + \partial_2 A_{31} + \partial_3 A_{12}). \tag{4.84}$$

Как уже отмечалось 2-форма в трехмерном пространстве взаимно однозначно определяется векторным полем $A_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma} X^\gamma$. Интеграл в левой части равенства сводится к интегралу по поверхности от скалярного произведения вектора X^α на единичную нормаль n^α к поверхности (4.78). Подынтегральное выражение в правой части (4.84) представляет собой дивергенцию от векторного поля

$$\partial_1 A_{23} + \partial_2 A_{31} + \partial_3 A_{12} = \partial_\alpha X^\alpha.$$

Таким образом окончательно получаем равенство

$$\int_{\partial\mathbb{U}} du \wedge dv \sqrt{|g|} X^\alpha n_\alpha = \int_{\mathbb{U}} dx^1 \wedge dx^2 \wedge dx^3\, \partial_\alpha X^\alpha, \tag{4.85}$$

где $\sqrt{|g|}$ – квадратный корень из индуцированной метрики (см. пример 4.6.4). Это есть *формула Гаусса–Остроградского* в трехмерном евклидовом пространстве. □

Глава 5

Метрика

Одним из важнейших понятий в геометрии и физике является дифференциально геометрическая метрика или просто метрика. Трудно переоценить ту роль, которую метрика играет в физических приложениях. С ее помощью строятся инварианты, определяется форма объема. Симметрии метрики являются симметриями пространства-времени, которые определяют сохраняющиеся токи. В частности, фундаментальные законы сохранения энергии-импульса и момента количества движения связаны с наличием у метрики определенных симметрий (см. раздел 19.2).

5.1 Определение и свойства

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$.

Определение. *Метрикой* $g \in \mathcal{T}_2(\mathbb{M})$ на многообразии $\mathbb{M}$ называется достаточно гладкое ковариантное тензорное поле типа $(0,2)$, которое является симметричным и невырожденным в каждой точке многообразия:

1) $g_{\alpha\beta} = g_{\beta\alpha}$ – симметричность,

2) $\det g_{\alpha\beta} \neq 0$ – невырожденность,

где $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ – выражение для метрики в координатном базисе. Метрика называется *римановой*, если матрица $g_{\alpha\beta}(x)$ положительно определена во всех точках $x \in \mathbb{M}$. В противном случае метрика называется *псевдоримановой*. □

Замечание. Метрика в данном определении не совпадает с тем же термином, определенном в разделе 2.3.1. Слово метрика широко используется в научной литературе в обоих значениях. Как правило, смысл термина ясен из контекста. Там, где мы хотим подчеркнуть различие терминов, мы будем говорить дифференциально геометрическая метрика для метрики, которая определена в настоящем разделе, и топологическая метрика, определенная в разделе 2.3.1. Далее в подавляющем большинстве случаев термин метрика будет употребляться в дифференциально геометрическом смысле. □

Метрика определяет билинейное невырожденное и симметричное отображение

$$g: \quad \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \ \to \ \mathcal{C}^\infty(\mathbb{M}),$$

которое называется *скалярным произведением* векторных полей. В компонентах скалярное произведение векторных полей $X = X^\alpha \partial_\alpha$ и $Y = Y^\alpha \partial_\alpha$ задается сверткой индексов:

$$(X, Y) := X^\alpha Y^\beta g_{\alpha\beta}, \tag{5.1}$$

для которого мы также будем иногда употреблять обозначение $(X,Y) = g(X,Y)$, чтобы отметить метрику, с помощью которой проводится свертка. Скалярное произведение, очевидно, симметрично $(X,Y) = (Y,X)$. Невырожденность метрики формулируется так: не существует отличного от нуля векторного поля $X \in \mathcal{X}(\mathbb{M})$ такого, что скалярное произведение $(X,Y) = 0$ для всех $Y \in \mathcal{X}(\mathbb{M})$ и во всех точках $x \in \mathbb{M}$. Для римановой метрики скалярное произведение является положительно определенным в каждой точке, т.е. $(X,X) \geq 0$, причем равенство $(X,X) = 0$ имеет место только для тривиального векторного поля $X = 0$. Скалярное произведение базисных векторных полей ∂_α равно компонентам метрики:

$$(\partial_\alpha, \partial_\beta) = g_{\alpha\beta}. \tag{5.2}$$

Поскольку метрика невырождена, то существует *обратная метрика*, т.е. симметричное невырожденное контравариантное тензорное поле $g^{\alpha\beta}\partial_\alpha \otimes \partial_\beta \in \mathcal{T}^2(\mathbb{M})$ типа $(2,0)$, компоненты которого в каждой точке многообразия удовлетворяет условию

$$g^{\alpha\beta}g_{\beta\gamma} = g_{\gamma\beta}g^{\beta\alpha} = \delta^\alpha_\gamma. \tag{5.3}$$

При преобразовании координат $x^\alpha \to x^{\alpha'}(x)$ компоненты метрики и ее обратной преобразуются по обычным правилам для тензорных полей:

$$g_{\alpha'\beta'} = \frac{\partial x^\alpha}{\partial x^{\alpha'}}\frac{\partial x^\beta}{\partial x^{\beta'}}g_{\alpha\beta}, \tag{5.4}$$

$$g^{\alpha'\beta'} = g^{\alpha\beta}\frac{\partial x^{\alpha'}}{\partial x^\alpha}\frac{\partial x^{\beta'}}{\partial x^\beta}. \tag{5.5}$$

Отсюда следует, что определитель метрики

$$g(x) = \det g_{\alpha\beta}(x) \tag{5.6}$$

преобразуется по правилу $g' = J^{-2}g$, где J – якобиан преобразования координат (2.58). То есть определитель метрики является скалярной плотностью степени -2. В общем случае, ввиду симметрии по индексам, метрика $g_{\alpha\beta}$ задается

$$[g_{\alpha\beta}] = \frac{n(n+1)}{2}$$

произвольными компонентами с единственным условием невырожденности. Здесь и в дальнейшем число компонент тензора мы будем обозначать квадратными скобками.

С помощью метрики и ее обратной можно изменить тип тензора путем опускания или подъема всех или части индексов произвольного тензорного поля.

Пример 5.1.1. Произвольному векторному полю $X \in \mathcal{X}(\mathbb{M})$ можно взаимно однозначно поставить в соответствие 1-форму и наоборот. В компонентах отображение задается следующими формулами:

$$X_\alpha = X^\beta g_{\beta\alpha}, \qquad X^\alpha = g^{\alpha\beta}X_\beta. \tag{5.7}$$

Это отображение $\mathcal{X}(\mathbb{M}) \leftrightarrow \Lambda_1(\mathbb{M})$, очевидно, линейно, не зависит от выбора системы координат и, значит, является изоморфизмом $\mathcal{C}^\infty(\mathbb{M})$-модулей. Операция подъема и опускания индексов естественным образом продолжается на тензоры произвольного типа. □

Замечание. Если на многообразии задана метрика, то между контравариантными и ковариантными тензорными полями одного ранга r существует взаимно однозначное соответствие, точнее, изоморфизм $\mathcal{C}^\infty(\mathbb{M})$-модулей:

$$\mathcal{T}^r(\mathbb{M}) \simeq \mathcal{T}_r(\mathbb{M}) \simeq \mathcal{T}^p_q(\mathbb{M}), \qquad \forall p,q: \ p+q=r.$$

По этой причине будем считать, что тензоры одного ранга, но с различным типом индексов, описывают один и тот же геометрический объект и будем обозначать их, как правило, одной буквой. Например,

$$T_{\alpha\beta} = T_\alpha{}^\gamma g_{\gamma\beta} = T^\gamma{}_\beta g_{\gamma\alpha} = T^{\gamma\delta} g_{\gamma\alpha} g_{\delta\beta},$$

где $T_{\alpha\beta}$ и $T^{\alpha\beta}$ – ковариантные и контравариантные компоненты тензора второго ранга. Компоненты $T_\alpha{}^\beta$ и $T^\alpha{}_\beta$ называются смешанными. Поскольку при наличии метрики появилась возможность опускать и поднимать индексы, то необходимо следить за порядком ковариантных и контравариантных индексов. В общем случае $T_\alpha{}^\beta \neq T^\beta{}_\alpha$. □

Определение. Скалярное произведение вектора в каждой точке $x \in \mathbb{M}$ с самим собой называется *квадратом* вектора и обозначается

$$X^2 := (X, X). \tag{5.8}$$

Длиной вектора X назовем выражение

$$|X| := \sqrt{|X^2|}. \tag{5.9}$$

Знак модуля в этом определении необходим в том случае, если метрика не является положительно определенной. Пусть X, Y – два векторных поля на римановом многообразии $\mathbb{M}$. Если в точке $x \in \mathbb{M}$ они отличны от нуля, то определим *угол* φ между ними в данной точке с помощью следующего соотношения

$$\cos\varphi := \frac{(X,Y)}{\sqrt{X^2 Y^2}}. \quad \square \tag{5.10}$$

Определение угла (5.10) корректно. Действительно, для положительно определенной метрики правая часть соотношения (5.10) не превышает единицу, поскольку справедливо неравенство треугольника

$$(X,Y) \leq \sqrt{X^2 Y^2}.$$

Определение угла инвариантно относительно выбора координат, т.к. правая часть (5.10) содержит только инвариантные комбинации векторных полей.

Определение. Если $(X,Y) = 0$, то два вектора называются *ортогональными* или *перпендикулярными*, независимо от положительной определенности метрики. □

Угол между векторами определяется только метрикой и не изменится, если метрику умножить на произвольную функцию, отличную от нуля (преобразование Вейля, раздел **??**).

Определение. Две метрики $g' = dx^\alpha \otimes dx^\beta g'_{\alpha\beta}$ и $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ на многообразии $\mathbb{M}$ называются *конформно эквивалентными* или инвариантными относительно преобразований Вейля, если они пропорциональны

$$g' = \mathsf{e}^{2\phi} g, \tag{5.11}$$

где $\phi(x)$ – некоторая функция на $\mathbb{M}$. Преобразование (5.11) называют *конформным преобразованием* метрики. □

Это определение не зависит от выбора системы координат и поэтому задает вейлевскую (конформную) эквивалентность глобально.

Замечание. Термин конформная эквивалентность метрик широко используется, однако он не имеет отношения к конформным преобразованиям в комплексном анализе. Конформные преобразования комплексных переменных – это подпсевдогруппа общих преобразований координат, в то время как при преобразовании (5.11) координаты не меняются. Поэтому мы будем употреблять термин вейлевская инвариантность, так как в двумерных моделях математической физики термин конформная инвариантность употребляется в своем первоначальном значении, как в комплексном анализе. □

Поскольку метрика невырождена, то у нее существует обратная метрика (5.3), которая естественным образом определяет скалярное произведение в пространстве 1-форм. Скалярное произведение двух 1-форм $A = dx^\alpha A_\alpha$ и $B = dx^\beta B_\beta$ по определению равно

$$(A, B) = g(A, B) := g^{\alpha\beta} A_\alpha B_\beta. \tag{5.12}$$

Это скалярное произведение согласовано с операцией подъема и опускания индексов (5.7):

$$g^{\alpha\beta} A_\alpha B_\beta = g_{\alpha\beta} A^\alpha B^\beta, \quad \text{где} \quad A^\alpha = g^{\alpha\beta} A_\beta, \quad B^\alpha = g^{\alpha\beta} B_\beta.$$

Скалярные произведения векторов (5.1) и 1-форм (5.12) естественным образом продолжаются на тензорные поля произвольного типа (r, s).

Метрика на многообразии определяет инвариантную квадратичную форму дифференциалов, которая называется *интервалом*

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta. \tag{5.13}$$

Интервал задает расстояние между двумя бесконечно близкими точками с координатами x^α и $x^\alpha + dx^\alpha$. Это расстояние зависит от точки многообразия, но не от выбора системы координат. В случае псевдоримановой метрики выражение (5.13) не задает метрику (расстояние) в топологическом смысле, определенную в разделе 2.3.1, т.к. квадратичная форма (5.13) не является положительно определенной. С помощью положительно определенной римановой метрики топологическое расстояние между двумя точками $p, q \in \mathbb{M}$ можно определить как точную нижнюю грань интеграла

$$l(p, q) = \inf \int_p^q ds, \qquad ds = \sqrt{ds^2} = dt \sqrt{g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}}, \tag{5.14}$$

где интегрирование ведется по всем кусочно дифференцируемым кривым $x(t)$, соединяющим точки p и q. Доказательство свойств расстояния при этом существенно опирается на положительную определенность римановой метрики. Топология, определяемая топологической метрикой (5.14), совпадает с топологией многообразия $\mathbb{M}$

Теорема 5.1.1. *На любом n-мерном дифференцируемом многообразии $\mathbb{M}$ существует риманова метрика.*

Доказательство. Выберем локально конечный атлас $\{\mathbb{U}_i\}$ на $\mathbb{M}$. Такой атлас всегда существует в силу теоремы 3.2.2. Обозначим координаты на каждой карте через x_i^α, $\alpha = 1, \ldots, n$. Пусть $\{f_i\}$ – разбиение единицы, подчиненное данному атласу, такое, что $\mathsf{supp}\, f_i \subset \mathbb{U}_i$. Выберем евклидову метрику на каждой карте и склеим карты с помощью разбиения единицы

$$ds_i^2 = \delta_{\alpha\beta} dx_i^\alpha dx_i^\beta, \tag{5.15}$$

$$ds^2 = \sum_i f_i ds_i^2, \tag{5.16}$$

где выражение $f_i ds_i^2$ определено

$$(f_i ds_i^2)(x) = \begin{cases} f_i(x) ds_i^2, & x \in \mathbb{U}_i, \\ 0, & x \notin \mathbb{U}_i. \end{cases}$$

Уравнения (5.15), (5.16) определяют гладкое симметричное ковариантное тензорное поле второго ранга на $\mathbb{M}$. Поскольку правая часть (5.16) содержит конечное число слагаемых в каждой точке $x \in \mathbb{M}$, то сумма корректно определена. Выберем координатную окрестность $\mathbb{U}$ с координатами x^α такую, что замыкание области $\overline{\mathbb{U}}$ является компактным. Тогда область $\mathbb{U}$ пересекается с конечным числом карт $\mathbb{U}_{i_1}, \ldots, \mathbb{U}_{i_r}$, т.к. атлас является локально конечным покрытием. Поэтому ограничение (5.16) на $\mathbb{U}$ можно записать в виде

$$ds^2 = \sum_{\lambda=1}^{r} f_{i_\lambda} ds_{i_\lambda}^2 = g_{\alpha\beta} dx^\alpha dx^\beta,$$

где

$$g_{\alpha\beta} = \sum_{\lambda=1}^{r} f_{i_\lambda} \delta_{\gamma\delta} \frac{\partial x_{i_\lambda}^\gamma}{\partial x^\alpha} \frac{\partial x_{i_\lambda}^\delta}{\partial x^\beta}.$$

Поскольку $0 \le f_i \le 1$ и $\sum_i f_i = 1$, то существует такой индекс j, что $f_j(x) > 0$. Поэтому

$$ds^2(x) \ge f_j ds_j^2.$$

Таким образом метрика ds^2 является положительно определенной на $\mathbb{M}$. □

Замечание. Доказательство этой теоремы свелось к переносу евклидовой метрики δ_{ij} из евклидова пространства $\mathbb{R}^n$ на многообразие $\mathbb{M}$ с помощью возврата отображения, которое фигурирует в определении многообразия, как гомеоморфизм. При этом положительная определенность евклидовой метрики существенна. Доказательство не проходит, если таким же образом попытаться перенести лоренцеву метрику из $\mathbb{R}^{1,n-1}$ на $\mathbb{M}$. □

5.2 Метрика на лоренцевых многообразиях

За счет выбора системы координат метрику всегда можно привести к диагональному виду в любой наперед заданной точке $x \in \mathbb{M}$. Действительно, при замене координат компоненты метрики преобразуются по-правилу (5.4). При этом матрица Якоби преобразования координат $\partial x^\alpha / \partial x^{\alpha'}$ в фиксированной точке многообразия может быть выбрана произвольным образом.

Пример 5.2.1. Однородное линейное преобразование координат $x^\alpha = x^{\alpha'} M_{\alpha'}{}^{\alpha}$ с постоянной матрицей $M_{\alpha'}{}^{\alpha} = \mathsf{const}$ дает

$$g_{\alpha'\beta'} = M_{\alpha'}{}^{\alpha} M_{\beta'}{}^{\beta} g_{\alpha\beta}. \qquad \square$$

Замечание. В общем случае метрику можно привести к диагональному виду в фиксированной точке, но не в окрестности. Это связано с тем, что n функций преобразований координат, которыми можно воспользоваться, недостаточно для фиксирования $n(n-1)/2$ функций, параметризующих недиагональные элементы метрики. Исключение составляют многообразия двух и трех измерений. В двумерном случае метрика имеет только одну недиагональную компоненту и ее можно привести к диагональному виду в окрестности произвольной точки. Более того, ее можно преобразовать к конформно плоскому виду. На трехмерном многообразии метрика имеет три недиагональные компоненты, что равно числу произвольных функций, параметризующих преобразования координат. Можно показать, что ее также можно привести к диагональному виду не только в заданной точке, но и в некоторой окрестности этой точки. $\square$

Привести метрику к диагональному виду в точке можно многими способами. При этом сигнатура не зависит от выбора системы координат в которой метрика диагональна. Действительно, если в точке $p \in \mathbb{M}$ метрика диагональна в двух системах координат и при этом имеет различную сигнатуру, то эти системы координат не могут быть связаны невырожденным преобразованием. В различных системах координат диагональные компоненты метрики могут иметь различные (ненулевые) значения, может меняться последовательность положительных и отрицательных компонент, однако число положительных и отрицательных компонент остается неизменным.

Если $g_{\alpha\beta}$ – матрица, составленная из компонент метрики в некоторой системе координат, то числа p и q равны, соответственно, числу положительных и отрицательных собственных значений. При общих преобразованиях координат собственные числа могут менять свою величину, но не знак. Нулевых собственных значений быть не может, т.к. в этом случае метрика была бы вырожденной.

Определение. *Сигнатурой метрики*, заданной на многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, называется пара натуральных чисел (p, q) таких, что $p + q = n$, где p и q – количество, соответственно, положительных и отрицательных чисел, стоящих на диагонали метрики после ее диагонализации в какой либо точке многообразия $x \in \mathbb{M}$. $\square$

Предложение 5.2.1. *Сигнатура метрики не зависит от точки линейно связного многообразия.*

Доказательство. Допустим, что в некоторых точках $p, q \in \mathbb{M}$ метрика имеет разную сигнатуру, и соединим эти точки произвольной кривой. Тогда из непрерывности метрики следует, что ее определитель обратился бы в нуль в некоторой точке кривой, что недопустимо. $\square$

Замечание. В зависимости от сигнатуры метрики скалярное произведение векторов может быть положительно определено или нет. Риманова метрика на многообразии размерности n имеет сигнатуру $(n, 0)$. Отрицательно определенная метрика имеет сигнатуру $(0, n)$. В римановой геометрии $(\mathbb{M}, g)$ положительно и отрицательно определенные метрики, по существу, можно не различать, т.к. все геометрические характеристики многообразий связаны простым преобразованием – изменением знака

метрики. Поэтому многообразия с отрицательно определенной метрикой мы также будем называть римановыми. В то же время многие модели математической физики неинвариантны относительно изменения знака метрики

$$g_{\alpha\beta} \to -g_{\alpha\beta}, \tag{5.17}$$

поскольку действие, которым описывается модель, помимо метрики содержит также другие поля. Поэтому решения уравнений Эйлера–Лагранжа с положительно и отрицательно определенной метрикой, вообще говоря, неэквивалентны. □

Сигнатура метрики инвариантна, не зависит от выбора системы координат и точки линейно связного многообразия. Будем считать, что после диагонализации номера координат выбраны таким образом, что сначала идут все положительные, а затем – отрицательные собственные значения. Будем писать

$$\mathsf{sign}\, g_{\alpha\beta} = (\underbrace{+\cdots+}_{p}\underbrace{-\cdots-}_{q}). \tag{5.18}$$

Определение. Пара $(\mathbb{M}, g)$ называется *римановым многообразием*, если метрика является знакоопределенной $p = n$ или $q = n$, и *псевдоримановым многообразием*, если метрика не является знакоопределенной, $p \neq 0$ и $q \neq 0$. Если положительный элемент на диагонали один,

$$\mathsf{sign}\, g_{\alpha\beta} = (+ - \cdots -),$$

то говорят, что метрика имеет *лоренцеву сигнатуру*. Если на многообразии задана лоренцева метрика, то будем говорить, что многообразие *лоренцево*. □

Замечание. Подчеркнем, что задание метрики на многообразии может быть произвольным. В частности, на одном многообразии можно задать несколько метрик одновременно, причем разной сигнатуры, если такие существуют. □

В настоящей монографии псевдоримановы многообразия размерности $n \geq 4$ с $p \geq 2$ и $q \geq 2$ не рассматриваются, так как они недостаточно хорошо изучены и не имеют широкого применения в математической физике.

Для лоренцевой метрики скалярное произведение двух векторов может быть положительно, отрицательно или равно нулю, а из условия $X^2 = 0$ не следует, что $X = 0$.

Определение. Назовем *пространством-временем* псевдориманово многообразие $\mathbb{M}$ с заданной метрикой $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ лоренцевой сигнатуры (лоренцево многообразие). В пространстве-времени векторное поле X в точке $x \in \mathbb{M}$ называется:

$$\begin{array}{lll} \textit{времениподобным,} & \text{если} & (X,X) > 0, \\ \textit{светоподобным (изотропным, нулевым),} & \text{если} & (X,X) = 0, \\ \textit{пространственноподобным,} & \text{если} & (X,X) < 0. \end{array} \tag{5.19}$$

Это определение распространятся на область $\mathbb{U} \subset \mathbb{M}$, если во всех точках $x \in \mathbb{U}$ выполнены соответствующие соотношения. □

Замечание. Определение (5.19) инвариантно относительно замены координат. В общем случае гладкое векторное поле может иметь различный тип в различных областях связного многообразия. □

Определение. Так же, как и в римановом пространстве, два вектора в пространстве-времени называются *ортогональными (перпендикулярными)*, если их скалярное произведение равно нулю. В частности, любой изотропный вектор перпендикулярен самому себе.

Определим тип координатного векторного поля ∂_0. Из (5.2) следует, что $(\partial_0, \partial_0) = g_{00}$. Если в данной системе координат $g_{00} > 0$, то векторное поле ∂_0 времениподобно. В этом случае назовем координату $x^0 = t$ *времениподобной* или *временем*. Другими словами, временем называется любой параметр вдоль интегральной кривой произвольного времениподобного векторного поля. Противоположно направленное векторное поле $-\partial_0$ также времениподобно и связано с ∂_0 преобразованием координат $(x^0, x^1, \dots, x^{n-1}) \to (-x^0, x^1, \dots, x^{n-1})$, которое назовем *обращением времени*. Выберем (произвольно) ориентацию координаты x^0 и будем говорить, что векторное поле ∂_0 направлено в будущее, а $-\partial_0$ – в прошлое. Тогда на времениподобных векторных полях можно ввести ориентацию. Произвольное времениподобное поле X направлено в будущее, если $(X, \partial_0) = X^0 > 0$. В противном случае, когда $X^0 < 0$, будем говорить, что времениподобное векторное поле направлено в прошлое. □

Замечание. В общем случае метрика, индуцированная на сечениях $t = \mathsf{const}$, может быть отрицательно определена или быть знаконеопределенной. То есть сечения пространства-времени, соответствующие постоянному времени, совсем не обязательно пространственноподобны (= все касательные векторы пространственноподобны). Это зависит от выбора остальных координат. В конце настоящего раздела мы рассмотрим простой пример. □

Можно доказать, что каждая точка многообразия $\mathbb{M}$ лоренцевой сигнатуры $(1, n-1)$ имеет такую координатную окрестность с координатами $(t, u^1, \dots, u^{n-1})$, что в этих координатах метрика имеет вид:

$$ds^2 = dt^2 + g_{\mu\nu}(t, u) du^\mu du^\nu, \qquad \mu, \nu = 1 \dots n-1. \tag{5.20}$$

При этом метрика $g_{\mu\nu}$, индуцированная на сечениях $t = \mathsf{const}$, является отрицательно определенной. Другими словами, все сечения $t = \mathsf{const}$ для метрики вида (5.20) являются пространственноподобными подмногообразиями, т.е. все касательные векторы к сечениями пространственноподобны. Эти сечения являются вложенными римановыми многообразиями с локальной системой координат $\{u^\mu\}$ и отрицательно определенной метрикой $g_{\mu\nu}$. Отметим также, что все векторы, касательные к этому подмногообразию ортогональны времениподобному векторному полю ∂_0.

Замечание. Здесь и в дальнейшем мы будем использовать следующие обозначения координат на лоренцевом многообразии. Греческие буквы из начала алфавита $\alpha, \beta, \dots$ будут использоваться для нумерации всех координат, а буквы из середины алфавита $\mu, \nu, \dots$ – для нумерации остальных координат, $\{x^\alpha\} = \{x^0, x^\mu\}$. Это правило легко запомнить по следующим включениям:

$$\{1, \dots, n-1\} \subset \{0, 1, \dots, n-1\}, \qquad \{\mu, \nu, \dots\} \subset \{\alpha, \beta, \dots\}.$$

Как правило, мы будем считать, что координата x^0 является временем, и все остальные координатные линии x^μ – пространственноподобны. Поэтому координаты x^μ будем называть пространственноподобными. □

В общем случае метрика на лоренцевом многообразии

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & g_{0\nu} \\ g_{\mu 0} & g_{\mu\nu} \end{pmatrix} \tag{5.21}$$

не имеет блочно диагонального вида (5.20). Приведем критерий того, что метрика (5.21), заданная на некотором многообразии $\mathbb{M}$, имеет лоренцеву сигнатуру.

Теорема 5.2.1. *Пусть в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ задана метрика (5.21) такая, что $g_{00} > 0$. Эта метрика имеет лоренцеву сигнатуру тогда и только тогда, когда матрица*

$$g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}} \tag{5.22}$$

отрицательно определена в каждой точке $\mathbb{U}$.

Доказательство. Достаточно рассмотреть произвольную точку из $\mathbb{U}$. Пусть на $\mathbb{U}$ задана метрика (5.21) для которой $g_{00} > 0$. Интервал в окрестности $\mathbb{U} \subset \mathbb{M}$ имеет вид

$$ds^2 = g_{00}dx^0dx^0 + 2g_{0\mu}dx^0dx^\mu + g_{\mu\nu}dx^\mu dx^\nu, \qquad g_{00} > 0.$$

Зафиксируем произвольную точку $x \in \mathbb{U}$. Введем вместо x^0 новую координату $\tilde{x}^0$, для которой в точке $x \in \mathbb{U}$ выполнено соотношение

$$dx^0 = d\tilde{x}^0 - \frac{dx^\mu g_{0\mu}}{g_{00}}.$$

В фиксированной точке этого всегда можно добиться линейным преобразованием координат. Тогда интервал примет вид

$$ds^2 = g_{00}d\tilde{x}^0d\tilde{x}^0 + \left(g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}\right)dx^\mu dx^\nu. \tag{5.23}$$

Если метрика (5.21) имеет лоренцеву сигнатуру, то существует такая система координат, что в точке $x \in \mathbb{M}$ метрика диагональна, причем $g_{00} > 0$, а все остальные диагональные компоненты $g_{\mu\mu}$ отрицательны. Поскольку метрика (5.23) связана с диагональной метрикой также невырожденным преобразованием координат, то матрица (5.22) отрицательно определена.

Обратно. Если матрица (5.22) отрицательно определена, то дальнейшим линейным преобразованием координат x^μ ее всегда можно преобразовать к диагональному виду в фиксированной точке, причем на диагонали будут стоять отрицательные числа. Следовательно, метрика имеет лоренцеву сигнатуру. □

Замечание. В разделе 22.2 мы докажем, что отрицательная определенность матрицы (5.22) эквивалентна отрицательной определенности "пространственного блока" $g^{\mu\nu}$ обратной метрики $g^{\alpha\beta}$. □

Матрица (5.22) на лоренцевом многообразии симметрична и невырождена, т.к. отрицательно определена. Она имеет следующий геометрический смысл. Вдоль координатных линий времени $t = x^0$ всегда можно определить единичное векторное поле

$$n := \frac{1}{\sqrt{g_{00}}}\partial_0, \qquad (n, n) = 1.$$

Рассмотрим произвольное векторное поле $X = X^0\partial_0 + X^\mu\partial_\mu$. У него есть составляющая, перпендикулярная времениподобному векторному полю n,

$$X_\perp = X - (X, n)n = X^\mu\partial_\mu - X^\mu g_{0\mu}\partial_0.$$

Если задано два произвольных векторных поля X и Y, то скалярное произведение их перпендикулярных составляющих равно

$$(X_\perp, Y_\perp) = \left(g_{\mu\nu} - \frac{g_{0\mu} g_{0\nu}}{g_{00}} \right) X^\mu Y^\nu.$$

Таким образом, матрица (5.22) играет роль метрики для перпендикулярных составляющих векторных полей. Из теоремы 5.2.1 следует, что, если $\mathbb{M}$ – лоренцево многообразие и x^0 – время, то составляющие векторных полей, перпендикулярные n, всегда пространственноподобны.

Замечание. Не следует думать, что матрицу (5.22) можно рассматривать, как метрику на некотором $(n-1)$-мерном подмногообразии, касательные векторы к которому всюду перпендикулярны времениподобному векторному полю n. Дело в том, что такие подмногообразия могут не существовать. Можно проверить, что коммутатор двух перпендикулярных векторных полей $[X_\perp, Y_\perp]$ в общем случае не будет ортогонален n. Для этого достаточно, не ограничивая общности, положить $g_{00} = 1$ и проверить, что коммутатор $[X_\perp, Y_\perp]$ имеет нетривиальную составляющую вдоль n. Тем самым множество всех векторных полей $X_\perp$ не находится в инволюции, и, согласно теореме Фробениуса, подмногообразия, касательные векторы к которому всюду перпендикулярны n, не существует. □

Приведем еще один критерий того, что метрика на многообразии является лоренцевой.

Теорема 5.2.2. *Для того, чтобы метрика (5.21) с $g_{00} > 0$ в точке $x \in \mathbb{M}$ имела лоренцеву сигнатуру необходимо и достаточно, чтобы выполнялись неравенства:*

$$\det \begin{pmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{pmatrix} < 0, \quad \det \begin{pmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{pmatrix} > 0, \quad \ldots, \quad \det (g_{\alpha\beta}) \begin{cases} > 0, & n - \textit{нечетно}, \\ < 0, & n - \textit{четно}. \end{cases} \tag{5.24}$$

Доказательство. См., например, []. □

Пример 5.2.2. Рассмотрим плоскость Минковского $\mathbb{R}^{1,1}$ с декартовой системой координат $\{x^0, x^1\}$ и метрикой

$$\eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Введем новую систему координат $\{\tilde{x}^0, \tilde{x}^1\} = \{x^0 - \sqrt{2}x^1, x^1\}$. Тогда координатные базисные векторы преобразуются по-правилу (см. рис. 5.1)

$$\{\partial_0, \partial_1\} \mapsto \{\tilde{\partial}_0, \tilde{\partial}_1\} = \{\partial_0, \sqrt{2}\partial_0 + \partial_1\}.$$

В новой системе координат сечения $\tilde{x}^0 = \mathsf{const}$ являются времениподобными прямыми, а метрика имеет вид

$$g_{\alpha\beta} = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix}.$$

Матрица, соответствующая (5.22), состоит из одного элемента $g_{11} - g_{01}^2/g_{00} = -1$.

Обратная метрика имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} -1 & \sqrt{2} \\ \sqrt{2} & -1 \end{pmatrix}.$$

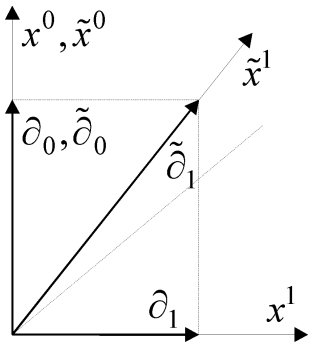


Рис. 5.1: Пример времениподобного сечения $\tilde{x}^0 = \mathsf{const}$.

Отметим, что в этих координатах обе компоненты обратной метрики на диагонали g^{00} и g^{11} отрицательны. □

Замечание. Рассмотренный пример показывает, что "пространственные" компоненты метрики $g_{\mu\nu}$ в общем случае не образуют отрицательно определенной матрицы. Роль пространственной части метрики играет выражение (5.22), и она должна быть отрицательно определена на лоренцевом многообразии. Кроме того, из примера следует, что, даже если координата x^0 на лоренцевом многообразии является временем, тем не менее временна́я компонента обратной метрики g^{00} может быть отрицательна. □

В дальнейшем мы всегда предполагаем, что координаты на лоренцевом многообразии выбраны таким образом, что координата x^0 является временем, и все сечения $x^0 = \mathsf{const}$ пространственноподобны. Такой выбор координат удобен, например, при постановке задачи Коши в различных моделях математической физики. На языке компонент метрики эти условия означают, что $g_{00} > 0$, и матрица $g_{\mu\nu}$ отрицательно определена. В этом случае нетрудно доказать, что временна́я компонента обратной метрики также положительна $g^{00} > 0$. В обратную сторону утверждение неверно: условий положительности $g_{00} > 0$ и $g^{00} > 0$ недостаточно для отрицательной определенности матрицы $g_{\mu\nu}$.

5.3 Векторные поля и вложения

В разделе 3.6.6 было установлено, что отличная от нуля 1-форма $A = dx^\alpha A_\alpha$ определяет в касательном расслоении $\mathbb{T}(\mathbb{M})$ распределение $(n-1)$-подпространств, натянутое на векторы Y^α, удовлетворяющие уравнению $Y^\alpha A_\alpha = 0$. Наличие метрики позволяет каждой 1-форме поставить в соответствие векторное поле $X^\alpha = g^{\alpha\beta} A_\beta$. По-построению, это векторное поле перпендикулярно любому вектору из $(n-1)$-мерного подпространства в касательном пространстве, т.к. $X^\alpha dx^\beta g_{\alpha\beta} = 0$. Таким образом, при наличии метрики, задание ненулевого векторного поля эквивалентно заданию $(n-1)$-мерного распределения на многообразии. В общем случае это распределение не является инволютивным, и поэтому у него не существует интегральных подмногообразий. То есть векторное поле не определяет ортогональные интегральные подмногообразия. Обратное утверждение верно. Каждое $(n-1)$-мерное подмногообразие (гиперповерхность) определяет нормальное векторное поле, заданное на подмногообразии.

Определение. Пусть $\mathbb{M}$ – риманово многообразие. Тогда *углом* между двумя пересекающимися гиперповерхностями в точках их пересечения называется угол между соответствующими нормальными векторами в данных точках. □

Определение. Если на многообразии задана метрика и векторное поле, то определим проекционные операторы на заданное векторное поле и перпендикулярные направления следующим образом. Пусть X^α – компоненты произвольного векторного поля, такого, что $X^2 \neq 0$. Тогда операторы

$$\Pi^{\text{L}}{}_\alpha^\beta = \frac{X_\alpha X^\beta}{X^2}, \qquad \Pi^{\text{T}}{}_\alpha^\beta = \delta_\alpha^\beta - \frac{X_\alpha X^\beta}{X^2}, \tag{5.25}$$

являются *проекционными операторами* на направление векторного поля и перпендикулярные направления, соответственно. При этом выполнены равенства, определяющие набор проекционных операторов:

$$(\Pi^{\text{L}})^2 = \Pi^{\text{L}}, \quad (\Pi^{\text{T}})^2 = \Pi^{\text{T}}, \quad \Pi^{\text{L}}\Pi^{\text{T}} = 0, \quad \Pi^{\text{L}} + \Pi^{\text{T}} = 1. \quad \square$$

Поскольку нетривиальные ковекторы

$$V_\alpha - X_\alpha \frac{(V,X)}{X^2} \quad \text{и} \quad X_\alpha$$

являются собственными векторами проекционных операторов Π^{L} и Π^{T} с нулевыми собственными значениями, то проекционные операторы вырождены:

$$\mathsf{det}\,\Pi^{\text{L}} = \mathsf{det}\,\Pi^{\text{T}} = 0.$$

Рассмотрим, как меняется метрика при отображениях. Пусть многообразие $\mathbb{M}$ вложено[1] в $\mathbb{N}$, $\mathsf{dim}\,\mathbb{M} \leq \mathsf{dim}\,\mathbb{N}$, и на пространстве мишени $\mathbb{N}$ задана метрика g. Тогда возврат отображения $f: \mathbb{M} \hookrightarrow \mathbb{N}$ задает метрику на $\mathbb{M}$. Бескоординатное определение индуцированной метрики f^*g имеет вид

$$(f^*g)(X,Y) = g(f_*X, f_*Y), \quad \forall X, Y \in \mathcal{X}(\mathbb{M}),$$

где f_* – дифференциал отображения, отображающий векторные поля X, Y на $\mathbb{M}$ в векторные поля f_*X, f_*Y на $\mathbb{N}$. Скалярные произведения на $\mathbb{M}$ и $\mathbb{N}$ заданы соответственно метриками f^*g и g. Это значит, что при вложении длины векторов и углы между ними сохраняются. Обозначив координаты на $\mathbb{M}$ и $\mathbb{N}$ через x^α и y^{A}, получим явное выражение для компонент индуцированной метрики f^*g:

$$(f^*g)_{\alpha\beta}(x) = \frac{\partial y^{\text{A}}}{\partial x^\alpha}\frac{\partial y^{\text{B}}}{\partial x^\beta} g_{\text{AB}}(y). \tag{5.26}$$

Определение. Пусть на многообразии $\mathbb{M}$ задана некоторая метрика h. Если при вложении $\mathbb{M}$ в многообразие $\mathbb{N}$ с метрикой g исходная метрика совпадает с индуцированной, $h = f^*g$, то такое вложение называется *изометрическим*. □

Представляет большой интерес задача о нахождении изометрических вложений заданного многообразия с определенной на нем метрикой в евклидово пространство.

[1] На самом деле достаточно погружения.

Пример 5.3.1. Пусть двумерная сфера $\mathbb{S}^2_r$ радиуса r вложена в трехмерное евклидово пространство $\mathbb{R}^3$. Вложение можно записать в виде

$$x = r\sin\theta\cos\varphi,$$
$$y = r\sin\theta\sin\varphi,$$
$$z = r\cos\varphi,$$

где $0 \le \theta \le \pi$ и $0 \le \varphi < 2\pi$ – координаты на сфере. Тогда метрика, индуцированная на сфере, примет вид

$$ds^2 = dx^2 + dy^2 + dz^2 = r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2. \tag{5.27}$$

То есть в выражение для евклидова интервала вместо дифференциалов dx, dy и dz необходимо просто подставить их выражения через дифференциалы полярного и азимутального углов. □

Пример 5.3.2. Пусть двумерный тор $\mathbb{T}^2$ радиуса r с направляющей окружностью радиуса $R > r$ вложен в трехмерное евклидово пространство $\mathbb{R}^3$ (см. рис. 5.2):

$$x = (R + r\cos\theta)\cos\varphi,$$
$$y = (R + r\cos\theta)\sin\varphi,$$
$$z = r\sin\theta,$$

где $0 \le \theta < 2\pi$ и $0 \le \varphi < 2\pi$ – координаты на торе. Тогда индуцированная метрика имеет вид

$$ds^2 = dx^2 + dy^2 + dz^2 = r^2 d\theta^2 + (R + r\cos\theta)^2 d\varphi^2. \quad \square \tag{5.28}$$

Рис. 5.2: Двумерный тор, вложенный в трехмерное евклидово пространство. ftorin

5.4 Выбор системы координат

Если на многообразии задана метрика, которую преобразованием координат нельзя привести к евклидовой или лоренцевой форме, тогда понятие декартовой системы координат отсутствует, и это создает определенные трудности для наглядного представления о многообразии. С практической точки зрения удобнее сначала совершить преобразование координат, после которого метрика примет какой-либо относительно простой вид, связанный, например, с симметрией задачи, и уже потом дать физическую интерпретацию пространству или пространству-времени.

Определение. Риманова метрика называется *конформно евклидовой* в некоторой области $\mathbb{U} \subset \mathbb{M}$, если существует такая система координат на $\mathbb{U}$, в которой метрика и ее обратная имеют вид

$$g_{\alpha\beta} = \mathrm{e}^{2\phi}\delta_{\alpha\beta}, \qquad g^{\alpha\beta} = \mathrm{e}^{-2\phi}\delta^{\alpha\beta}. \tag{5.29}$$

Здесь $\phi(x)$ – некоторая достаточно гладкая функция от x^α. Функция $\mathrm{e}^{2\phi}$ называется *конформным множителем*. В таком виде метрика и ее обратная невырождены при $\phi \neq \pm\infty$. Если у каждой точки $x \in \mathbb{M}$ существует окрестность, на которой метрика является конформно евклидовой, то метрика называется *локально конформно евклидовой*. □

Замечание. Подчеркнем, что $\phi(x)$ не является скалярным полем, хотя и не имеет ни одного индекса, т.к. в другой системе координат конформно евклидова метрика может не иметь вида (5.29), и, значит, функция ϕ не определена вовсе. □

Если метрика записана в конформно евклидовом виде, то в этой системе координат угол (5.10) между двумя произвольными векторами тот же, что и в евклидовом пространстве. Это обстоятельство позволяет более наглядно представить себе свойства многообразия.

Конформно евклидова метрика является частным случаем метрики. Это видно уже из того, что в общем случае она зависит от $n+1$ независимой функции (здесь n функций соответствует выбору системы координат и одна функция – конформному множителю), в то время как метрика общего вида зависит от $n(n+1)/2$ функций. Исключение представляет только случай двух измерений.

Теорема 5.4.1. *Если риманова метрика принадлежит классу $\mathcal{C}^3(\mathbb{M})$ на двумерном многообразии $\mathbb{M}$ (поверхности), то для любой точки $x \in \mathbb{M}$ существует окрестность $\mathbb{U}_x$, в которой можно выбрать конформно евклидовы координаты.*

Доказательство. См., например, [**?**]. □

Замечание. Конформно евклидовы координаты x^1, x^2 в теореме 5.4.1 определены на окрестности $\mathbb{U}_x$ с точностью до конформных преобразований комплексных координат, которые вводятся на $\mathbb{U}_x$: $z = x^1 + ix^2$, что оправдывает название "конформно евклидова" для метрики. □

Определение. Если в пространстве-времени метрика имеет вид (5.29), где символ Кронекера $\delta_{\alpha\beta}$ заменен на лоренцеву метрику $\eta_{\alpha\beta}$, то пространство-время называется *конформно минковским.* □

В общем случае, для того, чтобы задать метрику в некоторой окрестности, необходимо задать $n(n+1)/2$ функций от n переменных. Во многих случаях, например, в общей теории относительности, метрика ищется как решение некоторой системы уравнений, которая ковариантна относительно выбора системы координат. То есть система уравнений движения выглядит по-разному в различных системах отсчета, однако пространства решений находятся во взаимно однозначном соответствии. При этом соответствие устанавливается преобразованием координат. При решении ковариантной системы уравнений удобно выбрать ту или иную систему отсчета с тем, чтобы уменьшить число неизвестных функций. Поскольку переход между системами координат характеризуется n функциями от n переменных, то максимальное число функций, которое можно зафиксировать в метрике общего вида, равно n. В результате метрика будет определяться $n(n-1)/2$ функциями. В этом случае говорят, что выбрана система координат или *зафиксирована калибровка* для метрики[2]. Как правило, координаты выбираются таким образом, чтобы система уравнений имела наиболее простой вид.

Выбор системы координат не означает, что n функций, входящих в метрику, можно зафиксировать произвольным образом. Для того, чтобы определить, существует ли система отсчета, где метрика имеет заданный вид, необходимо проанализировать систему уравнений для функций перехода к такой системе координат. Если эта система уравнений имеет решение, то такая калибровка называется *допустимой.*

[2]Это название пришло из электродинамики, инвариантной относительно локальных преобразований, которые называются калибровочными (см. раздел 23.3).

Часто система уравнений на функции перехода, имеет только локальные решения. Это означает, что соответствующая система координат может быть выбрана только локально.

5.5 Разложение Ходжа

Если на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, задана положительно определенная риманова метрика, то можно доказать ряд важных результатов для дифференциальных форм, которые не имеют места в общем случае. Для формулировки этих результатов введем несколько понятий.

Определение. Форма $A \in \Lambda_r(\mathbb{M})$ является *коточной*, если ее можно представить в виде

$$A = \delta B, \quad \text{где} \quad B \in \Lambda_{r+1}(\mathbb{M}).$$

Форма $A \in \Lambda_r(\mathbb{M})$ называется *козамкнутой*, если

$$\delta A = 0.$$

Форма $A \in \Lambda_r(\mathbb{M})$ называется *гармонической*, если

$$\triangle_{\text{г}} A = 0.$$

□

Любая функция $f \in \mathcal{C}^1(\mathbb{M})$, очевидно, является козамкнутой, $\delta f = 0$. Любая коточная форма является козамкнутой в силу свойства (4.58).

Теорема 5.5.1. *Любая гармоническая форма $A \in \Lambda_r(\mathbb{M})$ на компактном ориентируемом римановом многообразии является замкнутой и козамкнутой.*

Доказательство. Из определения оператора Лапласа–Бельтрами (4.59) для скалярного произведения (4.54) следует равенство

$$\langle \triangle_{\text{г}} A, A\rangle = -\langle \delta A, \delta A\rangle - \langle dA, dA\rangle. \tag{5.30}$$

Отсюда следует, что условие гармоничности $\triangle_{\text{г}} A = 0$ влечет за собой равенства $\delta A = 0$ и $dA = 0$. При этом ориентируемость и компактность многообразия достаточны для существования интеграла, входящего в скалярное произведение. Положительная определенность метрики необходима для того, чтобы оба слагаемых в правой части равенства (5.30) обращались в нуль по отдельности, и из равенства $\langle \delta A, \delta A\rangle = 0$ вытекало $\delta A = 0$ (аналогично для dA). □

Следствие. Любая гармоническая функция на компактном ориентируемом римановом многообразии является замкнутой $df = 0$ и, следовательно, равна константе. □

Замечание. Сформулированная теорема справедлива также для компактных неориентируемых многообразий. Действительно, она верна для двулистного ориентируемого накрытия $\widetilde{\mathbb{M}} \to \mathbb{M}$, которое всегда существует и тоже компактно, и при проекции замкнутость и козамкнутость форм сохраняется. □

Теорема 5.5.2 (Разложение Ходжа). *Пусть $\mathbb{M}$ – компактное ориентируемое риманово многообразие. Тогда любую r-форму $A \in \Lambda_r(\mathbb{M})$ можно представить в виде суммы точной, коточной и гармонической формы*

$$A = dB + \delta C + H, \qquad B \in \Lambda_{r-1}(\mathbb{M}), \quad C \in \Lambda_{r+1}(\mathbb{M}), \quad H \in \Lambda_r(\mathbb{M}). \tag{5.31}$$

Это разложение единственно.

Доказательство. См., например, [**?**]. □

Единственность разложения вытекает из ортогональности трех подпространств относительно скалярного произведения форм:

$$\langle A, A\rangle = \langle dB, dB\rangle + \langle \delta C, \delta C\rangle + \langle H, H\rangle.$$

Следствие. Уравнение

$$\triangle_{\text{Г}} A = D, \qquad A, D \in \Lambda_r(\mathbb{M}),$$

на компактном ориентируемом многообразии имеет решение тогда и только тогда, когда r-форма D ортогональна пространству гармонических r-форм. □

Теорема 5.5.3. *На компактном ориентируемом римановом пространстве существует только конечное число линейно независимых гармонических форм.*

Доказательство. См., например, [**?**]. □

Определение. Размерность пространства гармонических r-форм

$$\mathcal{H}_r(\mathbb{M}) = \{H \in \Lambda_r(\mathbb{M}) : \quad \triangle_{\text{Г}} H = 0\}$$

называется r-тым *числом Бетти*

$$b_r(\mathbb{M}) = \mathsf{dim}\, \mathcal{H}_r(\mathbb{M}). \quad \square \tag{5.32}$$

Отметим несколько свойств чисел Бетти. Поскольку оператор Лапласа–Бельтрами коммутирует с оператором Ходжа (4.60), а оператор Ходжа устанавливает изоморфизм $\mathcal{H}_r(\mathbb{M})$ и $\mathcal{H}_{n-r}(\mathbb{M})$, то $b_r(\mathbb{M}) = b_{n-r}(\mathbb{M})$. Так как все гармонические 0-формы на компактном ориентируемом римановом многообразии являются константами, то $b_0(\mathbb{M}) = b_n(\mathbb{M}) = 1$. Отметим, что пространство $\mathcal{H}_n(\mathbb{M})$ состоит из n-форм, пропорциональных форме объема υ (4.72). Если многообразие $\mathbb{M}$ связно, то $b_n(\mathbb{M}) = 1$ или $b_n(\mathbb{M}) = 0$ соответственно для ориентируемых и неориентируемых многообразий.

Определение. Число

$$\chi(\mathbb{M}) = \sum_{r=0}^{n} (-1)^r b_r(\mathbb{M}) \tag{5.33}$$

называется *эйлеровой характеристикой* многообразия $\mathbb{M}$. □

Теорема 5.5.4. *Пусть $\mathbb{M}$ – компактное ориентируемое риманово многообразие и A – дифференцируемая 1-форма. Тогда*

$$\int \upsilon \delta A = 0.$$

В частности, для произвольной функции $f \in \mathcal{C}^2(\mathbb{M})$

$$\int \upsilon \triangle_{\text{Г}} f = 0.$$

Доказательство. Справедливо равенство

$$\int_{\mathbb{M}} \upsilon \delta A = \langle \delta A, 1 \rangle = \langle A, d1 \rangle = 0. \qquad \square$$

Теорема 5.5.5. *Пусть $\mathbb{M}$ – компактное риманово многообразие с положительной гауссовой кривизной $K > 0$. Тогда первое число Бетти равно нулю $b_1(\mathbb{M}) = 0$.*

Доказательство. Напомним, что $K = -\frac{1}{2}\widetilde{R}$, где $\widetilde{R}$ – скалярная кривизна, определенная в разделе 7.8. Доказательство приведено, например, в [**?**]. $\square$

Предложение 5.5.1. *Любая ковариантно постоянная форма является замкнутой, козамкнутой и гармонической.*

Доказательство. Это следует из выражений для внешней производной (4.35), оператора кограницы (4.56) и определения оператора Лапласа–Бельтрами (4.59). $\square$

Пример 5.5.1. Форма объема υ (4.72) замкнута, козамкнута и гармонична, поскольку $\widetilde{\nabla}_\alpha \sqrt{|g|} = 0$. $\square$

Глава 6

Связность на векторном расслоении и расслоении реперов

Большинство моделей математической физики сводится к решению некоторых дифференциальных уравнений, которыми являются уравнения движения или уравнения равновесия. Мы предполагаем, что дифференциальные уравнения являются ковариантными объектами, т.е. при преобразовании координат преобразуются по тензорным правилам. В этом случае физические следствия не будут зависеть от выбора системы отсчета. Чтобы строить такие модели используется понятие связности и соответствующей ковариантной производной, т.к. что обычные частные производные от тензорных полей на многообразии не приводят к тензорным объектам. Это относится не только к уравнениям общей теории относительности, но и к другим моделям математической физики. По сути дела ковариантные производные появляются и в плоском пространстве-времени Минковского при переходе к криволинейным системам координат.

6.1 Векторные расслоения

В разделе (3.4) были определены расслоения общего вида, когда типичным слоем является произвольное многообразие. Ниже мы рассмотрим частный случай расслоений, типичными слоями которых являются векторные пространства, и введем понятие связности. Рассмотренные ранее касательное $\mathbb{T}(\mathbb{M})$, кокасательное $\mathbb{T}^*(\mathbb{M})$ и тензорные $\mathbb{T}^r_s(\mathbb{M})$ расслоения с базой $\mathbb{M}$ представляет собой частные примеры векторных расслоений, а понятие связности приводит к оператору ковариантного дифференцирования произвольных тензорных полей – исключительно важному инструменту для построения моделей математической физики.

Определение. Дифференцируемое многообразие $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ называется *векторным расслоением* с *базой* $\mathbb{M}$, *проекцией* π и *типичным слоем* $\mathbb{V}$, где $\mathbb{M}$ – дифференцируемое многообразие и $\mathbb{V} \simeq \mathbb{R}^n$ – векторное пространство. Проекция π является дифференцируемым сюръективным отображением $\mathbb{E} \xrightarrow{\pi} \mathbb{M}$. При этом требуется, чтобы для любого атласа $\mathbb{M} = \bigcup_i \mathbb{U}_i$ существовали отображения $\{\chi_i\}$, удовлетворяющие следующим условиям:

1) Отображение χ_i есть диффеоморфизм $\chi_i : \ \pi^{-1}(\mathbb{U}_i) \to \mathbb{U}_i \times \mathbb{V}$ такой, что $\pi \circ \chi_i^{-1}(x, v) = x$ для всех точек $x \in \mathbb{U}_i$ и $v \in \mathbb{V}$.

2) Сужение отображения χ_i на каждый слой $\chi_{i,x} : \pi^{-1}(x) \to \mathbb{V}$ есть гомоморфизм векторных пространств для всех $x \in \mathbb{U}_i$. При этом, если $x \in \mathbb{U}_i \cap \mathbb{U}_j$, то отображение
$$f_{ij}(x) = \chi_{j,x}\chi_{i,x}^{-1} : \mathbb{V} \to \mathbb{V},$$
является автоморфизмом, т.е. $f_{ij}(x) \in \mathbb{GL}(n, \mathbb{R})$.
3) Если $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$, то отображение $f_{ij}(x) : \mathbb{U}_i \cap \mathbb{U}_j \to \mathbb{GL}(n, \mathbb{R})$ является достаточно гладким.

Набор отображений $\{f_{ij}\}$ называется *функциями перехода.* □

Пусть $p \in \pi^{-1}(x)$ – некоторая точка векторного расслоения $\mathbb{E}$ из слоя над $x \in \mathbb{U}_i \cap \mathbb{U}_j$. Тогда $\chi_i(p) = (x, v_i)$ и $\chi_j(p) = (x, v_j)$. Следовательно,
$$p = \chi_i^{-1}(x, v_i) = \chi_j^{-1}(x, v_j). \tag{6.1}$$

Зафиксируем базис $\hat{e}_a$, $a = 1, \ldots, n$ в векторном пространстве $\mathbb{V}$ и разложим по нему векторы: $v_i = v_i^a \hat{e}_a$ и $v_j = v_j^a \hat{e}_a$. Тогда из условия 2) следует равенство
$$v_j^a = v_i^b f_{ij\,b}{}^a,$$
где введена невырожденная матрица $f_{ij}(x) = \{f_{ij\,b}{}^a(x)\} \in \mathbb{GL}(n, \mathbb{R})$. Эта матрица дифференцируемо зависит от точки x в силу условия 3 из определения расслоения.

Замечание. В данном определении оба свойства из определения расслоения общего вида в разделе 3.4 объединены в одно свойство 1). Свойство 2) предполагает, что слой векторного расслоения $\pi^{-1}(x)$ в каждой точке $x \in \mathbb{M}$, сам является векторным пространством, *гомоморфным* типичному слою, а в областях пересечения карт допускаются автоморфизмы, то есть линейные отображения $\mathbb{V}$ на себя. С учетом свойства 1) получаем, что каждый слой изоморфен векторному пространству, $\pi^{-1}(x) \simeq \mathbb{V}$. В условии 2) гомоморфизм можно заменить на гомеоморфизм и не требовать изначально наличия структуры векторного пространства в слое $\pi^{-1}(x)$. Гомеоморфизма достаточно для переноса структуры векторного пространства из $\mathbb{V}$ на слой $\pi^{-1}(x)$, при этом каждый слой будет, конечно, изоморфен $\mathbb{V}$. Свойство 3) говорит о том, что автоморфизмы задаются матрицей $\mathbb{GL}(n, \mathbb{R})$, элементы которой гладко зависят от координат. Поэтому в определении векторного расслоения был выбран атлас вместо окрестности точки $x \in \mathbb{M}$ в определении расслоения общего вида. □

Если в векторном пространстве $\mathbb{V}$, $\mathsf{dim}\,\mathbb{V} = n$, зафиксирован базис $\{\hat{e}_a\}$, $a = 1, \ldots, n$, то его можно отождествить с евклидовым пространством $\mathbb{R}^n$. После этого топология и дифференцируемая структура евклидова пространства с помощью отображений $\{\chi_i\}$ естественным образом переносятся на пространство расслоения $\mathbb{E}$, превращая его в многообразие. Фактически, требование дифференцируемости отображений χ_i означает, что дифференцируемая структура на пространстве расслоения $\mathbb{E}$ согласована с дифференцируемой структурой, индуцированной отображениями χ_i.

В общем случае размерности базы $\mathbb{M}$ и векторного пространства $\mathbb{V}$ (типичного слоя) могут не совпадать. Если $\mathsf{dim}\,\mathbb{M} = m$ и $\mathsf{dim}\,\mathbb{V} = n$, то $\mathsf{dim}\,\mathbb{E} = m + n$.

В силу условия 1) в определении расслоения многообразие $\pi^{-1}(\mathbb{U}_i)$ является тривиальным расслоением над каждой областью $\mathbb{U}_i$.

Пример 6.1.1. Построим двумерные расслоения $\mathbb{E}(\mathbb{S}^1, \pi, \mathbb{R})$, базой которых является окружность $\mathbb{S}^1$, а типичным слоем – вещественная прямая $\mathbb{R}$, которая рассматривается, как векторное пространство. Отождествим каждую точку окружности с

полярным углом $\mathbb{S}^1 \supset \alpha \in [0, 2\pi)$. Покроем окружность двумя картами $\mathbb{S}^1 \supset \mathbb{U}_1 = (2\pi - \delta, 2\pi) \cup [0, \pi + \delta)$ и $\mathbb{S}^1 \supset \mathbb{U}_2 = (\pi - \delta, 2\pi) \cup [0, \delta)$, где $0 < \delta < \pi/2$, которые пересекаются по двум областям $\mathbb{U}_1 \cap \mathbb{U}_2 = \mathbb{U}_{12} \cup \mathbb{U}_{21}$, где $\mathbb{U}_{12} = (\pi - \delta, \pi + \delta)$ и $\mathbb{U}_{21} = (2\pi - \delta, 2\pi) \cup [0, \delta)$, при этом координатные функции определяют диффеоморфизмы:

$$\varphi_1(\mathbb{U}_1) = (-\delta, \pi + \delta) \in \mathbb{R},$$
$$\varphi_2(\mathbb{U}_2) = (\pi - \delta, 2\pi + \delta) \in \mathbb{R}.$$

Определим векторное расслоение $\mathbb{E}$ его локальным вложением в $\mathbb{R}^3$:

$$\pi^{-1}(\mathbb{U}_1) = \{p = (x, v^1, v^2) \in \mathbb{R}^3 : \quad x \in \varphi_1(\mathbb{U}_1),\ v^2 \cos(kx/2) = v^1 \sin(kx/2)\},$$
$$\pi^{-1}(\mathbb{U}_2) = \{p = (x, v^1, v^2) \in \mathbb{R}^3 : \quad x \in \varphi_2(\mathbb{U}_2),\ v^2 \cos(kx/2) = v^1 \sin(kx/2)\},$$

где $k = 0, 1, 2, \ldots$ и v^1, v^2 – декартовы координаты на плоскости $\mathbb{R}^2$. Для большей ясности мы обозначили точки окружности $\alpha \in \mathbb{S}^1$ и их координаты $x = \varphi(\alpha) \in \mathbb{R}$ разными буквами. Словами, при движении точки по окружности $\mathbb{S}^1$ вещественная прямая $\mathbb{R} \subset \mathbb{R}^2$ поворачивается в плоскости $\mathbb{R}^2$ на угол $kx/2$. Если $v \in \mathbb{R}$ – точка на прямой, то ей соответствует точка $\boldsymbol{v} = (v^1, v^2) = \Big(v \cos(kx/2), v \sin(kx/2)\Big) \in \mathbb{R}^2$ на плоскости. Направляющим вектором прямой является $\boldsymbol{n} = \Big(\cos(kx/2), \sin(kx/2)\Big) \in \mathbb{R}^2$. Теперь определим отображения:

$$\chi_1 : \quad \pi^{-1}(\mathbb{U}_1) \ni p \mapsto (\alpha, v) \in \mathbb{S}^1 \times \mathbb{R},$$
$$\chi_2 : \quad \pi^{-1}(\mathbb{U}_2) \ni p \mapsto (\alpha, v) \in \mathbb{S}^1 \times \mathbb{R},$$

где $v = (\boldsymbol{n}, \boldsymbol{v}) = v^1 \cos(kx/2) + v^2 \sin(kx/2)$. Для определения гомоморфизмов χ_{ij} необходимо учесть, что точки x и $x + 2\pi$ в пересечении $\mathbb{U}_{21}$ соответствуют одной и той же точке окружности, и

$$\cos\Big(k(2\pi + x)/2\Big) = (-1)^k \cos(kx/2), \qquad \sin\Big(k(2\pi + x)/2\Big) = (-1)^k \sin(kx/2).$$

Поэтому $f_{12} = \mathsf{id}(\mathbb{R})$ и $f_{21} = (-1)^k \mathsf{id}(\mathbb{R})$. Отсюда следует, что все расслоения для четных k диффеоморфны между собой и диффеоморфны тривиальному расслоению $\mathbb{E} = \mathbb{S}^1 \times \mathbb{R}$, которое получается при $k = 0$. Расслоения при нечетных k также диффеоморфны между собой и диффеоморфны листу Мебиуса, соответствующего $k = 1$. □

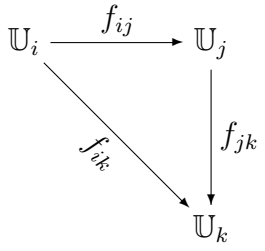
Последний пример показывает, что отображения χ_i однозначно определяют функции перехода. В обратную сторону. Если задано два векторных расслоения с одинаковой базой и функциями перехода, то они диффеоморфны между собой.

При задании расслоения функции перехода должны удовлетворять условиям совместности:

1) для всех $x \in \mathbb{U}_i$, $f_{ii} = \mathsf{id}$;

2) для всех $x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k \neq \emptyset$, $\quad f_{ij} f_{jk} f_{ki} = \mathsf{id}$.

Эти условия совместности являются, очевидно, необходимыми. Второе свойство означает коммутативность диаграммы



где $f_{ik} = f_{ki}^{-1}$. При задании расслоения с помощью функций перехода необходимо требовать выполнения данных свойств.

Теорема 6.1.1. *Пусть $\{\mathbb{U}_i\}$ – открытое координатное покрытие (атлас) дифференцируемого многообразия $\mathbb{M}$, и $\mathbb{V}$ – векторное пространство. Если для всех пар пересекающихся координатных областей $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$ заданы гладкие отображения $f_{ij} : \mathbb{U}_i \cap \mathbb{U}_j \to \mathbb{GL}(\mathbb{V})$, которые удовлетворяет условиям совместности 1) и 2), тогда существует векторное расслоение $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ с функциями перехода f_{ij}.*

Доказательство. См., например, [**?**]. □

Поскольку слоями векторных расслоений являются векторные пространства, то из заданных векторных расслоений с одной и той же базой можно строить новые векторные расслоения, производя со слоями операции, которые допустимы для векторных пространств, поточечно.

Пример 6.1.2 (Прямая сумма $\mathbb{E}_1 \oplus \mathbb{E}_2$ и тензорное произведение $\mathbb{E}_1 \otimes \mathbb{E}_2$ расслоений.). Пусть $\mathbb{E}_1(\mathbb{M}, \pi_1, \mathbb{V}_1)$ и $\mathbb{E}_2(\mathbb{M}, \pi_2, \mathbb{V}_2)$ – два векторных расслоения с одной и той же базой $\mathbb{M}$. Определим прямую сумму векторных расслоений

$$\mathbb{E}_1 \oplus \mathbb{E}_2(\mathbb{M}, \pi, \mathbb{V}_1 \oplus \mathbb{V}_2) := \bigcup_{x \in \mathbb{M}} \pi_1^{-1}(x) \oplus \pi_2^{-1}(x). \tag{6.2}$$

Если точка $p \in \mathbb{E}_1 \oplus \mathbb{E}_2$, то она имеет вид $p = p_1 \oplus p_2$, где $p_1 \in \pi_1^{-1}(x)$, $p_2 \in \pi_2^{-1}(x)$ для некоторого $x \in \mathbb{M}$. Тем самым определена проекция $\pi(p) = x$. Типичным слоем прямой суммы векторных расслоений является векторное пространство $\mathbb{V}_1 \oplus \mathbb{V}_2$. Пусть $\{\mathbb{U}_i\}$ – координатное покрытие $\mathbb{M}$ и $\{f_{ij}^{(1)}\}$ и $\{f_{ij}^{(2)}\}$ – функции перехода для расслоений $\mathbb{E}_1$ и $\mathbb{E}_2$, соответственно. Функции перехода для прямой суммы векторных расслоений имеют вид

$$h_{ij} := \begin{pmatrix} f_{ij}^{(1)} & 0 \\ 0 & f_{ij}^{(2)} \end{pmatrix},$$

которые задают автоморфизм в $\mathbb{V}_1 \oplus \mathbb{V}_2$.

Аналогично определяется тензорное произведение векторных расслоений:

$$\mathbb{E}_1 \otimes \mathbb{E}_2(\mathbb{M}, \pi, \mathbb{V}_1 \oplus \mathbb{V}_2) := \bigcup_{x \in \mathbb{M}} \pi_1^{-1}(x) \otimes \pi_2^{-1}(x). \tag{6.3}$$

Функции перехода для тензорного произведения имеют вид

$$h_{ij} := f_{ij}^{(1)} \otimes f_{ij}^{(2)},$$

которые действуют в тензорном произведении $\mathbb{V}_1 \otimes \mathbb{V}_2$. □

Размерности прямой суммы и тензорного произведения расслоений равны:

$$\begin{aligned} \mathsf{dim}\,(\mathbb{E}_1 \oplus \mathbb{E}_2) &= \mathsf{dim}\,\mathbb{M} + \mathsf{dim}\,\mathbb{V}_1 + \mathsf{dim}\,\mathbb{V}_2, \\ \mathsf{dim}\,(\mathbb{E}_1 \otimes \mathbb{E}_2) &= \mathsf{dim}\,\mathbb{M} + \mathsf{dim}\,\mathbb{V}_1 \times \mathsf{dim}\,\mathbb{V}_2. \end{aligned}$$

В заключение данного раздела обсудим касательное расслоение, построенное в разделе 3.6.2.

Теорема 6.1.2. *Касательное расслоение $\mathbb{T}(\mathbb{M})$ является векторным расслоением $\mathbb{E}(\mathbb{M}, \pi, \mathbb{R}^n)$ с базой $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, типичным слоем $\mathbb{R}^n$ и проекцией $\pi : (x, X) \to x$, где $X \in \mathbb{T}_x(\mathbb{M})$.*

Доказательство. Проверка свойств 1)–3) в определении векторного расслоения. □

Замечание. Для касательного расслоения размерности базы и типичного слоя совпадают:

$$\dim \mathbb{M} = \dim \mathbb{R}^n = n.$$ □

Касательное расслоение имеет свою специфику. Вообще говоря, в определении векторного расслоения ничего не говорится о том, какое именно линейное преобразование векторного пространства происходит в областях пересечения карт. Говорится лишь о том, что оно возможно. Для касательного расслоения реализован естественный способ гладко сопоставить каждому преобразованию координат автоморфизм касательного пространства. А именно, каждому преобразованию координат, соответствует тождественное преобразование касательного пространства, $f_{ij} = \mathrm{id}\left(\mathbb{T}_x(\mathbb{M})\right)$ для всех $x \in \mathbb{U}_i \cap \mathbb{U}_j$. Если в касательном пространстве в точке $x \in \mathbb{M}$ (слое над x) выбран координатный (голономный) базис $\{\partial_\alpha\}$, то при преобразовании координат в базе компоненты вектора умножаются на матрицу Якоби, а голономный базис касательного пространства – на обратную матрицу Якоби (3.9). Поэтому сам вектор остается без изменения. Если вектор касательного пространства рассматривается, как набор компонент $\{X^\alpha\}$, считая, что базис векторного пространства фиксирован, то при преобразовании координат компоненты вектора умножаются на матрицу Якоби $\partial_\alpha x^{\alpha'}$ справа. Тогда функциями перехода являются матрицы Якоби. Обе точки зрения на функции перехода допустимы. Это зависит от того, как рассматривать базис векторного пространства $\mathbb{V}$. В общем случае базис касательного пространства e_a, $a = 1, \ldots, n$, можно зафиксировать произвольным образом (неголономный базис). Тогда мы считаем, что при преобразовании координат базис и координаты векторов относительно этого базиса: $X = X^a e_a$ не меняются. Такой подход часто бывает удобнее и соответствует использованию репера $e_a = e^\alpha{}_a \partial_\alpha$, который образует базис касательного пространства.

Дифференцируемое сечение векторного расслоения $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$, которое всюду отлично от нуля, существует не всегда. Существование таких сечений отражает определенные топологические свойства базы $\mathbb{M}$.

Понятие (псевдо) риманова многообразия естественным образом обобщается на произвольные векторные расслоения.

Определение. Если в каждой точке $x \in \mathbb{M}$ задана невырожденная симметричная билинейная форма g на слое $\pi^{-1}(x)$, и значение $g(V_1, V_2)$ является достаточно гладкой функцией на $\mathbb{M}$ для произвольных достаточно гладких сечений $V_1, V_2 \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$, то $\mathbb{E}$ называется *(псевдо) римановым векторным расслоением.* □

Пример 6.1.3. Тензорное расслоение на римановом многообразии является римановым векторным расслоением. Например, для тензорных полей $X = \{X_\alpha{}^\beta\}$ и $\{Y_\alpha{}^\beta\}$ типа $(1,1)$ квадратичная форма задается римановой метрикой

$$g(X, Y) = g^{\alpha\beta} X_\alpha{}^\gamma Y_\beta{}^\delta g_{\gamma\delta}.$$ □

6.2 Связность на векторном расслоении

В настоящем разделе мы определим связность на векторном расслоении в инвариантной форме. Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = m$, и векторное пространство $\mathbb{V}$, $\dim \mathbb{V} = n$. Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ – векторное расслоение. Сечениями этого расслоения

являются векторные поля $V(x)$. Обозначим множество гладких сечений расслоения $\mathbb{E}$ через $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$. На множестве сечений векторного расслоения поточечно вводится сложение и умножение на гладкие функции. Таким образом, множество сечений $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ является $\mathcal{C}^\infty(\mathbb{M})$-модулем.

Все дальнейшие конструкции мы будем иллюстрировать записью соответствующих выражений в компонентах, поскольку это, во-первых, наглядно и, во-вторых, необходимо при проведении вычислений. Выберем в каждом слое $\pi^{-1}(x)$, $x \in \mathbb{U} \subset \mathbb{M}$, базис $e_a(x) \in \mathcal{V}_{\mathbb{E}}(\mathbb{U})$, $a = 1, \ldots, n$. Этот базис называется *реперо*м. Для каждой точки x существует окрестность $\mathbb{U} \subset \mathbb{M}$, в которой репер задается гладкими функциями от x. Вообще говоря, такой репер глобально существует не для всех многообразий, и это зависит от топологических свойств. В компонентах сечение векторного расслоения имеет вид $V = V^a(x)e_a$.

Кокасательное расслоение $\mathbb{T}^*(\mathbb{M})$ имеет ту же базу, что и $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$, поэтому можно построить тензорное произведение расслоений $\mathbb{T}^*(\mathbb{M}) \otimes \mathbb{E}$. Пусть $\mathcal{V}_{\mathbb{T}^*\otimes\mathbb{E}}(\mathbb{M})$ – множество сечений тензорного произведения расслоений. В компонентах сечение этого расслоения задается векторным полем с дополнительным ковариантным индексом $dx^\alpha V_\alpha{}^a(x)e_a$, где dx^α, $\alpha = 1, \ldots, m$, – координатный базис 1-форм.

Определение. *Связностью* ∇ *на векторном расслоении* $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ называется отображение

$$\nabla : \quad \mathcal{V}_{\mathbb{E}}(\mathbb{M}) \to \mathcal{V}_{\mathbb{T}^*\otimes\mathbb{E}}(\mathbb{M}), \tag{6.4}$$

которое удовлетворяет двум условиям:

1) для любых двух сечений $V_1, V_2 \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$

$$\nabla(V_1 + V_2) = \nabla V_1 + \nabla V_2;$$

2) для произвольного сечения $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ и произвольной функции $f \in \mathcal{C}^\infty(\mathbb{M})$

$$\nabla(fV) = df \otimes V + f\nabla V. \quad \square \tag{6.5}$$

Если в определении связности положить $f = \mathsf{const}$, то из условий 1) и 2) следует линейность отображения ∇.

Определение. Сечение $\nabla V \in \Lambda_1(\mathbb{M}) \otimes \mathbb{V}$ представляет собой 1-форму на $\mathbb{M}$ со значениями в векторном пространстве $\mathbb{V}$. Рассмотрим касательное векторное поле к базе $X \in \mathcal{X}(\mathbb{M})$. Тогда определено значение значение 1-формы ∇V на векторном поле X

$$\nabla_X V = \nabla V(X), \tag{6.6}$$

которое называется *ковариантной производной* векторного поля $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ вдоль касательного векторного поля $X \in \mathcal{X}(\mathbb{M})$. $\square$

Ковариантная производная вдоль векторного поля ∇_X является отображением $\mathcal{V}_{\mathbb{E}}(\mathbb{M}) \to \mathcal{V}_{\mathbb{E}}(\mathbb{M})$. Пусть $X, Y \in \mathcal{X}(\mathbb{M})$ – два произвольных векторных поля на $\mathbb{M}$, и $f \in \mathcal{C}^\infty(\mathbb{M})$ – гладкая функция. Тогда ковариантная производная имеет следующие свойства:

$$\begin{aligned}
\nabla_{X+Y}V &= \nabla_X V + \nabla_Y V,\\
\nabla_{fX}V &= f\nabla_X V,\\
\nabla_X(V_1 + V_2) &= \nabla_X V_1 + \nabla_X V_2,\\
\nabla_X(fV) &= (Xf)V + f\nabla_X V.
\end{aligned}$$

Построим выражение для ковариантной производной в компонентах. Запишем отображение (6.4) для векторов репера:

$$\nabla e_a = dx^\alpha \otimes \omega_{\alpha a}{}^b e_b, \tag{6.7}$$

где $\omega_{\alpha a}{}^b(x)$ – некоторые функции на координатной окрестности $\mathbb{U}$. Введем 1-формы

$$\omega_a{}^b = dx^\alpha \omega_{\alpha a}{}^b. \tag{6.8}$$

Тогда соотношение (6.7) можно переписать в виде

$$\nabla e_a = \omega_a{}^b \otimes e_b. \tag{6.9}$$

Рассмотрим, как меняются компоненты $\omega_{\alpha a}{}^b$ при преобразовании репера

$$e'_a = S_a{}^b e_b, \tag{6.10}$$

где $S_a{}^b(x) \in \mathbb{GL}(n,\mathbb{R})$ – некоторая матрица преобразования, элементы которой могут зависеть от точки $x \in \mathbb{U}$. В новом базисе имеем

$$\nabla e'_a = \omega'^b_a \otimes e'_b.$$

С другой стороны, из определения связности (6.5) следует

$$\nabla(S_a{}^b e_b) = dS_a{}^b \otimes e_b + S_a{}^b \nabla e_b.$$

Сравнивая эти выражения, получаем правило преобразования:

$$\omega'^b_a = S_a{}^c \omega_c{}^d S^{-1}{}^b_d + dS_a{}^c S^{-1}{}^b_c, \tag{6.11}$$

или, в комопонентах,

$$\omega'_{\alpha a}{}^b = S_a{}^c \omega_{\alpha c}{}^d S^{-1}{}^b_d + \partial_\alpha S_a{}^c S^{-1}{}^b_c. \tag{6.12}$$

Замечание. Это – одна из самых важных формул дифференциальной геометрии, которая широко используется в калибровочных моделях математической физики. Подчеркнем, что преобразование 1-форм $\omega_a{}^b$ содержит неоднородное слагаемое и поэтому не является тензорным. Следовательно, компоненты $\omega_{\alpha a}{}^b$ не определяют никакого тензора. Они определяют связность. □

Теорема 6.2.1. *На любом векторном расслоении* $\mathbb{E}(\mathbb{M},\pi,\mathbb{V})$ *существует связность.*

Доказательство. Доказательство проводится путем явного построения связности с использованием разложения единицы [**?**]. □

Пусть на многообразии $\mathbb{M}$ задан атлас $\{\mathbb{U}_i\}$. Тогда связность в каждой карте задается компонентами $\omega_{\alpha a}{}^b$, которые в областях пересечения карт связаны преобразованием (6.12). Верно также и обратное утверждение. Набор компонент $\omega_{\alpha a}{}^b$, связанных преобразованием (6.12) в областях пересечения карт, однозначно определяет связность на векторном расслоении.

Определение. Компоненты $dx^\alpha \omega_{\alpha a}{}^b$ называются *локальной формой связности* или, просто, связностью.

Определив действие связности на репер, можно выписать явное выражение для действия связности на произвольное векторное поле $V^a e_a \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$:

$$\nabla V = (dV^a + V^b \omega_b{}^a) \otimes e_a = dx^\alpha (\partial_\alpha V^a + V^b \omega_{\alpha b}{}^a) \otimes e_a, \tag{6.13}$$

где мы воспользовались определением (6.5). Это выражение называется *ковариантной производной* векторного поля. Выражение в скобках,

$$\nabla_\alpha V^a = \partial_\alpha V^a + V^b \omega_{\alpha b}{}^a, \tag{6.14}$$

называется *ковариантной производной компонент* векторного поля. □

Замечательным свойством ковариантной производной является то, что при локальном изменении базиса (6.10) ковариантная производная меняется по тензорному закону. Действительно, при изменении базиса (репера) (6.10) компоненты вектора $V = V^a e_a = V'^a e'_a$ преобразуются по векторному представлению группы $\mathbb{GL}(n, \mathbb{R}^n)$:

$$V'^a = V^b S^{-1}{}_b{}^a, \qquad S_b{}^a(x) \in \mathbb{GL}(n, \mathbb{R}^n). \tag{6.15}$$

Используя преобразование связности (6.12), нетрудно проверить справедливость следующей формулы преобразования ковариантной производной

$$\nabla'_\alpha V'^a = \partial_\alpha V'^a + V'^b \omega'_{\alpha b}{}^a = (\nabla_\alpha V^b) S^{-1}{}_b{}^a,$$

где ковариантная производная со штрихом ∇'_α берется с локальной формой связности $\omega'_{\alpha b}{}^a$. Тем самым ковариантная производная компонент векторного поля преобразуется по векторному представлению группы $\mathbb{GL}(n, \mathbb{R}^n)$.

Ковариантная производная от векторного поля V вдоль касательного векторного поля X (6.6) в компонентах имеет вид

$$\nabla_X V^a = X^\alpha \nabla_\alpha V^a.$$

Рассмотрим некоторые свойства связности. Легко доказывается

Предложение 6.2.1. *Пусть на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ задана связность ∇ и зафиксирована точка базы $x \in \mathbb{M}$. Тогда в окрестности точки x существует такой репер, что компоненты локальной формы связности в этой точке обращаются в нуль $\omega_{\alpha a}{}^b(x) = 0$.*

Замечание. Подчеркнем, что в данном предложении речь идет о фиксированной точке многообразия, а не об окрестности. □

Продолжим изучение локальной формы связности. Перепишем соотношение (6.11) в виде

$$\omega' S = S\omega + dS, \tag{6.16}$$

где, для краткости, мы опустили векторные индексы, предполагая всюду суммирование "с десяти до четырех" (имеется ввиду расположение чисел на циферблате часов). Внешняя производная от этого равенства дает

$$d\omega' S - \omega' \wedge dS = dS \wedge \omega + S d\omega.$$

Подстановка выражения для dS из (6.16) приводит к равенству

$$d\omega' - \omega' \wedge \omega' = S(d\omega - \omega \wedge \omega) S^{-1}, \tag{6.17}$$

которое приводит к следующему важному понятию.

Определение. 2-форма на $\mathbb{M}$

$$R_a{}^b = d\omega_a{}^b - \omega_a{}^c \wedge \omega_c{}^b \tag{6.18}$$

называется *локальной формой кривизны* связности ∇. В компонентах:

$$R_a{}^b = \frac{1}{2} dx^\alpha \wedge dx^\beta R_{\alpha\beta a}{}^b,$$

где

$$R_{\alpha\beta a}{}^b = \partial_\alpha \omega_{\beta a}{}^b - \partial_\beta \omega_{\alpha a}{}^b - \omega_{\alpha a}{}^c \omega_{\beta a}{}^c + \omega_{\beta a}{}^c \omega_{\alpha c}{}^b. \quad \square \tag{6.19}$$

Локальная форма кривизны – это 2-форма на многообразии $\mathbb{M}$ со значениями в тензорном произведении $\mathbb{V}^* \otimes \mathbb{V}$, где $\mathbb{V}^*$ – векторное пространство, сопряженное к $\mathbb{V}$.

Формула (6.17) показывает, что при локальном преобразовании репера форма кривизны преобразуется по тензорному закону, т.е. однородно, несмотря на то, что преобразование формы связности (6.11) содержит неоднородное слагаемое.

Пусть на многообразии $\mathbb{M}$ задано два векторных поля $X, Y \in \mathcal{X}(\mathbb{M})$. Тогда значение формы кривизны на этих полях $R(X,Y)_a{}^b$ определяет линейное преобразование из $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ в $\mathcal{V}_{\mathbb{E}}(\mathbb{M})$. Если $V = V^a e_a \in \mathcal{V}_{\mathbb{E}}(\mathbb{E})$, то

$$R(X,Y): \quad V^a \to V^b R(X,Y)_b{}^a.$$

Это отображение обладает очевидными свойствами:

1) $R(X,Y) = -R(Y,X)$;
2) $R(fX,Y) = fR(X,Y)$;
3) $R(X,Y)(fV) = fR(X,Y)V$;

где $f \in \mathcal{C}^\infty(\mathbb{M})$.

Предложение 6.2.2. *Пусть на векторном расслоении $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$ задана связность ∇ с локальной формой кривизны R. Тогда для любых векторных полей $X, Y \in \mathcal{X}(\mathbb{M})$ справедлива формула*

$$R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}. \tag{6.20}$$

Доказательство. Прямая проверка в компонентах. $\square$

Теорема 6.2.2 (**Тождества Бианки**). *Форма кривизны R удовлетворяет тождествам Бианки:*

$$dR_a{}^b + R_a{}^c \wedge \omega_c{}^b - \omega_a{}^c \wedge R_c{}^b = 0. \tag{6.21}$$

Доказательство. Внешняя производная от определения кривизны (6.18) равна

$$\begin{aligned} dR &= -d\omega \wedge \omega + \omega \wedge d\omega \\ &= -(R + \omega \wedge \omega) \wedge \omega + \omega \wedge (R + \omega \wedge \omega) \\ &= -R \wedge \omega + \omega \wedge R. \end{aligned}$$ $\square$

В компонентах тождества Бианки имеют вид:

$$dx^\alpha \wedge dx^\beta \wedge dx^\gamma \partial_\alpha R_{\beta\gamma a}{}^b = dx^\alpha \wedge dx^\beta \wedge dx^\gamma (\omega_{\alpha a}{}^c R_{\beta\gamma c}{}^b - R_{\alpha\beta a}{}^c \omega_{\gamma c}{}^b)$$

или

$$\nabla_\alpha R_{\beta\gamma a}{}^b + \nabla_\beta R_{\gamma\alpha a}{}^b + \nabla_\gamma R_{\alpha\beta a}{}^b = 0,$$

где

$$\nabla_\alpha R_{\beta\gamma a}{}^b = \partial_\alpha R_{\beta\gamma a}{}^b + R_{\beta\gamma a}{}^c \omega_{\alpha c}{}^b - \omega_{\alpha a}{}^c R_{\beta\gamma c}{}^b.$$

– ковариантная производная от тензора (2-формы) кривизны.

Замечание. Построенная конструкция имеет важное приложение в калибровочных моделях теории поля. Если матрицы $S_a{}^b(x) \in \mathbb{GL}(n,\mathbb{R})$ в (6.15) образуют представление ρ некоторой полупростой группы Ли $\mathbb{G}$, то эта группа называется калибровочной. В этом случае локальная форма связности $\omega_{\alpha a}{}^b$ называется *полем Янга-Миллса*, а локальная форма кривизны $R_{\alpha\beta a}{}^b$ – *напряженностью* поля Янга-Миллса. Векторные поля $V = V^a e_a$ называются *полями материи*, преобразующимися по представлению ρ калибровочной группы. Эти поля называются скалярными полями, так как представляют собой набор функций на многообразии, не меняющихся при преобразовании координат $V^a \in \mathcal{C}^\infty(\mathbb{M})$ для всех значений индекса a. □

Определение. Рассмотрим сечение векторного расслоения $\mathbb{E}$, т.е. векторное поле $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$. Сечение V называется *параллельным или горизонтальным* относительно связности ∇, если ковариантная производная от него равна нулю

$$\nabla V = 0. \quad \square \tag{6.22}$$

Нулевое сечение, очевидно, всегда является параллельным. В то же время отличное от нуля параллельное сечение в общем случае может отсутствовать. Рассмотрим этот вопрос подробно. Разложим векторное поле по локальному базису: $V = V^a e_a$. Тогда уравнение (6.22) эквивалентно пфаффовой системе уравнений на компоненты:

$$dV^a + V^b \omega_b{}^a = 0. \tag{6.23}$$

Для анализа этой системы уравнений введем обозначение для левой части

$$A^a = dV^a + V^b \omega_b{}^a.$$

Внешняя производная от этой 1-формы равна

$$dA^a = A^b \wedge \omega_b{}^a + V^b R_b{}^a.$$

Из этого выражения и теоремы Фробениуса (3.11.4), записанной в терминах дифференциальных форм, следует, что, если кривизна связности равна нулю, то пфаффова система уравнений (6.22) является вполне интегрируемой. В этом случае существует n линейно независимых параллельных сечений векторного расслоения. Мы видим, что для существования отличных от нуля локальных решений системы уравнений (6.23) локальная форма кривизны связности должна обращаться в нуль. Другими словами, равенство нулю формы кривизны является условием разрешимости системы уравнений в частных производных (6.23).

Определение. Пусть на многообразии $\mathbb{M}$ задана дифференцируемая кривая $\gamma = \{x^\alpha(t)\}$, $\alpha = 1,\ldots,m$. Обозначим касательный вектор к кривой (вектор скорости) через v. Тогда сечение векторного расслоения $V \in \mathcal{V}_{\mathbb{E}}(\mathbb{M})$ называется *параллельным вдоль кривой* γ, если ковариантная производная от него вдоль v равна нулю

$$\nabla_v V = 0. \quad \square \tag{6.24}$$

Касательный вектор к кривой (вектор скорости) имеет вид $v = \dot{x}^\alpha \partial_\alpha$. В компонентах условие параллельности сечения вдоль кривой γ сводится к системе уравнений на компоненты сечения

$$\dot{V}^a + \dot{x}^\alpha V^b \omega_{\alpha b}{}^a = 0. \tag{6.25}$$

Поскольку это – система обыкновенных дифференциальных уравнений, то локально у нее существует единственное решение при заданных начальных данных. Таким образом, если задан вектор V_p в некоторой точке p на кривой γ, то он однозначно определяет векторное поле вдоль кривой γ, которое называется *параллельным переносом* вектора V_p вдоль кривой γ. Очевидно, что параллельный перенос задает изоморфизм слоев векторного расслоения $\mathbb{E}$ в различных точках кривой γ.

Связность ∇ на векторном расслоении $\mathbb{E}(\mathbb{M},\pi,\mathbb{V})$ индуцирует связность на сопряженном векторном расслоении $\mathbb{E}^*(\mathbb{M},\pi,\mathbb{V}^*)$, которую мы также обозначим ∇. Это делается следующим образом. Пусть заданы сечения $V\in\mathcal{V}_{\mathbb{E}}(\mathbb{M})$ и $V^*\in\mathcal{V}_{\mathbb{E}^*}(\mathbb{M})$. Тогда каждое сечение $V^*(x)$ задает поточечное линейное отображение

$$V^*(x):\quad \mathcal{V}_{\mathbb{E}}(\mathbb{M})\ni V(x)\ \mapsto\ (V^*,V)\in\mathcal{C}^\infty(\mathbb{M}).$$

Определим индуцированную связность на сопряженном расслоении $\mathbb{E}^*$ следующей формулой

$$d(V^*,V)=(V^*,\nabla V)+(\nabla V^*,V),\qquad \forall V\in\mathcal{V}_{\mathbb{E}}(\mathbb{M}). \tag{6.26}$$

Найдем выражение для индуцированной связности в компонентах. Пусть e^a – базис сопряженного векторного пространства $\mathbb{V}^*$: $(e^a,e_b)=\delta^a_b$. Этот базис называется *корепером*. Тогда подстановка корепера e^a и репера e_b в определение (6.26) дает

$$\nabla e^a=-e^b\otimes\omega_b{}^a. \tag{6.27}$$

Отсюда следует выражение для ковариантной производной сечения $V^*=e^aV_a$ сопряженного расслоения

$$\nabla V^*=e^a\otimes(dV_a-\omega_a{}^bV_b). \tag{6.28}$$

Эта ковариантная производная отличается знаком у слагаемого со связностью от соответствующего выражения для сечения векторного расслоения (6.13).

Определение. Пусть на векторных расслоениях $\mathbb{E}_1(\mathbb{M},\pi_1,\mathbb{V}_1)$ и $\mathbb{E}_2(\mathbb{M},\pi_2,\mathbb{V}_2)$ с одинаковой базой $\mathbb{M}$ независимо заданы связности ∇_1 и ∇_2. Пусть заданы сечения $V_1\in\mathcal{V}_{\mathbb{E}_1}(\mathbb{M})$ и $V_2\in\mathcal{V}_{\mathbb{E}_2}(\mathbb{M})$. Тогда на прямой сумме $\mathbb{E}_1\oplus\mathbb{E}_2$ и тензорном произведении $\mathbb{E}_1\otimes\mathbb{E}_2$ расслоений определена связность

$$\begin{aligned}\nabla(V_1\oplus V_2)&=\nabla_1V_1\oplus\nabla_2V_2,\\ \nabla(V_1\otimes V_2)&=\nabla_1V_1\otimes V_2+V_1\otimes\nabla_2V_2,\end{aligned} \tag{6.29}$$

которая называется *индуцированной*. □

6.3 Аффинная связность

Касательное расслоение $\mathbb{T}(\mathbb{M})$, $\dim\mathbb{M}=n$, является векторным расслоением, типичным слоем которого является векторное пространство $\mathbb{R}^n$. Поэтому определение связности на векторном расслоении без изменений переносится на касательные расслоения. Связность на касательном расслоении называется *аффинной связностью*. Если в касательном пространстве и сопряженном к нему пространстве 1-форм выбрать координатный базис ∂_α и dx^α, то формулы для компонент связности (6.7) и (6.27) определяют аффинную связность

$$\begin{aligned}\nabla(\partial_\alpha)&=\ \ dx^\beta\otimes\Gamma_{\beta\alpha}{}^\gamma\partial_\gamma,\\ \nabla(dx^\alpha)&=-dx^\gamma\otimes dx^\beta\Gamma_{\beta\gamma}{}^\alpha.\end{aligned} \tag{6.30}$$

У компонент аффинной связности $\Gamma_{\alpha\beta}{}^{\gamma}$ все индексы являются координатными, и для них принято обозначение Γ вместо ω. Формула преобразования компонент аффинной связности при преобразовании координат получается из формулы преобразования компонент связности на векторном расслоении (6.12) после замены матрицы $S_b{}^a(x)$ на матрицу Якоби $\partial_\alpha x^{\alpha'}$. При этом необходимо учесть, что при преобразовании координат преобразуется также базис 1-форм, что приводит к дополнительному общему множителю:

$$\Gamma_{\beta'\gamma'}{}^{\alpha'} = (\partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma \Gamma_{\beta\gamma}{}^{\alpha} + \partial^2_{\beta'\gamma'} x^\alpha)\partial_\alpha x^{\alpha'}. \tag{6.31}$$

Определение. Совокупность функций $\Gamma_{\alpha\beta}{}^{\gamma}$, преобразующихся по правилу (6.31) при преобразовании координат, называется *компонентами локальной формы аффинной связности* или, короче, аффинной связностью на $\mathbb{M}$. □

Замечание. Аффинная связность, так же как и связность на произвольном векторном расслоении, определенная в предыдущем разделе, не имеет ничего общего с метрикой. Ее можно задать независимо от метрики, которая может вообще отсутствовать на многообразии. □

Аффинная связность ∇ осуществляет отображение

$$\nabla: \quad \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \ni X, Y \ \mapsto \nabla_X Y \in \mathcal{X}(\mathbb{M}), \tag{6.32}$$

где $\nabla_X Y$ – ковариантная производная от векторного поля Y вдоль векторного поля X. Как и в разделе 6.2, можно убедиться, что это отображение обладает следующими свойствами:

$$\begin{aligned} \nabla_{X+Y} Z &= \nabla_X Z + \nabla_Y Z, \\ \nabla_{fX} Y &= f \nabla_X Y, \\ \nabla_X (Y+Z) &= \nabla_X Y + \nabla_X Z, \\ \nabla_X (fY) &= (Xf) Y + f \nabla_X Y, \end{aligned} \tag{6.33}$$

где $f \in \mathcal{C}^\infty(\mathbb{M})$ и $X, Y, Z \in \mathcal{X}(\mathbb{M})$. Верно также обратное утверждение: отображение ∇ со свойствами (6.33) определяет аффинную связность на касательном расслоении. Иногда его принимают в качестве определения аффинной связности.

Связность, определенная на касательном и кокасательном расслоении индуцирует связность на произвольных тензорных расслоениях с помощью правил (6.29). В компонентах ковариантная производная от тензорных полей будет выписана в следующем разделе. Индуцированная ковариантная производная линейна,

$$\nabla(aY \oplus bZ) = a \nabla Y \oplus b \nabla Z, \qquad a, b \in \mathbb{R},$$

где $Y, Z \in \mathcal{T}(\mathbb{M})$ – произвольные тензорные поля. Справедливо правило Лейбница

$$\nabla(Y \otimes Z) = (\nabla Y) \otimes Z + Y \otimes (\nabla Z). \tag{6.34}$$

Нетрудно проверить, что ковариантная производная перестановочна с каждым свертыванием: $\nabla C = C \nabla$, где C – оператор свертки. Кроме того, ковариантная производная вдоль векторного поля ∇_X сохраняет тип тензорных полей $\nabla_X \mathcal{T}^r_s(\mathbb{M}) \subset \mathcal{T}^r_s(\mathbb{M})$. Таким образом, ковариантная производная вдоль векторного поля ∇_X является дифференцированием в тензорной алгебре, которое было определено в разделе 3.14. В силу теоремы 3.14.1 ковариантная производная вдоль векторного поля ∇_X отличается от производной Ли L_X на дифференцирование, порожденное тензорным полем типа $(1,1)$.

Тензор кривизны (6.19) для аффинной связности (6.30) в компонентах принимает вид

$$R_{\alpha\beta\gamma}{}^{\delta} = \partial_\alpha \Gamma_{\beta\gamma}{}^{\delta} - \partial_\beta \Gamma_{\alpha\gamma}{}^{\delta} - \Gamma_{\alpha\gamma}{}^{\epsilon}\Gamma_{\beta\epsilon}{}^{\delta} + \Gamma_{\beta\gamma}{}^{\epsilon}\Gamma_{\alpha\epsilon}{}^{\delta}. \quad \square \tag{6.35}$$

Рассмотрим тензор кривизны (6.35), как отображение касательных пространств

$$\mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \times \mathcal{X}(\mathbb{M}) \to \mathcal{X}(\mathbb{M})$$

и обозначим

$$R(X, Y, Z) = X^\alpha Y^\beta Z^\gamma R_{\alpha\beta\gamma}{}^{\delta}\partial_\delta.$$

Тогда нетрудно проверить справедливость следующего тождества:

$$R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

которое является частным случаем предложения 6.2.2. Это соотношение иногда принимают в качестве определения кривизны R аффинной связности ∇.

Определение. 2-форма со значениями в касательном пространстве,

$$T = \frac{1}{2} dx^\alpha \wedge dx^\beta T_{\alpha\beta}{}^{\gamma}\partial_\gamma,$$

определенная равенством:

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y], \quad \forall X, Y \in \mathcal{X}(\mathbb{M}). \tag{6.36}$$

называется *тензором кручения*. Компоненты тензора кручения равны антисимметричной части компонент аффинной связности:

$$T_{\alpha\beta}{}^{\gamma} = \Gamma_{\alpha\beta}{}^{\gamma} - \Gamma_{\beta\alpha}{}^{\gamma}. \quad \square \tag{6.37}$$

Поскольку при преобразовании координат неоднородное слагаемое в (6.31) симметрично по индексам β', γ', то компоненты кручения действительно преобразуются по тензорному закону. Таким образом, аффинная связность определяет два тензора: тензор кривизны и тензор кручения.

6.4 Связность на расслоении реперов

Существует тесная взаимосвязь между касательным расслоением и расслоением реперов. Начнем с конструктивного построения расслоения реперов. Предположим, что задано дифференцируемое многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$.

Определение. *Репером* в точке $x \in \mathbb{M}$ называется упорядоченный набор объектов $p = (x, e_1, \ldots, e_n) = \{x, e_a\}$, $a = 1, \ldots, n$, где $\{e_a\}$ – набор линейно независимых касательных векторов в точке x. Множество всех реперов на $\mathbb{M}$ обозначим через $\mathbb{L}(\mathbb{M})$. Ниже мы построим дифференцируемую структуру на $\mathbb{L}(\mathbb{M})$, тем самым превратив множество всех реперов в дифференцируемое многообразие размерности $\dim \mathbb{L}(\mathbb{M}) = n + n^2$ такое, что естественная проекция

$$\pi(x, e_1, \ldots, e_n) = x, \tag{6.38}$$

является гладким отображением $\mathbb{L}(\mathbb{M}) \to \mathbb{M}$. Тогда $\mathbb{L}(\mathbb{M})$ называется *расслоением реперов* на $\mathbb{M}$. $\square$

Замечание. Забегая вперед, скажем, что расслоение реперов представляет собой главное расслоение $\mathbb{L}(\mathbb{M}) = \mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{GL}(n,\mathbb{R})\Big)$ с базой $\mathbb{M}$, проекцией π и структурной группой $\mathbb{GL}(n,\mathbb{R})$, с которым ассоциировано касательное расслоением $\mathbb{T}(\mathbb{M})$. □

Процедура построения дифференцируемой структуры на $\mathbb{L}(\mathbb{M})$ аналогична построению дифференцируемой структуры на тензорных расслоениях. Пусть $\mathbb{U} \subset \mathbb{M}$ – некоторая карта с координатами x^α, $\alpha = 1, \ldots, n$. Координатный базис касательного расслоения ∂_α образует координатный репер в каждой точке. Тогда произвольный репер можно разложить по координатному базису

$$e_a = e^\alpha{}_a \partial_\alpha. \tag{6.39}$$

По-определению, совокупность чисел $e^\alpha{}_a$ образует невырожденную $n \times n$ матрицу, т.е. представляет собой элемент группы $\mathbb{GL}(n,\mathbb{R})$ в каждой точке $x \in \mathbb{M}$. Поэтому можно определить отображение $\chi: \ \pi^{-1}(\mathbb{U}) \to \mathbb{U} \times \mathbb{GL}(n,\mathbb{R})$ такое, что для каждой точки $(x, e_1, \ldots, e_n) \in \mathbb{L}(\mathbb{M})$ и $x \in \mathbb{U}$, $e^\alpha{}_a \in \mathbb{GL}(n,\mathbb{R})$ имеем равенство

$$\chi(x, e_1, \ldots, e_n) = (x, e^\alpha{}_a),$$

где репер e_a задан равенством (6.39). Очевидно, что отображение χ является взаимно однозначным.

Теперь выберем атлас $\{\mathbb{U}_i\}$ на $\mathbb{M}$ с построенными выше отображениями $\{\chi_i\}$. Прообразы всех открытых подмножеств топологических произведений $\mathbb{U}_i \times \mathbb{GL}(n,\mathbb{R})$ для отображений χ_i образует базу топологии на $\mathbb{L}(\mathbb{M})$. При этом все отображения χ_i будут гомеоморфизмами.

Благодаря отображению χ_i, прообраз проекции $\pi^{-1}(\mathbb{U}_i)$ становится координатной окрестностью расслоения реперов $\mathbb{L}(\mathbb{M})$. Пусть в пересекающихся областях $\mathbb{U}_i \cap \mathbb{U}_j \neq \emptyset$ заданы координаты x^α и $x^{\alpha'}$ соответственно. Тогда координатами репера $(x, e_1, \ldots, e_n)$ в пересечении $\mathbb{U}_i \cap \mathbb{U}_j$ будут наборы из $n+n^2$ чисел $\{x^\alpha, e^\beta{}_a\}$ и $\{x^{\alpha'}, e^{\beta'}{}_a\}$, которые связаны преобразованием

$$\begin{aligned} x^{\alpha'} &= x^{\alpha'}(x), \\ e^{\alpha'}{}_a &= e^\alpha{}_a \partial_\alpha x^{\alpha'}. \end{aligned} \tag{6.40}$$

Если преобразование координат на $\mathbb{M}$ класса $\mathcal{C}^\infty$, то мы получаем бесконечно дифференцируемую структуру на $\mathbb{L}(\mathbb{M})$. Таким образом расслоение реперов становится дифференцируемым $(n+n^2)$-мерным многообразием класса $\mathcal{C}^\infty$. При этом естественная проекция $\pi: \ \mathbb{L}(\mathbb{M}) \to \mathbb{M}$ является гладким сюръективным отображением.

Сечение расслоения реперов $\mathbb{L}(\mathbb{M})$ мы будем обозначать $e_a(x)$. Его также можно разложить по координатному базису $e_a(x) = e^\alpha{}_a(x)\partial_\alpha$. Мы предполагаем, что все компоненты этого разложения являются гладкими функциями: $e^\alpha{}_a(x) \in \mathcal{C}^\infty(\mathbb{U})$. В разделе 6.2 под репером понималось именно сечение расслоения реперов.

Замечание. В формуле (6.39) компоненты разложения $e^\alpha{}_a$ рассматриваются, как координаты (числа) на пространстве главного расслоения $\mathbb{L}(\mathbb{M})$, а не как функции $e^\alpha{}_a(x)$. Формулы (6.40) – это также преобразование координат на $\mathbb{L}(\mathbb{M})$. В физической литературе сечение расслоения реперов $e_a(x)$ принято называть просто репером. Для краткости мы также будем употреблять термин репер для обозначения сечения расслоения реперов там, где это не вызывает двусмысленности. □

Отображение χ задает локально структуру прямого произведения на расслоении реперов $\mathbb{L}(\mathbb{M})$. То есть прообраз $\pi^{-1}(\mathbb{U})$ диффеоморфен прямому произведению $\mathbb{U}\times\mathbb{GL}(n,\mathbb{R})$.

Сужение отображения χ на каждый слой $\chi_x:\ \pi^{-1}(x)\to\mathbb{GL}(n,\mathbb{R})$ является гомоморфизмом групп. Если $x\in\mathbb{U}_i\cap\mathbb{U}_j$ тогда отображения $f_{ij}(x)=\chi_{j,x}\circ\chi^{-1}_{i,x}$ (функции перехода) представляют собой автоморфизмы группы $\mathbb{GL}(n,\mathbb{R})$. Из (6.40) следует, что преобразование $\chi_{j,x}\circ\chi^{-1}_{i,x}$ является умножением справа в группе $\mathbb{GL}(n,\mathbb{R})$ на матрицу Якоби $J_{ij\,\alpha}{}^{\alpha'}=\partial x^{\alpha'}_{(j)}/\partial x^{\alpha}_{(i)}$. Это значит, что набор матриц Якоби J_{ij} представляет собой совокупность функций склейки на расслоении реперов.

Замечание. Типичным слоем расслоения реперов является группа $\mathbb{GL}(n,\mathbb{R})$, на которой нет структуры векторного пространства. Поэтому расслоение реперов не является векторным расслоением.

Поскольку каждый слой $\pi^{-1}(x)$ диффеоморфен и гомоморфен $\mathbb{GL}(n,\mathbb{R})$, то он изоморфен группе Ли: $\pi^{-1}(x)\simeq\mathbb{GL}(n,\mathbb{R})$. □

Структурная группа $\mathbb{GL}(n,\mathbb{R})$ естественным образом действует на расслоении реперов и представляет собой группу преобразований $\mathbb{L}(\mathbb{M})$. Пусть $S_a{}^b\in\mathbb{GL}(n,\mathbb{R})$. Тогда правое действие r_S структурной группы на $\mathbb{L}(\mathbb{M})$ определим следующим образом

$$r_S(x,e_a)=(x,e'_a),\quad\text{где}\quad e'_a=S_a{}^b e_b.$$

Замечание. Мы записываем правое действие структурной группы в виде матричного умножения слева. В этом нет неоднозначности, т.к. суммирование всегда проводится по одному нижнему и одному верхнему индексу, $S_a{}^b e_b=e_b S_a{}^b$. Наши обозначения вызваны принятым правилом суммирования “с десяти до четырех”. Это правило является следствием общепринятой записи $dx^\alpha\partial_\alpha$. Другая запись $\partial_\alpha dx^\alpha$ может вызвать недоразумения. Кроме того, векторы репера принято нумеровать нижним индексом, поскольку координатные векторы ∂_α образуют координатный репер. □

Очевидно, что каждое преобразование r_S задает диффеоморфизм $\mathbb{L}(\mathbb{M})$ и сохраняет слои: $\pi\circ r_S=\pi$. Действие r_S называется *правым действием структурной группы*. Если $S,T\in\mathbb{GL}(n,\mathbb{R})$, то

$$r_{ST}=r_S\circ r_T.$$

Действие группы преобразований $\mathbb{GL}(n,\mathbb{R})$ свободно, т.е. любой элемент группы, отличный от единичного, перемещает все точки многообразия $\mathbb{L}(\mathbb{M})$.

Обозначим матрицу, обратную к реперу $e^\alpha{}_a$, через $e_\beta{}^b$:

$$e^\alpha{}_a e_\alpha{}^b=\delta^b_a,\qquad e_\alpha{}^a e^\beta{}_a=\delta^\beta_\alpha.$$

Тогда при преобразовании координат (6.40) на пространстве расслоения $\mathbb{L}(\mathbb{M})$ она преобразуется, как ковекторное поле

$$e_{\alpha'}{}^a=\partial_{\alpha'}x^\alpha e_\alpha{}^a.$$

Отсюда следует, что 1-формы (корепер)

$$e^a=dx^\alpha e_\alpha{}^a \tag{6.41}$$

не зависят от выбора координат на $\mathbb{M}$. Это значит, что e^a являются дифференциальными 1-формами, заданными в каждой координатной окрестности согласованным

образом и, следовательно, определяющими некоторую 1-форму на всем пространстве расслоения $\mathbb{L}(\mathbb{M})$. Напомним, что числа $e^\alpha{}_a$ в (6.41) рассматриваются в качестве координат на $\mathbb{L}(\mathbb{M})$. Поэтому $e_\alpha{}^a$ представляют собой набор хорошо определенных функций на $\mathbb{L}(\mathbb{M})$, поскольку $\det e^\alpha{}_a \neq 0$.

Пфаффова система уравнений

$$e^a = 0 \tag{6.42}$$

определяет n^2-мерное касательное подпространство $\mathbb{V}_p(\mathbb{L})$ в касательном пространстве $\mathbb{T}_p(\mathbb{L})$ к каждой точке расслоения $p = (x, e) \in \mathbb{L}(\mathbb{M})$, которое называется *вертикальным пространством*. Система уравнений (6.42) в каждой координатной окрестности эквивалентна системе уравнений

$$dx^\alpha = 0.$$

Эта система уравнений является вполне интегрируемой, и ее максимальное интегральное многообразие задается уравнениями

$$x^\alpha = \mathrm{const},$$

т.е. является слоем $\pi^{-1}(x)$ в точке $x \in \mathbb{M}$. Это значит, что вертикальные подпространства являются касательными пространствами к слоям.

Векторное поле на пространстве расслоения $X \in \mathcal{X}(\mathbb{L})$ в компонентах имеет вид

$$X = X^\alpha \frac{\partial}{\partial x^\alpha} + X^\alpha{}_a \frac{\partial}{\partial e^\alpha{}_a}.$$

Поэтому векторное поле Y принадлежит вертикальному подпространству $Y \in \mathbb{V}_p(\mathbb{L})$ в каждой точке $p = (x, e) \in \mathbb{L}(\mathbb{M})$ тогда и только тогда, когда оно имеет компоненты только вдоль координат слоя:

$$Y = Y^\alpha{}_a \frac{\partial}{\partial e^\alpha{}_a}$$

Дифференциал π_* проекции (6.38) отображает касательные пространства $\mathbb{T}_p(\mathbb{L}) \to \mathbb{T}_x(\mathbb{M})$. Очевидно, что все векторы из вертикального подпространства $\mathbb{V}_p(\mathbb{L})$ отображаются в нулевой вектор $\mathbb{T}_x(\mathbb{M})$. Верно и обратное утверждение. Если какой либо вектор из $\mathbb{T}_p(\mathbb{L})$ отображается в нулевой вектор, то он принадлежит вертикальному подпространству. Это значит, что вертикальные подпространства являются ядром дифференциала проекции π_*.

Пусть на базе $\mathbb{M}$ задана аффинная связность ∇ или, что эквивалентно, 1-формы $\Gamma_\alpha{}^\beta = dx^\gamma \Gamma_{\gamma\alpha}{}^\beta$. Тогда ковариантный дифференциал векторного поля $e_a(x)$ равен

$$De_a = (de^\alpha{}_a + e^\beta{}_a \Gamma_\beta{}^\alpha) \otimes \partial_\alpha.$$

Если рассматривать $e^\alpha{}_a$, как координаты на $\mathbb{L}(\mathbb{M})$ (что мы и делаем), то выражение

$$De^\alpha{}_a = de^\alpha{}_a + e^\beta{}_a \Gamma_\beta{}^\alpha$$

является 1-формой на координатной окрестности $\pi^{-1}(\mathbb{U})$ расслоения реперов $\mathbb{L}(\mathbb{M})$. Поскольку 1-форма $De^\alpha{}_a$ при преобразовании координат является вектором по индексу α, то 1-форма

$$\omega_a{}^b = e_\alpha{}^b De^\alpha{}_a \tag{6.43}$$

не зависит от выбора системы координат и, значит, задает дифференциальную 1-форму на всем расслоении реперов $\mathbb{L}(\mathbb{M})$. 1-форма $\omega_a{}^b$ называется *формой связности* на расслоении реперов. Так как $(x^\alpha, e^\beta{}_a)$ являются координатами на $\mathbb{L}(\mathbb{U})$, то

$(dx^\alpha, de^\beta{}_a)$ представляют собой координатный базис кокасательного пространства к $\mathbb{L}(\mathbb{U})$. Поэтому $(x^\alpha, e^\beta{}_a, dx^\alpha, de^\beta{}_a)$ задают систему координат кокасательного расслоения $\mathbb{T}^*(\mathbb{L})$ к $\mathbb{L}(\mathbb{U})$. Теперь e^a и $\omega_a{}^b$ задают $n + n^2$ дифференциальных 1-форм на расслоении реперов $\mathbb{L}(\mathbb{U})$. В каждой координатной окрестности $\pi^{-1}(\mathbb{U})$ их можно выразить в виде линейной комбинации 1-форм dx^α и $de^\alpha{}_a$, и наоборот. Поскольку 1-формы e^a и $\omega_a{}^b$ линейно независимы, то они задают поле корепера на всем пространстве $\mathbb{L}(\mathbb{M})$. Дуальный к нему базис определяет поле репера на пространстве расслоения $\mathbb{L}(\mathbb{M})$ глобально.

Замечание. В общем случае сечение расслоения реперов $e_a(x)$ может не существовать глобально. Например, на двумерной сфере $\mathbb{S}^2$ не существует непрерывного векторного поля, которое не обращается в нуль ни в одной точке (теорема 11.2.1). Следовательно, глобального сечения $\mathbb{L}(\mathbb{S}^2)$ не существует. В то же время, поскольку на произвольном многообразии $\mathbb{M}$ существует аффинная связность, то на расслоении реперов $\mathbb{L}(\mathbb{M})$ всегда существует корепер e^a глобально. В этом смысле многообразие $\mathbb{L}(\mathbb{M})$ устроено проще, чем база $\mathbb{M}$. □

В локальной системе координат из определения 1-форм e^a и $\omega_a{}^b$ следует

$$\begin{aligned} dx^\alpha &= e^a e^\alpha{}_a, \\ de^\alpha{}_a &= -e^\beta{}_a \Gamma_\beta{}^\alpha + \omega_a{}^b e^\alpha{}_b. \end{aligned}$$

Взятие внешней производной от равенств (6.41) и (6.43) приводит к *структурным уравнениям* для аффинной связности

$$de^a - e^b \wedge \omega_b{}^a = \frac{1}{2} e^b \wedge e^c \, T_{bc}{}^a, \tag{6.44}$$

$$d\omega_a{}^b - \omega_a{}^c \wedge \omega_c{}^b = \frac{1}{2} e^c \wedge e^d \, R_{cda}{}^b, \tag{6.45}$$

где $T_{bc}{}^a = e^\beta{}_b e^\gamma{}_c T_{\beta\gamma}{}^\alpha e_\alpha{}^a$ и $R_{cda}{}^b = e^\gamma{}_c e^\delta{}_d e^\alpha{}_a R_{\gamma\delta\alpha}{}^\beta e_\beta{}^b$ – компоненты тензоров кручения и кривизны. Компоненты тензоров кручения $T_{\beta\gamma}{}^\alpha$ и кривизны $R_{\gamma\delta\alpha}{}^\beta$ в координатном базисе были определены ранее (6.37) и (6.35). Введем обозначения для 2-форм кручения и кривизны:

$$\begin{aligned} T^a &:= \frac{1}{2} e^b \wedge e^c T_{bc}{}^a, \\ R_a{}^b &:= \frac{1}{2} e^c \wedge e^d R_{cda}{}^b. \end{aligned} \tag{6.46}$$

Взятие внешней производной от уравнений структуры (6.44), (6.45) приводит к *тождествам Бианки* для аффинной связности:

$$\begin{aligned} dT^a + T^b \wedge \omega_b{}^a &= e^b \wedge R_b{}^a, \\ dR_a{}^b + R_a{}^c \wedge \omega_c{}^b - \omega_a{}^c \wedge R_c{}^b &= 0. \end{aligned} \tag{6.47}$$

Второе из этих тождеств совпадает с полученной ранее формулой (6.21). В компонентах тождества Бианки будут подробно рассмотрены в разделе 7.10.

Замечание. Структурные уравнения (6.44) и (6.45) можно принять за определение тензоров кручения и кривизны. □

В координатном базисе $e^a \mapsto dx^\alpha$, $\omega_a{}^b \mapsto \Gamma_\alpha{}^\beta = dx^\gamma \Gamma_{\gamma\alpha}{}^\beta$, и структурные уравнения принимают вид

$$\begin{aligned} -dx^\beta \wedge \Gamma_\beta{}^\alpha &= \frac{1}{2} dx^\beta \wedge dx^\gamma \, T_{\beta\gamma}{}^\alpha, \\ d\Gamma_\alpha{}^\beta - \Gamma_\alpha{}^\gamma \wedge \Gamma_\gamma{}^\beta &= \frac{1}{2} dx^\gamma \wedge dx^\delta \, R_{\gamma\delta\alpha}{}^\beta, \end{aligned} \tag{6.48}$$

где в правых частях стоят тензоры кручения $T_{\beta\gamma}{}^\alpha$ и кривизны $R_{\gamma\delta\alpha}{}^\beta$ и учтено равенство $ddx^\alpha = 0$.

Теорема 6.4.1. *Пусть на расслоении реперов $\mathbb{L}(\mathbb{M})$ заданы n^2 дифференциальных 1-форм $\omega_a{}^b$, которые вместе с кореперо́м e^a удовлетворяют структурным уравнениям (6.44), (6.45), где $T_{bc}{}^a$ и $R_{cda}{}^b$ – некоторые функции, определенные на $\mathbb{L}(\mathbb{M})$. Тогда существует аффинная связность ∇ такая, что выполнено равенство (6.43).*

Доказательство. См., например, [**?**]. □

Замечание. От функций, определяющих кручение и кривизну, нет надобности требовать выполнения тождеств Бианки, т.к. они являются следствием структурных уравнений, выполнение которых является условием теоремы. □

Ранее было отмечено, что пфаффова система уравнений $e^a = 0$ определяет распределение вертикальных подпространств $\mathbb{V}_p(\mathbb{L}) \subset \mathbb{T}_p(\mathbb{L})$ на расслоении реперов $\mathbb{L}(\mathbb{M})$. В каждой точке $(x, e) \in \mathbb{L}(\mathbb{M})$ система пфаффовых уравнений

$$\omega_a{}^b = 0 \tag{6.49}$$

определяет n-мерное касательное подпространство $\mathbb{H}_p(\mathbb{L}) \subset \mathbb{T}_p(\mathbb{L})$, которое называется *горизонтальным*. Таким образом, мы имеем распределение горизонтальных подпространств $\mathbb{H}(\mathbb{L})$, однозначно определенное аффинной связностью ∇ на расслоении реперов $\mathbb{L}(\mathbb{M})$. Оно имеет следующие свойства.

Теорема 6.4.2. *Пусть задано расслоение реперов $\mathbb{L}(\mathbb{M})$ и на базе $\mathbb{M}$ определена аффинная связность ∇. Тогда определены распределения вертикальных $\mathbb{V}_p(\mathbb{L})$ и горизонтальных $\mathbb{H}_p(\mathbb{L})$ подпространств, которые удовлетворяют свойствам:*

1) В каждой точке $(x, e) \in \mathbb{L}(\mathbb{M})$ касательное пространство $\mathbb{T}_p(\mathbb{L})$ разлагается в прямую сумму

$$\mathbb{T}_p(\mathbb{L}) = \mathbb{V}_p(\mathbb{L}) \oplus \mathbb{H}_p(\mathbb{L}). \tag{6.50}$$

При этом образ горизонтального пространства $\mathbb{H}_p(\mathbb{L})$ при проекции π изоморфен касательному пространству к базе $\mathbb{T}_x(\mathbb{M})$.

2) Горизонтальные пространства инвариантны относительно правого действия группы преобразований r_S, $S \in \mathbb{GL}(n, \mathbb{R})$. То есть

$$(r_S)_* \mathbb{H}_p(\mathbb{L}) = \mathbb{H}_{r_S(x,p)}(\mathbb{L}).$$

Доказательство. Докажем свойство 1). Поскольку сумма размерностей горизонтального $\mathbb{H}_p(\mathbb{L})$ и вертикального $\mathbb{V}_p(\mathbb{L})$ подпространств равна размерности касательного пространства $\mathbb{T}_p(\mathbb{L})$, то достаточно доказать, что $\mathbb{H}_p(\mathbb{L}) \cap \mathbb{V}_p(\mathbb{L}) = \emptyset$. От противного. Пусть $X \in \mathbb{H}_p(\mathbb{L}) \cap \mathbb{V}_p(\mathbb{L})$. Тогда, по-определению, выполнены уравнения

$$e^a(X) = 0, \quad \omega_a{}^b(X) = 0.$$

Поскольку 1-формы $\{e^a, \omega_b{}^c\}$ образуют базис кореперов на $\mathbb{L}(\mathbb{M})$, то $X = 0$. Это доказывает разложение (6.50). Поскольку проекция π является гладким и сюрьективным отображением, то ее дифференциал π_* является сюрьективным гомоморфизмом. Поскольку $\pi_*\Big(\mathbb{V}_p(\mathbb{L})\Big) = 0$, то отображение $\pi_* : \ \mathbb{H}_p(\mathbb{L}) \to \mathbb{T}_x(\mathbb{M})$ является изоморфизмом.

Для доказательства свойства 2) необходимо проверить, что правое действие группы преобразований на 1-формы $\omega_a{}^b$ имеет вид

$$(r_S)^*\omega_a{}^b = S_a{}^c\omega_c{}^d S^{-1b}_d.$$

Действительно, это следует из определения (6.43) и равенства

$$de'^\alpha{}_a = \frac{\partial e'^\alpha{}_a}{\partial e^\beta{}_b} de^\beta{}_b = S_a{}^b de^\alpha{}_b.$$

Поскольку распределение горизонтальных подпространств $\mathbb{H}(\mathbb{L})$ является аннигилятором подпространства $\omega_a{}^b$, то из правила действия левых преобразований следует свойство 2). □

Верно также обратное утверждение. Если в касательном расслоении $\mathbb{T}(\mathbb{L})$ задано дифференцируемое n-мерное распределение горизонтальных подпространств $\mathbb{H}$, удовлетворяющих свойствам 1) и 2), то на $\mathbb{M}$ существует аффинная связность ∇ такая, что $\mathbb{H}$ является распределением горизонтальных подпространств расслоения реперов $\mathbb{L}(\mathbb{M})$ для связности ∇. Поэтому, с точки зрения расслоения реперов, задание аффинной связности эквивалентно заданию распределения горизонтальных подпространств со свойствами 1), 2). Далее эти свойства лягут в основу определения связности на главных расслоениях общего вида.

Замечание. Поскольку аффинная связность ∇ зависит дифференцируемо от $p \in \mathbb{L}(\mathbb{M})$, то распределение горизонтальных подпространств $\mathbb{H}_p(\mathbb{L})$ также дифференцируемо зависит от точки p. □

В формулах (6.44)–(6.47) все формы рассматривались на пространстве расслоения $\mathbb{L}(\mathbb{M})$. В частности компоненты тензоров кручения $T_{bc}{}^a$ и кривизны $R_{cda}{}^b$ в структурных уравнениях (6.44) и (6.45) явно зависят от координат слоя $e^\alpha{}_a$. Эти формы можно спустить на базу с помощью возврата отображения.

Определение. Рассмотрим сечение $\sigma : \mathbb{M} \to \mathbb{L}(\mathbb{M})$, которому в локальных координатах соответствует репер $e^\alpha{}_a(x)$. Тогда дифференциал сечения σ_* отображает касательные векторы к базе в касательные векторы к сечению:

$$\mathcal{X}(\mathbb{M}) \ni X = X^\alpha\partial_\alpha \ \mapsto \ \sigma_* X = X^\alpha\partial_\alpha + X^\beta\partial_\beta e^\alpha{}_a \frac{\partial}{\partial e^\alpha{}_a} \in \mathcal{X}\Big(\sigma(\mathbb{M})\Big).$$

Следовательно, возврат отображения σ^* спускает форму связности для каждого сечения на базу:

$$\sigma^*\omega_a{}^b := dx^\alpha\omega_{\alpha a}{}^b = dx^\alpha e_\gamma{}^b(\partial_\alpha e^\gamma{}_a + e^\beta{}_a\Gamma_{\alpha\beta}{}^\gamma). \tag{6.51}$$

Форма $dx^\alpha\omega_{\alpha a}{}^b$ называется *локальной формой связности* на расслоении реперов. Она получается из формы связности (6.43) простой заменой: $e^\alpha{}_a \mapsto e^\alpha{}_a(x)$ и $de^\alpha{}_a \mapsto dx^\beta\partial_\beta e^\alpha{}_a$. Аналогично спускаются на базу 2-формы кручения и кривизны (6.46):

$$T^a = \frac{1}{2}dx^\alpha \wedge dx^\beta\, T_{\alpha\beta}{}^a = de^a - e^b \wedge \omega_b{}^a, \tag{6.52}$$

$$R_a{}^b = \frac{1}{2}dx^\alpha \wedge dx^\beta\, R_{\alpha\beta a}{}^b = d\omega_a{}^b - \omega_a{}^c \wedge \omega_c{}^b, \tag{6.53}$$

где введены 1-формы локальной связности и корепера:

$$\omega_a{}^b = dx^\alpha \omega_{\alpha a}{}^b, \qquad e^a = dx^\alpha e_\alpha{}^a, \tag{6.54}$$

и

$$T_{\alpha\beta}{}^a = \partial_\alpha e_\beta{}^a - e_\alpha{}^b \omega_{\beta b}{}^a - (\alpha \leftrightarrow \beta), \tag{6.55}$$

$$R_{\alpha\beta a}{}^b = \partial_\alpha \omega_{\beta a}{}^b - \omega_{\alpha a}{}^c \omega_{\beta c}{}^b - (\alpha \leftrightarrow \beta). \tag{6.56}$$

2-формы (6.52) и (6.53) называются *локальными формами* кручения и кривизны. □

Замечание. Для локальных форм мы сохранили прежние обозначения, чтобы не вводить новых. Фактически, в приложениях при проведении вычислений используются локальные формы, заданные на базе $\mathbb{M}$, а не на расслоении реперов $\mathbb{L}(\mathbb{M})$. □

Тождества Бианки (6.47) после спуска на базу имеют точно такой же вид, только вместо компонент корепера $e_\alpha{}^a$ надо рассматривать сечение $e_\alpha{}^a(x)$, вместо формы связности $\omega_a{}^b$, определенную в (6.43), – локальную форму связности $dx^\alpha \omega_{\alpha a}{}^b(x)$ и вместо форм кручения и кривизны – соответствующие им локальные формы.

Определение. Поля $e_\alpha{}^a(x)$ и $\omega_{\alpha a}{}^b(x)$, заданные на $\mathbb{M}$, в аффинной геометрии называются *переменными Картана.* □

Переменные Картана используются в приложениях, т.к. часто позволяют упростить вычисления. Формулы (6.55) и (6.53) дают выражение для тензоров кручения и кривизны в переменных Картана.

Глава 7

Аффинная геометрия. Локальное рассмотрение

В современной дифференциальной геометрии принят бескоординатный язык. Это оправдано тем, что все геометрические объекты, например, тензорные поля, являются понятиями, которые не зависят от выбора системы координат. Этот язык удобен для определений и формулировки утверждений. Однако, он абстрактен, и требуется немалое время для его усвоения. Кроме того, при проведении вычислений, особенно в моделях математический физики, используется координатный подход, в котором геометрические объекты отождествляются с набором своих компонент. Этот язык действительно необходим для проведения вычислений и является более наглядным. В настоящем разделе мы дадим определения и опишем основные свойства объектов аффинной геометрии в компонентах. По сути дела это означает, что мы рассматриваем топологически тривиальное многообразие $\mathbb{M}$, которое диффеоморфно евклидову пространству $\mathbb{R}^n$ и покрывается одной картой. Это не значит, что координатный подход является менее строгим. Как только установлено, что некоторый объект или утверждение не зависят от выбора системы координат в $\mathbb{R}^n$, то все сказанное сразу переносится на нетривиальные многообразия.

7.1 Локальное определение аффинной связности

Ввиду важности понятия аффинной связности покажем, как она вводится локально без обращения к общей теории связностей в векторных расслоениях. Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$ в некоторой карте заданы: функция $f \in \mathcal{C}^1(\mathbb{M})$, дифференцируемое векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ и дифференцируемая 1-форма $A = dx^\alpha A_\alpha \in \Lambda_1(\mathbb{M})$. Пусть на этом многообразии задана также аффинная связность ∇. Тогда в компонентах ковариантные производные имеют вид:

$$\nabla_\alpha f = \partial_\alpha f \tag{7.1}$$

$$\nabla_\alpha X^\beta = \partial_\alpha X^\beta + X^\gamma \Gamma_{\alpha\gamma}{}^\beta, \tag{7.2}$$

$$\nabla_\alpha A_\beta = \partial_\alpha A_\beta - \Gamma_{\alpha\beta}{}^\gamma A_\gamma, \tag{7.3}$$

где $\Gamma_{\alpha\beta}{}^\gamma(x)$ – компоненты локальной формы аффинной связности, которые, мы предполагаем, являются гладкими функциями от координат x^α. То есть ковариантная производная для скалярного поля совпадает с обычной частной производной, а для векторного поля и 1-формы появляются дополнительные слагаемые, линейные по

компонентам. Отметим, что дополнительные слагаемые в (7.2) и (7.3) имеют разные знаки.

Покажем, что ковариантная производная от компонент векторного поля преобразуется, как тензор. Рассмотрим преобразование координат (2.55), тогда компоненты векторного поля преобразуется по правилу (3.11). Дифференцируя соотношение (3.11) по $x^{\beta'}$, получаем два слагаемых

$$\partial_{\beta'} X^{\alpha'} = \partial_{\beta'} x^{\beta} \partial_{\beta} X^{\alpha} \partial_{\alpha} x^{\alpha'} + \partial_{\beta'} x^{\beta} X^{\alpha} \partial^2_{\beta\alpha} x^{\alpha'}. \tag{7.4}$$

Первое слагаемое в правой части соответствует тензорному закону преобразования для производной от векторного поля, в то время как второе слагаемое этот закон нарушает.

Чтобы получить тензорный закон преобразования в общем случае вводится понятие ковариантной производной, которая содержит дополнительное слагаемое в (7.2). Если потребовать, чтобы ковариантная производная после преобразования координат имела тот же вид с некоторыми новыми компонентами $\Gamma_{\beta'\gamma'}{}^{\alpha'}$,

$$\partial_{\beta'} X^{\alpha'} + \Gamma_{\beta'\gamma'}{}^{\alpha'} X^{\gamma'} = \partial_{\beta'} x^{\beta} (\partial_{\beta} X^{\alpha} + \Gamma_{\beta\gamma}{}^{\alpha} X^{\gamma}) \partial_{\alpha} x^{\alpha'},$$

то с учетом (7.4) получим следующий закон преобразования компонент аффинной связности

$$\Gamma_{\beta'\gamma'}{}^{\alpha'} = \partial_{\beta'} x^{\beta} \partial_{\gamma'} x^{\gamma} \Gamma_{\beta\gamma}{}^{\alpha} \partial_{\alpha} x^{\alpha'} - \partial_{\beta'} x^{\beta} \partial_{\gamma'} x^{\gamma} \partial^2_{\beta\gamma} x^{\alpha'}, \tag{7.5}$$

С учетом тождества

$$\partial_{\alpha} (\partial_{\beta} x^{\beta'} \partial_{\beta'} x^{\gamma}) = \partial_{\alpha} \delta^{\beta}_{\gamma} = 0 \quad \Leftrightarrow \quad \partial^2_{\alpha\beta} x^{\beta'} = -\partial_{\alpha} x^{\alpha'} \partial_{\beta} x^{\beta'} \partial^2_{\alpha'\beta'} x^{\gamma} \partial_{\gamma} x^{\beta'}$$

закон преобразования (7.5) можно переписать в эквивалентной форме (6.31), отличающейся вторым слагаемым.

Закон преобразования компонент аффинной связности отличается от тензорного закона наличием неоднородных слагаемых в (6.31) и (7.5), которые содержат вторые производные от функций перехода. Если ограничить класс допустимых преобразований координат аффинными (линейными неоднородными), то аффинная связность будет преобразовываться как тензор.

Сами по себе компоненты аффинной связности не являются компонентами тензора, однако они позволяют строить новые тензорные поля из заданных с помощью ковариантного дифференцирования. Собственно, название “ковариантное” и отражает это обстоятельство. Нетрудно проверить, что ковариантная производная от 1-формы (7.3) является тензорным полем второго ранга типа $(0,2)$. Ковариантное дифференцирование было продолжено на тензоры произвольного ранга с помощью формул (6.29). В компонентах это выглядит следующим образом. Помимо обычной частной производной в ковариантную производную входят слагаемые с компонентами аффинной связности со знаком плюс для каждого контравариантного и минус для каждого ковариантного индекса:

$$\begin{aligned}\nabla_{\alpha} K_{\beta_1\ldots\beta_s}{}^{\gamma_1\ldots\gamma_r} = \partial_{\alpha} K_{\beta_1\ldots\beta_s}{}^{\gamma_1\ldots\gamma_r} - \Gamma_{\alpha\beta_1}{}^{\delta} K_{\delta\beta_2\ldots\beta_s}{}^{\gamma_1\ldots\gamma_r} - \ldots - \Gamma_{\alpha\beta_s}{}^{\delta} K_{\beta_1\ldots\beta_{s-1}\delta}{}^{\gamma_1\ldots\gamma_r} \\ + K_{\beta_1\ldots\beta_s}{}^{\delta\gamma_2\ldots\gamma_r} \Gamma_{\alpha\delta}{}^{\gamma_1} + \ldots + K_{\beta_1\ldots\beta_s}{}^{\gamma_1\ldots\gamma_{r-1}\delta} \Gamma_{\alpha\delta}{}^{\gamma_r}.\end{aligned}$$

Пример 7.1.1. Для тензора второго ранга типа $(1,1)$ ковариантная производная имеет вид

$$\nabla_{\alpha} K_{\beta}{}^{\gamma} = \partial_{\alpha} K_{\beta}{}^{\gamma} - \Gamma_{\alpha\beta}{}^{\delta} K_{\delta}{}^{\gamma} + \Gamma_{\alpha\delta}{}^{\gamma} K_{\beta}{}^{\delta}. \tag{7.6}$$

В частном случае, ковариантная производная от символа Кронекера (3.56) тождественно равна нулю

$$\nabla_\alpha \delta^\gamma_\beta = 0,$$

поскольку равна нулю частная производная $\partial_\alpha \delta^\gamma_\beta = 0$, а слагаемые со связностью сокращаются. То есть символ Кронекера ковариантно постоянен (параллелен) относительно произвольной аффинной связности. □

Можно показать с помощью прямых вычислений (если доказательство в общем виде, которое было дано ранее, кого то не убеждает), что определенная таким образом ковариантная производная от тензора произвольного типа (r, s) действительно дает тензор типа $(r, s+1)$.

Поскольку произвольный тензор типа (r, s) является полилинейным отображением (3.58), то вид ковариантной производной от тензорного поля в компонентах однозначно определяется четырьмя условиями: видом ковариантной производной от функции, векторного и ковекторного полей (7.1)–(7.3) и правилом Лейбница (6.34).

Пример 7.1.2. Найдем вид ковариантной производной от тензорного поля второго ранга $K_\alpha{}^\beta$. Ковариантная производная от функции $K_\gamma{}^\beta X^\gamma A_\beta \in \mathcal{C}^\infty(\mathbb{M})$, где $X^\gamma \in \mathcal{X}(\mathbb{M})$ и $A_\beta \in \Lambda_1(\mathbb{M})$ – произвольные векторное и ковекторное поля, совпадает с обычной:

$$\nabla_\alpha (K_\gamma{}^\beta X^\gamma A_\beta) = \partial_\alpha (K_\gamma{}^\beta X^\gamma A_\beta).$$

Воспользовавшись правилом Лейбница и видом ковариантных производных (7.2), (7.3), получим формулу для ковариантной производной от тензора второго ранга (7.6). □

Ковариантная производная коммутирует с произвольным свертыванием:

$$\nabla_\alpha (\delta^{\beta_j}_{\alpha_i} K_{\beta_1 \cdots \beta_s}{}^{\alpha_1 \cdots \alpha_r}) = \delta^{\beta_j}_{\alpha_i} \nabla_\alpha K_{\beta_1 \cdots \beta_s}{}^{\alpha_1 \cdots \alpha_r}, \qquad 1 \le i \le r,\ 1 \le j \le s.$$

что следует из ковариантного постоянства символа Кронекера. В частном случае,

$$\partial_\alpha (X^\beta A_\beta) = \nabla_\alpha (X^\beta A_\beta) = (\nabla_\alpha X^\beta) A_\beta + X^\beta (\nabla_\alpha A_\beta).$$

Определение. Произвольному векторному полю можно поставить в соответствие бесконечно малую величину

$$DX^\alpha = dx^\beta \nabla_\beta X^\alpha = dx^\beta (\partial_\beta X^\alpha + \Gamma_{\beta\gamma}{}^\alpha X^\gamma), \tag{7.7}$$

которая называется *ковариантным дифференциалом*. Аналогично определяется ковариантный дифференциал для произвольного тензорного поля. □

Для функции ковариантный дифференциал совпадает с обычным.

Определение. Рассмотрим произвольную дифференцируемую кривую $\gamma = \{x^\alpha(t)\}$ в римановом пространстве с положительно определенной метрикой. Тогда длина дуги кривой отлична от нуля

$$ds = dt\sqrt{\dot{x}^2} \neq 0, \qquad \text{где } \dot{x}^2 = \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}. \tag{7.8}$$

Выражение

$$\frac{DX^\alpha}{ds} = u^\beta \nabla_\beta X^\alpha, \tag{7.9}$$

где

$$u^\alpha := \frac{\dot{x}^\alpha}{\sqrt{\dot{x}^2}}, \qquad u^2 = 1, \tag{7.10}$$

– единичный касательный вектор к кривой, называется *ковариантной производной вектора X вдоль кривой γ.* □

Если метрика имеет лоренцеву сигнатуру, то ковариантную производную можно определить либо вдоль времениподобных, либо вдоль пространственноподобных кривых. В последнем случае $\dot{x}^2$ заменяется на $-\dot{x}^2$ под знаком корня.

Аналогично определяется ковариантная производная вдоль кривой от произвольного тензорного поля.

Пусть на многообразии задано векторное поле $X \in \mathcal{X}(\mathbb{M})$. Тогда *ковариантная производная вдоль векторного поля X* от тензорного поля $K_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\beta_s}$ типа (r,s) дает тензорное поле того же типа с компонентами

$$(\nabla_X K)_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\beta_s} = X^\gamma \nabla_\gamma K_{\alpha_1\ldots\alpha_r}{}^{\beta_1\ldots\beta_s}. \tag{7.11}$$

Пример 7.1.3. В частности, для двух векторных полей справедливо равенство

$$(\nabla_X Y)^\alpha = X^\beta \nabla_\beta Y^\alpha. \quad \square \tag{7.12}$$

Рассмотренная выше ковариантная производная от тензора вдоль кривой является частным случаем ковариантной производной вдоль векторного поля, которое касается кривой. Ковариантная производная вдоль кривой определена только в тех точках многообразия $\mathbb{M}$, через которые проходит кривая.

Наличие неоднородного слагаемого в законе преобразования аффинной связности (7.5) или (6.31) не позволяет их складывать как тензорные поля и умножать даже на числа. Например, если на многообразии задано две аффинные связности $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$, то их сумма в общем случае связностью не является. В то же время нетрудно проверить, что сумма

$$\Gamma_{\alpha\beta}{}^\gamma = f\Gamma_{1\alpha\beta}{}^\gamma + (1-f)\Gamma_{2\alpha\beta}{}^\gamma, \tag{7.13}$$

где $f \in \mathcal{C}^1(\mathbb{M})$, задает компоненты некоторой аффинной связности.

Разность двух аффинных связностей является тензорным полем типа $(1,2)$, т.к. неоднородные слагаемые сокращаются.

Определение. *Вариацией* аффинной связности $\delta\Gamma_{\alpha\beta}{}^\gamma$ называется разность двух связностей, заданных на многообразии $\mathbb{M}$. □

Вариация связности является тензорным полем типа $(1,2)$.

7.2 Кручение и неметричность

Продолжим локальное изучение аффинной связности. В общем случае компоненты связности $\Gamma_{\alpha\beta}{}^\gamma$ никакой симметрии по индексам не имеет и никак не связана с метрикой $g_{\alpha\beta}$, поскольку эти понятия определяют на многообразии $\mathbb{M}$ разные геометрические операции. А именно, метрика многообразия определяет в каждой точке скалярное произведение векторов из касательного пространства, а аффинная связность определяет ковариантное дифференцирование и параллельный перенос тензоров. Геометрия на многообразии $\mathbb{M}$ определяется метрикой и аффинной связностью. Будем говорить, что на $\mathbb{M}$ задана *аффинная геометрия*, если заданы достаточно гладкие метрика и аффинная связность, т.е. задано три объекта $(\mathbb{M}, g, \Gamma)$.

Замечание. В общем случае метрика и аффинная связность задаются произвольным образом и являются совершенно независимыми геометрическими объектами. Поэтому при построении физических моделей их можно рассматривать как независимые поля, имеющие разную физическую интерпретацию. В настоящее время принято считать, что метрика описывает гравитационное взаимодействие. Физический смысл аффинной связности пока неясен. Это связано с тем, что физическая интерпретация связности зависит от конкретной модели. Соответствующие модели сложны с математической точки зрения и в настоящее время изучены недостаточно хорошо. □

По-определению, кручение многообразия в локальной системе координат равно антисимметричной части аффинной связности (6.37). Из закона преобразования связности (6.31) следует, что кручение является тензорным полем типа $(1,2)$.

Определение. Если на многообразии задана аффинная геометрия, то можно построить *тензор неметричности* $Q_{\alpha\beta\gamma}$. Он равен ковариантной производной от метрики

$$-Q_{\alpha\beta\gamma} := \nabla_\alpha g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} - \Gamma_{\alpha\beta}{}^{\delta} g_{\delta\gamma} - \Gamma_{\alpha\gamma}{}^{\delta} g_{\beta\delta}. \quad \square \tag{7.14}$$

Тензор неметричности, по-построению, симметричен относительно перестановки двух последних индексов

$$Q_{\alpha\beta\gamma} = Q_{\alpha\gamma\beta}.$$

Заметим, что для определения неметричности необходимы оба объекта: и метрика, и связность.

Таким образом по заданной метрике и аффинной связности построено два тензорных поля: кручение и тензор неметричности. Докажем, что по заданной метрике, кручению и тензору неметричности можно однозначно восстановить соответствующую аффинную связность. Уравнение (7.14) всегда можно решить относительно связности Γ. Действительно, линейная комбинация

$$\nabla_\alpha g_{\beta\gamma} + \nabla_\beta g_{\gamma\alpha} - \nabla_\gamma g_{\alpha\beta}$$

приводит к следующему общему решению для аффинной связности с опущенным верхним индексом

$$\begin{aligned}\Gamma_{\alpha\beta\gamma} := \Gamma_{\alpha\beta}{}^{\delta} g_{\delta\gamma} = \frac{1}{2}(\partial_\alpha g_{\beta\gamma} + \partial_\beta g_{\gamma\alpha} - \partial_\gamma g_{\alpha\beta}) + \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}) \\ + \frac{1}{2}(Q_{\alpha\beta\gamma} + Q_{\beta\gamma\alpha} - Q_{\gamma\alpha\beta}).\end{aligned} \tag{7.15}$$

Правая часть этого равенства симметрична по индексам α и β за исключением одного слагаемого, $\frac{1}{2}T_{\alpha\beta\gamma}$, что согласуется с определением тензора кручения (**??**). Заметим, что слагаемые с метрикой и тензором неметричности в (7.15) имеют одинаковый порядок индексов и знаков, а слагаемые с кручением при том же порядке индексов отличаются знаками. Таким образом, для того, чтобы на многообразии $\mathbb{M}$ задать аффинную геометрию, необходимо и достаточно задать три тензорных поля: метрику, кручение и неметричность. Подчеркнем еще раз, что все три объекта можно задать совершенно независимым образом, и в моделях математической физики их можно рассматривать как независимые динамические переменные.

Определение. Второе слагаемое в (7.15)

$$K_{\alpha\beta\gamma} = \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}) \tag{7.16}$$

называется тензором *кокручения.* □

Заметим, что уравнение (7.14) симметрично по индексам β, γ, т.е. число алгебраических уравнений в точности равно числу компонент симметричной части связности $\Gamma_{\beta\gamma}{}^{\alpha}$. Поэтому весь произвол в решении этой системы уравнений определяется тензором кручения.

Вычислив ковариантную производную от тождества $g^{\alpha\beta}g_{\beta\gamma} = \delta^{\alpha}_{\gamma}$, находим, что ковариантная производная от обратной метрики,

$$\nabla_\alpha g^{\beta\gamma} = Q_\alpha{}^{\beta\gamma}, \qquad Q_\alpha{}^{\beta\gamma} = g^{\beta\delta}g^{\gamma\epsilon}Q_{\alpha\delta\epsilon}, \tag{7.17}$$

отличается знаком от (7.14). Если тензор неметричности отличен от нуля, то подъем и опускание индексов не коммутирует с ковариантным дифференцированием.

Пример 7.2.1. $\nabla_\alpha X_\beta = (\nabla_\alpha X^\gamma)g_{\gamma\beta} - Q_{\alpha\gamma\beta}X^\gamma.$ □

Для задания аффинной геометрии на многообразии необходимо задать метрику и аффинную связность. Рассмотрим частные случаи аффинной геометрии.

Определение. При попытке объединить гравитацию с электромагнетизмом Г. Вейль рассмотрел тензор неметричности специального вида [**?**]

$$Q_{\alpha\beta\gamma} = W_\alpha g_{\beta\gamma}, \tag{7.18}$$

где W_α – *форма Вейля*, отождествленная с электромагнитным потенциалом, (при этом предполагалось, что кручение тождественно равно нулю). Будем говорить, что на многообразии задана *геометрия Римана–Картана–Вейля*, если на нем задана метрика, кручение и неметричность специального вида (7.18).

Если тензор неметричности тождественно равен нулю, а метрика и кручение нетривиальны, то будем говорить, что на многообразии задана *геометрия Римана–Картана.* В этом случае из (7.15) следует, что аффинная связность однозначно определяется метрикой и кручением

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}(\partial_\alpha g_{\beta\gamma} + \partial_\beta g_{\alpha\gamma} - \partial_\gamma g_{\alpha\beta}) + \frac{1}{2}(T_{\alpha\beta\gamma} - T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}). \tag{7.19}$$

Такую связность называют *метрической*, поскольку ковариантная производная от метрики тождественно равна нулю,

$$\nabla_\alpha g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} - \Gamma_{\alpha\beta}{}^{\delta} g_{\delta\gamma} - \Gamma_{\alpha\gamma}{}^{\delta} g_{\beta\delta} = 0. \tag{7.20}$$

Это равенство называют *условием метричности*, и оно обеспечивает коммутативность ковариантного дифференцирования с опусканием и подъемом индексов. Условие метричности (7.20) эквивалентно условию $\nabla_\alpha g^{\beta\gamma} = 0$ в силу (7.17).

Если тензор кручения равен нулю, $T_{\alpha\beta}{}^{\gamma} = 0$, а неметричность имеет вид (7.18), то будем говорить, что задана геометрия *Римана–Вейля.* В геометрии Римана–Вейля выражение для аффинной связности можно записать в следующем виде:

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}\left[(\partial_\alpha + W_\alpha)g_{\beta\gamma} + (\partial_\beta + W_\beta)g_{\alpha\gamma} - (\partial_\gamma + W_\gamma)g_{\alpha\beta}\right]. \tag{7.21}$$

При этом она симметрична по двум первым индексам.

Если и тензор неметричности, и кручение тождественно равны нулю, а метрика нетривиальна и положительно определена, то будем говорить, что на многообразии задана *геометрия Римана*. В этом случае метрическая связность также симметрична по двум первым индексам и однозначно определяется метрикой

$$\widetilde{\Gamma}_{\alpha\beta}{}^{\gamma} = \frac{1}{2} g^{\gamma\delta}(\partial_\alpha g_{\beta\delta} + \partial_\beta g_{\alpha\delta} - \partial_\delta g_{\alpha\beta}) \tag{7.22}$$

Эта связность называется *связностью Леви–Чивита* или *символами Кристоффеля*. Заметим, что в определении метрики мы требовали невырожденность матрицы $g_{\alpha\beta}$. Из выражения для символов Кристоффеля видно, что это условие необходимо, в частности, для того, чтобы метрика определяла связность Леви–Чивиты. Если метрика не является положительно определенной, то геометрия называется *псевдоримановой*. Для связности при этом сохраняются прежние названия. □

Замечание. Если на многообразии задана аффинная геометрия общего вида $(\mathbb{M}, g, \Gamma)$, то на нем определены две связности: аффинная связность Γ и связность Леви–Чивиты $\widetilde{\Gamma}$, поскольку задана метрика. В такой ситуации над связностью Леви–Чивиты и построенных с ее помощью геометрических объектах мы будем писать знак тильды. □

Пример 7.2.2. Связность Леви–Чивита имеет наглядный геометрический смысл для двумерных поверхностей, изометрически вложенных в трехмерное евклидово пространство. Рассмотрим два вектора X_x и Y_y, касательных к поверхности в близких точках x и y. Леви-Чивита [**?**] предложил считать вектор Y_y параллельным вектору X_x, если его проекция в евклидовом пространстве на касательную плоскость в точке x параллельна вектору X_x. Можно проверить, что такая связность согласована с метрикой, индуцированной вложением, и имеет нулевое кручение. □

В римановой геометрии ковариантная производная определяется компонентами связности с двумя нижними и одним верхним индексом, $\widetilde{\Gamma}_{\alpha\beta}{}^{\gamma} = \widetilde{\Gamma}_{\alpha\beta\delta} g^{\gamma\delta}$ и поэтому одинакова для двух метрик, отличающихся постоянным множителем.

Символы Кристоффеля (7.22) симметричны по первым двум индексам. Это свойство выполняется в голономном базисе. В неголономном базисе эта симметрия нарушается (см. раздел 7.9).

Из выражения для символов Кристоффеля (7.22) или условия метричности (7.20) следует, что

$$\partial_\alpha g_{\beta\gamma} = \widetilde{\Gamma}_{\alpha\beta\gamma} + \widetilde{\Gamma}_{\alpha\gamma\beta}. \tag{7.23}$$

Это значит, что для того, чтобы символы Кристоффеля в некоторой системе координат были равны нулю, необходимо и достаточно, чтобы в этой системе координат компоненты метрики были постоянны. Поскольку символы Кристоффеля не являются компонентами тензора, то в другой системе координат они могут быть нетривиальны.

Пример 7.2.3. Символы Кристоффеля для евклидова пространства в декартовой системе координат равны нулю, но, скажем, в сферической или цилиндрической системе координат они отличны от нуля. □

В случае, когда тензор неметричности и кручение тождественно равны нулю, и при этом в окрестности любой точки существует система координат, в которой мет-

рика является единичной диагональной матрицей, и, следовательно, символы Кристоффеля равны нулю, то геометрия называется *локально евклидовой* (или *псевдоевклидовой*, если часть единиц входят в метрику с отрицательным знаком). Соответствующая система координат называется *декартовой*.

Пример 7.2.4. Рассмотрим двумерный тор, вложенный в трехмерное евклидово пространство. Если предположить, что кручение и неметричность на торе равны нулю, а метрика индуцирована вложением, то тензор кривизны тора равен нулю. В этом случае тор представляет собой нетривиальное локально евклидово многообразие. □

Нетрудно посчитать число независимых компонент у связности, тензора кручения и неметричности:

$$[\Gamma_{\alpha\beta}{}^{\gamma}] = n^3, \qquad [T_{\alpha\beta}{}^{\gamma}] = \frac{n^2(n-1)}{2}, \qquad [Q_{\alpha\beta\gamma}] = \frac{n^2(n+1)}{2}.$$

Отсюда следует, что суммарное число независимых компонент кручения и неметричности равно числу компонент аффинной связности.

7.3 Ковариантная производная тензорных плотностей

Ковариантная производная тензорного поля естественным образом обобщается на тензорные плотности произвольной степени p. Для определенности рассмотрим тензорную плотность $X_{\alpha}{}^{\beta}$ степени p и типа $(1,1)$. По-определению, при преобразовании координат она преобразуется по-правилу

$$X_{\alpha'}{}^{\beta'} = J^p \frac{dx^{\alpha}}{dx^{\alpha'}} X_{\alpha}{}^{\beta} \frac{dx^{\beta'}}{dx^{\beta}},$$

где J – якобиан преобразования (2.58). Потребуем, чтобы ковариантная производная от тензорной плотности была тензорной плотностью той же степени,

$$\nabla_{\alpha'} X_{\beta'}{}^{\gamma'} = J^p \frac{dx^{\alpha}}{dx^{\alpha'}} \frac{dx^{\beta}}{dx^{\beta'}} \nabla_{\alpha} X_{\beta}{}^{\gamma} \frac{dx^{\gamma'}}{dx^{\gamma}}. \tag{7.24}$$

Отсюда следует, что ковариантная производная от тензорной плотности имеет вид

$$\nabla_{\alpha} X_{\beta}{}^{\gamma} = \partial_{\alpha} X_{\beta}{}^{\gamma} - \Gamma_{\alpha\beta}{}^{\delta} X_{\delta}{}^{\gamma} + X_{\beta}{}^{\delta} \Gamma_{\alpha\delta}{}^{\gamma} + p \Gamma_{\alpha} X_{\beta}{}^{\gamma}, \tag{7.25}$$

где $\Gamma_{\alpha} = \Gamma_{\alpha\beta}{}^{\beta}$ – след аффинной связности. Сравнивая эту ковариантную производную с ковариантной производной (7.6) от тензорного поля типа $(1,1)$, находим, что различие состоит в появлении дополнительного слагаемого, пропорционального степени тензорной плотности и следу аффинной связности. Чтобы проверить выполнение закона преобразования (7.24) заметим, что в законе преобразования следа аффинной связности (7.44) содержится неоднородное слагаемое, которое компенсирует производную от якобиана

$$\partial_{\alpha} J^p = p J^p \frac{dx^{\beta}}{dx^{\beta'}} \partial_{\alpha} \left(\frac{dx^{\beta'}}{dx^{\beta}} \right).$$

Отсюда следует, что ковариантная производная от тензорной плотности произвольного типа отличается от ковариантной производной соответствующего тензора одним дополнительным слагаемым, пропорциональным степени тензорной плотности и следу аффинной связности.

Определение. Параллельный перенос для тензорных плотностей определяется так же, как и для тензоров. Мы говорим, что тензорная плотность *параллельно переносится* вдоль заданной кривой, если ее ковариантная производная вдоль этой кривой равна нулю. □

Ковариантная производная от линейной комбинации тензорных плотностей X и Y одинакового ранга и степени равна линейной комбинации ковариантных производных

$$\nabla_\alpha(aX + bY) = a\nabla_\alpha X + b\nabla_\alpha Y, \qquad a, b \in \mathbb{R}.$$

Естественным образом определяется тензорное произведение $X \otimes Y$ тензорных плотностей произвольных рангов и степеней. При этом степени тензорных плотностей складываются. Нетрудно проверить, что для ковариантного дифференцирования справедливо правило Лейбница

$$\nabla_\alpha(X \otimes Y) = (\nabla_\alpha X) \otimes Y + X \otimes (\nabla_\alpha Y).$$

Так же, как и для тензоров, свертка по индексам тензорных плотностей перестановочна с операцией ковариантного дифференцирования.

Пример 7.3.1. Рассмотрим частные случаи. Определители метрики g и репера $\sqrt{|g|}$ являются тензорными плотностями степени -2 и -1, соответственно. Используя выражение для следа аффинной связности (7.45), получаем

$$\nabla_\alpha g = \partial_\alpha g - 2\Gamma_\alpha g = -Q_\alpha\sqrt{|g|}, \tag{7.26}$$

$$\nabla_\alpha\sqrt{|g|} = \partial_\alpha\sqrt{|g|} - \Gamma_\alpha\sqrt{|g|} = -\frac{1}{2}Q_\alpha\sqrt{|g|}. \tag{7.27}$$

Отсюда следует, что для метрической связности определители метрики и репера ковариантно постоянны. То есть в геометрии Римана–Картана и, в частности, в римановой геометрии имеем

$$\widetilde{\nabla}_\alpha g = 0, \qquad \widetilde{\nabla}_\alpha\sqrt{|g|} = 0.$$

При этом условие метричности для связности является достаточным, но не необходимым. Из выражения (7.27) следует, что для сохранения элемента объема при параллельном переносе вдоль произвольной кривой, необходимо и достаточно, чтобы след тензора неметричности был равен нулю. В геометрии Римана–Картана–Вейля

$$\nabla_\alpha\sqrt{|g|} = -\frac{n}{2}W_\alpha\sqrt{|g|},$$

где n – размерность многообразия. То есть форма Вейля в этом случае определяет изменение элемента объема при параллельном переносе. □

7.4 Параллельный перенос

В разделе 6.2 был определен параллельный перенос векторов. С помощью аффинной связности на многообразии $\mathbb{M}$ можно определить *параллельный перенос* касательных векторов, а также тензоров произвольного ранга вдоль кривой.

Определение. Пусть дифференцируемая кривая $\gamma = x(t) = \{x^\alpha(t)\}$, $t \in [0,1]$ соединяет две точки многообразия $p, q \in \mathbb{M}$: $x(0) = p$, $x(1) = q$. Касательный вектор к кривой (вектор скорости) имеет компоненты $u^\alpha = \dot{x}^\alpha$ и предполагается отличным от нуля. Инвариантное условие параллельного переноса (6.24),

$$\nabla_u X = 0, \tag{7.28}$$

в компонентах (6.25) принимает вид

$$\dot{X}^\alpha = -\dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}{}^\alpha. \tag{7.29}$$

Это – система линейных дифференциальных уравнений на компоненты векторного поля $X^\alpha(t)$, которая решается с некоторым начальным условием $X^\alpha(0) = X_0^\alpha$. Решение этой задачи мы называем *параллельным переносом* вектора X_0 в точку $x(t)$ вдоль кривой γ. Результат параллельного переноса вектора вдоль кривой однозначен в силу единственности решения системы обыкновенных дифференциальных уравнений с (7.29) с заданными начальными условиями. □

Условие параллельного переноса (7.29) можно переписать в интегральной форме

$$X^\alpha(t) = X_0^\alpha - \int_0^t ds\, \dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}{}^\alpha, \tag{7.30}$$

где все функции в подынтегральном выражении рассматриваются, как функции от параметра s вдоль кривой и точка обозначает дифференцирование по s. При $t = 1$ получим компоненты вектора в точке q.

Результат параллельного переноса вектора из точки p в точку q не зависит ни от параметризации кривой, что очевидно, ни от выбора системы координат. Действительно, в новой штрихованной системе координат справедливо равенство

$$X^{\alpha'}(t) = X_0^{\alpha'} - \int_0^t ds\, \dot{x}^{\beta'} X^{\gamma'} \Gamma_{\beta'\gamma'}{}^{\alpha'}.$$

Это интегральное уравнение связано с уравнением (7.30) преобразованием координат. Чтобы доказать это, представим неоднородное слагаемое в преобразовании компонент аффинной связности (7.5), умноженное на $\dot{x}^{\beta'} X^{\gamma'}$, в виде

$$\dot{x}^\beta X^\gamma \partial^2_{\beta\gamma} x^{\alpha'} = X^\gamma \partial_t (\partial_\gamma x^{\alpha'})$$

и проинтегрируем по частям.

В то же время результат параллельного переноса вектора из точки p в точку q в общем случае зависит от кривой γ, соединяющей эти точки.

Параллельный перенос вектора вдоль кривой можно естественным образом обобщить на кусочно гладкие кривые, как последовательный параллельный перенос от одного излома к другому.

Аналогично определяется параллельный перенос 1-форм и тензоров произвольного ранга.

При параллельном переносе вектора из точки x в близкую точку $x + dx$, где инфинитезимальное приращение dx^α мы рассматриваем, как касательный вектор к кривой в точке x, компоненты вектора получат приращение

$$\delta X^\alpha = -dx^\beta X^\gamma \Gamma_{\beta\gamma}{}^\alpha. \tag{7.31}$$

Эти приращения линейны по дифференциалам dx^α и самому векторному полю. Формула (7.31) позволяет дать следующую интерпретацию ковариантной производной (7.2). Пусть на многообразии задано произвольное векторное поле $X(x)$. Чтобы получить ковариантную производную от него в точке x, необходимо взять значение векторного поля в точке $x + dx$, вычесть из него результат параллельного переноса вектора из точки x в точку $x + dx$ и разделить эту разность на dx.

Пример 7.4.1. При параллельном переносе скаляра получаем постоянное скалярное поле вдоль кривой $\delta f = 0$. Результат параллельного переноса скаляра для любой аффинной связности не зависит от пути переноса □

Пример 7.4.2. При параллельном переносе 1-форма получает приращение

$$\delta A_\alpha = dx^\beta \Gamma_{\beta\alpha}{}^\gamma A_\gamma, \tag{7.32}$$

которое отличается знаком от приращения вектора (7.31). □

Пример 7.4.3. При параллельном переносе тензора типа $(1,1)$ он получает приращение

$$\delta T^\alpha{}_\beta = dx^\gamma(-\Gamma_{\gamma\delta}{}^\alpha T^\delta{}_\beta + \Gamma_{\gamma\beta}{}^\delta T^\alpha{}_\delta).$$

Аналогично для тензора произвольного ранга в правой части будем иметь по одному слагаемому со знаком минус и плюс для каждого контра- и ковариантного индекса соответственно. □

Пример 7.4.4. В геометрии Римана–Картана, по-определению, ковариантная производная от метрики равна нулю. Поэтому можно считать, что метрика получается в результате параллельного переноса симметричной невырожденной матрицы из произвольной точки многообразия. При этом результат параллельного переноса заданной матрицы не зависит от пути для любой метрической связности. □

Пусть на многообразии $\mathbb{M}$ помимо аффинной связности задана метрика. Рассмотрим, как меняется скалярное произведение (X, Y) двух векторов при параллельном переносе вдоль произвольной кривой γ.

Предложение 7.4.1. *Зависимость скалярного произведения (X, Y) двух векторов, которые параллельно переносятся вдоль γ, от точки кривой определяется только тензором неметричности:*

$$\partial_u(X, Y) = \nabla_u(X^\alpha Y^\beta g_{\alpha\beta}) = -u^\gamma X^\alpha Y^\beta Q_{\gamma\alpha\beta}. \tag{7.33}$$

Отсюда следует, что в геометрии Римана–Картана ($Q = 0$) скалярное произведение двух векторов при параллельном переносе вдоль произвольной кривой сохраняется. В частности, квадрат вектора скорости u^2 кривой γ в геометрии Римана–Картана постоянен вдоль нее.

Доказательство. Следует из определения тензора неметричности (7.14). □

Следствие. В римановой геометрии и геометрии Римана–Картана длины векторов и углы между ними сохраняются при параллельном переносе вдоль произвольной кривой γ.

Задание аффинной связности позволяет сравнивать компоненты тензоров в бесконечно близких точках, причем делать это ковариантным образом. Трудность сравнения тензоров в различных точках связана с тем, что при преобразовании координат тензоры в разных точках преобразуются по-разному, и их сравнение (сложение компонент) теряет всякий смысл.

Выберем точку многообразия $p = \{p^\alpha\} \in \mathbb{M}$ и замкнутую кривую γ_p с началом и концом в точке p. Возьмем произвольный вектор X_p в точке p и перенесем его параллельно вдоль кривой γ_p. В результате получим новый вектор $X_p + \triangle X_p$, который может отличаться от исходного. Рассмотрим класс замкнутых кривых $\gamma_p = \{x^\alpha(t)\}$ малой "длины", для которых выполнено неравенство

$$\oint_{\gamma_p} dt \sqrt{(\dot{x}^1)^2 + \ldots + (\dot{x}^n)^2} < \epsilon \ll 1. \tag{7.34}$$

Это выражение нековариантно, т.е. мы работаем в фиксированной системе координат некоторой координатной окрестности. Тогда все рассматриваемые кривые заведомо лежат в малой окрестности точки $p \in \mathbb{U}_{p,\epsilon} \subset \mathbb{M}$, координаты точек которой удовлетворяют неравенствам:

$$\mathbb{U}_{p,\epsilon} := \{x \in \mathbb{M} : \quad |x^\alpha - p^\alpha| < \epsilon,\ \alpha = 1, \ldots, n\}.$$

Мы также предполагаем, что окрестность $\mathbb{U}_{p,\epsilon}$ односвязна. Рассмотрим двумерную поверхность $\mathbb{S}_\gamma \subset \mathbb{M}$ в многообразии $\mathbb{M}$, с границей γ_p и целиком лежащую в $\mathbb{U}_{p,\epsilon}$, которая задана параметрически: $\{x^\alpha(u, v)\}$, где u, v – координаты на поверхности.

Теорема 7.4.1 (Геометрический смысл кривизны). *Пусть на многообразии $\mathbb{M}$ в некоторой координатной окрестности задана аффинная связность $\Gamma_{\alpha\beta}{}^\gamma$ с тензором кривизны $R_{\alpha\beta\gamma}{}^\delta$. Перенесем вектор X_p вдоль замкнутой кривой γ_p, для которой выполнено неравенство (7.34). Тогда компоненты вектора получат приращение порядка не выше ϵ^2:*

$$\triangle X_p^\delta = -X_p^\gamma R_{\alpha\beta\gamma}{}^\delta(p) \int_{\mathbb{S}_\gamma} du \wedge dv \frac{\partial x^\alpha}{\partial u} \frac{\partial x^\beta}{\partial v} + \mathsf{o}(\epsilon^2), \tag{7.35}$$

где $\mathbb{S}_\gamma \subset \mathbb{U}_{p,\epsilon}$ – поверхность, ограниченная кривой γ_p.

Доказательство. При параллельном переносе вектора вдоль замкнутого пути приращение его компонент задается интегралом (7.30)

$$\triangle X_p^\alpha = -\oint_{\gamma_p} dt\, \dot{x}^\beta X_\parallel^\gamma \Gamma_{\beta\gamma}{}^\alpha = -\oint_{\gamma_p} dx^\beta X_\parallel^\gamma \Gamma_{\beta\gamma}{}^\alpha, \tag{7.36}$$

где $X_\parallel^\gamma(t)$ – результат параллельного переноса вектора X_p вдоль кривой γ до точки $x(t)$. Разложим подынтегральную функцию в ряд Тейлора

$$X_\parallel^\gamma \Gamma_{\beta\gamma}{}^\alpha = (X_\parallel^\gamma \Gamma_{\beta\gamma}{}^\alpha)_p + \left[\partial_\delta (X_\parallel^\gamma \Gamma_{\beta\gamma}{}^\alpha)\right]_p (x^\delta - p^\delta) + \ldots$$

Интеграл от первого слагаемого равен нулю, т.к. кривая замкнута, и, поэтому порядок приращения вектора при параллельном переносе не может превышать ϵ^2. Это

значит, что при вычислении интеграла в (7.36) с точностью ϵ^2 в подынтегральном выражении достаточно учитывать члены порядка ϵ.

Разнесем вектор X_p из точки p на всю окрестность $\mathbb{U}_{p,\epsilon}$ с помощью следующего соотношения

$$X^\alpha(x) := X^\alpha_p - (x^\beta - p^\beta) X^\gamma_p \Gamma_{\beta\gamma}{}^\alpha(p). \tag{7.37}$$

Это – параллельный перенос вектора X_p на окрестность $\mathbb{U}_{p,\epsilon}$ с точностью ϵ. С указанной точностью параллельный перенос вектора X_p не зависит от пути, так как приращение компонент вектора по замкнутому пути порядка ϵ^2. В результате мы получаем гладкое векторное поле $X(x)$ на $\mathbb{U}_{p,\epsilon}$. Поэтому в подынтегральном выражении (7.36) можно заменить векторное поле вдоль кривой $X^\alpha_\parallel$, которое в общем случае не является непрерывным, на гладкое векторное поле X. Тогда можно воспользоваться формулой Грина (4.82):

$$\oint_{\gamma_p} dt\, \dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}{}^\alpha = \int_{\mathbb{S}_\gamma} du \wedge dv \left[\frac{\partial}{\partial u} \left(\frac{\partial x^\beta}{\partial v} X^\gamma \Gamma_{\beta\gamma}{}^\alpha \right) - \frac{\partial}{\partial v} \left(\frac{\partial x^\beta}{\partial u} X^\gamma \Gamma_{\beta\gamma}{}^\alpha \right) \right].$$

Для векторного поля (7.37) выполнено равенство

$$\frac{\partial X^\gamma}{\partial u} = -\frac{\partial x^\delta}{\partial u} X^\beta_p \Gamma_{\delta\beta}{}^\gamma(p)$$

и аналогичную формулу для частной производной по v. В результате с точностью ϵ^2 получаем формулу (7.35). □

Замечание. С точностью ϵ^2 приращение компонент вектора (7.35) ковариантно. По построению, это приращение не зависит от выбора поверхности $\mathbb{S}_\gamma \subset \mathbb{U}_{p,\epsilon}$, натянутой на контур γ_p. □

Пример 7.4.5. Рассмотрим "параллелограмм" γ_p со сторонами $\epsilon^\alpha_1, \epsilon^\alpha_2 < \epsilon$ и вершинами в точках $\{p^\alpha\}$, $\{p^\alpha + \epsilon^\alpha_1\}$, $\{p^\alpha + \epsilon^\alpha_1 + \epsilon^\alpha_2\}$ и $\{p^\alpha + \epsilon^\alpha_2\}$. Пусть на него натянута поверхность $\mathbb{S}_\gamma = \{p^\alpha + \epsilon^\alpha_1 u + \epsilon^\alpha_2 v\}$; $u, v \in [0,1]$. Тогда при обходе по "параллелограмму" вектор получит приращение

$$\triangle X^\delta_p = -\epsilon^\alpha_1 \epsilon^\beta_2 X^\gamma_p R_{\alpha\beta\gamma}{}^\delta(p),$$

что следует из (7.35). См. рис.7.1. □

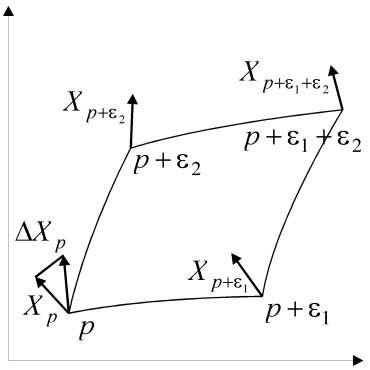


Рис. 7.1: Геометрический смысл кривизны: при параллельном переносе вектора X_p вдоль замкнутого контура он получает приращение $\triangle X_p$.

Замечание. Теорема о геометрическом смысле кривизны не зависит от того задана на многообразии $\mathbb{M}$ метрика или нет, т.к. параллельный перенос определяется только связностью. □

Пусть на многообразии $\mathbb{M}$ задана метрика $g_{\alpha\beta}$. Тогда на поверхности $\mathbb{S}_\gamma$ индуцируется метрика $g_{ij} = \partial_i x^\alpha \partial_j x^\beta$ ($i,j = 1,2$, где $\{u^i\} = (u,v)$). Следовательно, определена форма объема (площади поверхности) (4.72)

$$\upsilon = \frac{1}{2} du^i \wedge du^j \varepsilon_{ij}.$$

Тогда интеграл в (7.35) можно переписать в виде

$$\frac{1}{2}\int_{\mathbb{S}_\gamma} du^i \wedge du^j \frac{\partial x^\alpha}{\partial u^i}\frac{\partial x^\beta}{\partial u^j} = \frac{1}{2}\int_{\mathbb{S}_\gamma} dudv \sqrt{|g|}\varepsilon^{ij}\frac{\partial x^\alpha}{\partial u^i}\frac{\partial x^\beta}{\partial u^j}\,\mathsf{sgn}\,,$$

где $\mathsf{sgn} = \mathsf{sign}\,(\det g_{ij})$ и $\sqrt{|g|} = \sqrt{|\det g_{ij}|}$. С точностью ϵ^2 подынтегральное выражение можно вынести из под знака интегрирования и формула для приращения компонент вектора при параллельном переносе вдоль малого замкнутого пути (7.35) принимает вид

$$\triangle X^\delta_p = -\frac{1}{\sqrt{|g_p|}}\left(\frac{\partial x^\alpha}{\partial u}\frac{\partial x^\beta}{\partial v}\right)_p X^\gamma_p R_{\alpha\beta\gamma}{}^\delta(p) S, \tag{7.38}$$

где

$$S = \int_{\mathbb{S}_\gamma} dudv\sqrt{|g|}$$

– площадь поверхности $\mathbb{S}_\gamma$. Полученная формула ковариантна относительно преобразований координат на многообразии x^α и инвариантна относительно выбора координат u, v на поверхности.

Замечание. Геометрический смысл тензора кривизны лег в основу его физической интерпретации, как поверхностной плотности вектора Франка, характеризующего распределение дисклинаций в упругой среде со спиновой структурой (см. [**?**]). □

Формулы переноса вектора вдоль малого замкнутого контура без труда обобщаются на произвольные тензорные поля.

Пример 7.4.6. Рассмотрим такой же "параллелограмм" со сторонами ϵ_1 и ϵ_2 как и в примере 7.4.5. При параллельном переносе вдоль него компоненты тензора T типа $(1,1)$ получает приращение

$$\triangle T^{\ \alpha}_{p\beta} = -\epsilon_1^\gamma \epsilon_2^\delta \Big(R_{\gamma\delta\epsilon}{}^\alpha(p) T_p{}^\epsilon{}_\beta - R_{\gamma\delta\beta}{}^\epsilon(p) T_p{}^\alpha{}_\epsilon \Big).$$

Для тензора произвольного ранга в правой части равенства будем иметь по одному слагаемому со знаками плюс и минус соответственно для каждого контравариантного и ковариантного индекса. □

7.5 Геометрический смысл кручения

Рассмотрим многообразие $\mathbb{M}$ с заданной аффинной связностью. Пусть $\mathbb{U} \subset \mathbb{M}$ – координатная окрестность и $\Gamma_{\alpha\beta}{}^\gamma$ – компоненты связности. Пусть в произвольной точке $p \in \mathbb{U}$ заданы два касательных вектора $X_p = \{X^\alpha_p\}$ и $Y_p = \{Y^\alpha_p\}$, рис. 7.2. Чтобы дать

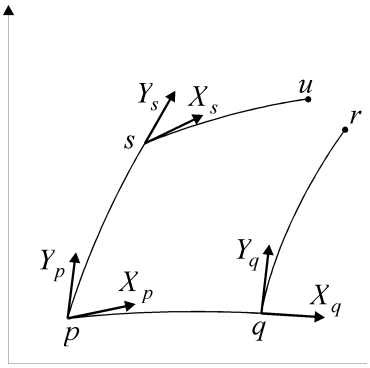


Рис. 7.2: Геометрический смысл кручения: "параллелограмм" с направляющими $X_p\epsilon$ и $Y_p\epsilon$ разомкнут.

геометрическую интерпретацию тензору кручения, произведем следующее построение. Перенесем вектор Y_p вдоль геодезической линии, которая касается вектора X_p, в близкую точку q, соответствующую параметру $t_q = t_p + \epsilon$, $\epsilon \ll 0$. Затем проведем геодезическую в точке q, которая касается вектора Y_q, и отметим на ней точку r, соответствующую параметру $t_q + \epsilon$. Затем проведем то же построение, но в обратном порядке: перенесем вектор X_p вдоль геодезической для Y_p в точку s, соответствующую параметру $t_p + \epsilon$, выпустим из нее геодезическую вдоль X_s и отметим точку u, соответствующую параметру $t_s + \epsilon$. Если кручение аффинной связности равно нулю, то точки r и u совпадут. При ненулевом кручении разность координат точек r и u имеет порядок ϵ^2 и, как будет показано, определяется тензором кручения.

Начнем построение. Рассмотрим геодезическую линию $\{x^\alpha(t)\}$ в $\mathbb{M}$, которая проходит через p и касается вектора X_p. Перенесем вектор X_p вдоль геодезической. В результата получим векторное поле $X(t)$ на геодезической, компоненты которого, по-построению, удовлетворяет уравнению

$$\nabla_{\dot{x}} X^\alpha = \frac{DX^\alpha}{dt} = \dot{x}^\beta(\partial_\beta X^\alpha + X^\gamma \Gamma_{\beta\gamma}{}^\alpha) = \dot{X}^\alpha + \dot{x}^\beta X^\gamma \Gamma_{\beta\gamma}{}^\alpha = 0, \tag{7.39}$$

где DX – ковариантный дифференциал векторного поля (7.7) и t – произвольный параметр вдоль геодезической. Тогда геодезическая будет интегральной кривой полученного векторного поля, проходящей через точку p.

Замечание. Параметр t может не совпадать с каноническим параметром вдоль геодезической, а векторное поле $X(t)$ – с полем скорости, соответствующим каноническому параметру. □

Геодезическая удовлетворят уравнению с начальными условиями:

$$\dot{x}^\alpha|_{t=t_p} = X^\alpha, \qquad x^\alpha|_{t=t_p} = x_p^\alpha. \tag{7.40}$$

Для наших целей вычисления достаточно провести во втором порядке по ϵ. Ниже мы будем писать знак равенства, понимая его с точностью $\mathsf{o}(\epsilon^2)$. В этом порядке координаты точки q равны

$$x_q^\alpha = x_p^\alpha + X_p^\alpha \epsilon + \frac{1}{2}\dot{X}_p^\alpha \epsilon^2 = x_p^\alpha + X_p^\alpha \epsilon - \frac{1}{2} X_p^\beta X_p^\gamma \Gamma_{p\,\beta\gamma}{}^\alpha \epsilon^2,$$

где мы учли уравнения (7.39) и (7.40). Теперь параллельно перенесем вектор Y_p в точку q:

$$Y_q^\alpha = Y_p^\alpha - (x_q^\beta - x_p^{\text{в}}) Y_p^\gamma \Gamma_{p\beta\gamma}{}^\alpha = Y_p^\alpha - X_p^\beta Y_p^\gamma \Gamma_{p\,\beta\gamma}{}^\alpha \epsilon,$$

где достаточно учесть только линейный член разложения. Теперь найдем координаты точки r вдоль геодезической, выпущенной из точки q:

$$\begin{aligned} x_r^\alpha &= x_q^\alpha + Y_q^\alpha\epsilon - \frac{1}{2}Y_q^\beta Y_q^\gamma \Gamma_{q\,\beta\gamma}{}^\alpha \epsilon^2 = \\ &= x_p^\alpha + (X_p^\alpha + Y_p^\beta)\epsilon^2 - \left(\frac{1}{2}X_p^\beta X_p^\gamma + X_p^\beta Y_p^\gamma + \frac{1}{2}Y_p^\beta Y_p^\gamma\right)\Gamma_{p\,\beta\gamma}{}^\alpha\epsilon^2. \end{aligned} \tag{7.41}$$

Аналогично, координаты точки u во втором порядке по ϵ равны

$$x_u^\alpha = x_p^\alpha + (X_p^\alpha + Y_p^\beta)\epsilon^2 - \left(\frac{1}{2}X_p^\beta X_p^\gamma + Y_p^\beta X_p^\gamma + \frac{1}{2}Y_p^\beta Y_p^\gamma\right)\Gamma_{p\,\beta\gamma}{}^\alpha\epsilon^2.$$

Таким образом получаем равенство

$$x_u^\alpha - x_r^\alpha = X_p^\beta Y_p^\gamma T_{p\,\beta\gamma}{}^\alpha \epsilon^2, \tag{7.42}$$

где $T_{p\,\beta\gamma}{}^\alpha$ – компоненты тензора кручения в точке p.

Формула (7.42) позволяет дать геометрическую интерпретацию тензору кручения: главная часть разомкнутости бесконечно малого "параллелограмма" с направляющими $X_p\epsilon$ и $Y_p\epsilon$ квадратична по ϵ и определяется тензором кручения.

Замечание. Геометрический смысл кручения лег в основу физической интерпретации кручения, как поверхностной плотности вектора Бюргерса, характеризующего дислокации в упругой среде (см. [**?**]). □

7.6 Свойства аффинной связности

Во многих важных моделях математической физики пространство-время рассматривается, как многообразие $\mathbb{M}$, $\dim\mathbb{M} = n$, на котором задана аффинная геометрия, т.е. метрика и аффинная связность. В настоящем разделе приведены тождества, включающие аффинную связность, которые полезны для приложений при проведении вычислений. Кроме того, определены инвариантные дифференциальные операторы второго порядка и приведена формула интегрирования по частям. Все формулы настоящего раздела доказываются прямыми вычислениями.

Определим *след аффинной связности*, который получается после свертки последней пары индексов:

$$\Gamma_\alpha := \Gamma_{\alpha\beta}{}^\beta. \tag{7.43}$$

Он не является ковекторным полем, поскольку закон преобразования содержит неоднородное слагаемое:

$$\Gamma_{\alpha'} = \partial_{\alpha'}x^\alpha\Gamma_\alpha - \partial_{\alpha'}x^\alpha\partial_{\beta'}x^\beta\partial^2_{\alpha\beta}x^{\beta'} = \partial_{\alpha'}x^\alpha\Gamma_\alpha + \partial^2_{\alpha'\beta'}x^\alpha\partial_\alpha x^{\beta'}. \tag{7.44}$$

Из выражения (7.15) для компонент аффинной связности следует, что след аффинной связности (7.43) равен

$$\Gamma_\alpha = \widetilde{\Gamma}_\alpha + \frac{1}{2}Q_\alpha, \tag{7.45}$$

где след символов Кристоффеля, $\widetilde{\Gamma}_\alpha := \widetilde{\Gamma}_{\alpha\beta}{}^\beta = \widetilde{\Gamma}_{\beta\alpha}{}^\beta$, имеет вид

$$\widetilde{\Gamma}_\alpha = \frac{1}{2}g^{\beta\gamma}\partial_\alpha g_{\beta\gamma} = -\frac{1}{2}g_{\beta\gamma}\partial_\alpha g^{\beta\gamma} = \frac{1}{2}\frac{\partial_\alpha g}{g} = \frac{\partial_\alpha\sqrt{|g|}}{\sqrt{|g|}} = e^\beta{}_b\partial_\alpha e_\beta{}^b, \tag{7.46}$$

и введен след тензора неметричности по последним индексам:

$$Q_\alpha := Q_{\alpha\beta}{}^\beta. \tag{7.47}$$

В тождествах (7.46)

$$g := \det g_{\alpha\beta}, \qquad \sqrt{|g|} = \det e_\alpha{}^a.$$

Замечание. След аффинной связности не зависит от кручения. □

В пространстве Римана–Картана–Вейля след неметричности пропорционален форме Вейля W_α из (7.18):

$$Q_\alpha = nW_\alpha.$$

Свертка аффинной связности (7.15) по первому и третьему индексам дает

$$\Gamma_{\beta\alpha}{}^\beta = \widetilde{\Gamma}_\alpha + T_\alpha + \frac{1}{2}Q_\alpha = \Gamma_\alpha + T_\alpha, \tag{7.48}$$

где введен след тензора кручения

$$T_\alpha := T_{\beta\alpha}{}^\beta. \tag{7.49}$$

Предложение 7.6.1. *Справедливо равенство*

$$\nabla_\alpha \varepsilon_{\alpha_1\ldots\alpha_n} = -\frac{1}{2}Q_\alpha \varepsilon_{\alpha_1\ldots\alpha_n}. \tag{7.50}$$

Доказательство. Из определения ковариантной производной и связи между полностью антисимметричным тензором и тензорной плотностью (3.62) следует равенство

$$\nabla_\alpha \varepsilon_{\alpha_1\ldots\alpha_n} = \hat{\varepsilon}_{\alpha_1\ldots\alpha_n}\partial_\alpha\sqrt{|g|} - \Gamma_{\alpha\alpha_1}{}^\beta \varepsilon_{\beta\alpha_2\ldots\alpha_n} - \ldots - \Gamma_{\alpha\alpha_n}{}^\beta \varepsilon_{\alpha_1\ldots\alpha_{n-1}\beta}.$$

Это выражение антисимметрично по индексам $\alpha_1, \ldots, \alpha_n$ и, следовательно, пропорционально полностью антисимметричному тензору:

$$\nabla_\alpha \varepsilon_{\alpha_1\ldots\alpha_n} = \varepsilon_{\alpha_1\ldots\alpha_n} X_\alpha,$$

где X_α – компоненты некоторого ковектора. Свертка полученного равенства с контравариантным тензором $\varepsilon^{\alpha_1\ldots\alpha_n}$ с учетом равенств (7.45), (7.46) и (28.52) определяет

$$X_\alpha = -\frac{1}{2}Q_\alpha.$$ □

Следствие. Для метрической связности полностью антисимметричный тензор является ковариантно постоянным. □

В дифференциальной геометрии естественным образом вводится понятие градиента и дивергенции. Назовем *градиентом* скалярного поля (функции) f ковекторное поле (1-форму):

$$\mathsf{grad}\, f = dx^\alpha \nabla_\alpha f = df = dx^\alpha \partial_\alpha f. \tag{7.51}$$

Назовем *дивергенцией* векторного поля скалярное поле

$$\mathsf{div}\, X = \nabla_\alpha X^\alpha = \partial_\alpha X^\alpha + \Gamma_{\beta\alpha}{}^\beta X^\alpha, \tag{7.52}$$

построенное с помощью ковариантной производной. С учетом выражения аффинной связности через кручение и неметричность (7.15) дивергенция векторного поля (7.52) принимает вид

$$\nabla_\alpha X^\alpha = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}X^\alpha\right) + T_\alpha X^\alpha + \frac{1}{2}Q_\alpha X^\alpha. \tag{7.53}$$

В частности, в римановой геометрии

$$\widetilde{\nabla}_\alpha X^\alpha = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}X^\alpha\right). \tag{7.54}$$

Эта полезная формула позволяет переписать инвариантный интеграл от дивергенции произвольного векторного поля в римановой геометрии в виде

$$\int d^n x\sqrt{|g|}\widetilde{\nabla}_\alpha X^\alpha = \int d^n x\partial_\alpha(\sqrt{|g|}X^\alpha)$$

и воспользоваться формулой Стокса (см. раздел 4.6.2). В пространстве аффинной связности после применения формулы Стокса возникают дополнительные объемные интегралы с кручением и неметричностью (7.53).

Замечание. В разделе 4.5 мы определили дивергенцию (4.63) для внешних форм. Пусть на многообразии $\mathbb{M}$ задана метрика $g_{\alpha\beta}$ и 1-форма $A = dx^\alpha A_\alpha$, соответствующая вектору X, где $A_\alpha := X^\beta g_{\alpha\beta}$. Тогда из сравнения формул (4.56) и (7.52) следует, что $\mathsf{div}\, A = \mathsf{div}\, X$ тогда и только тогда, когда

$$T_\alpha X^\alpha + \frac{1}{2}Q_\alpha X^\alpha = 0.$$

В частности, равенство имеет место в римановой геометрии. □

Дивергенция от антисимметричного тензорного поля, $F^{\alpha\beta} = -F^{\beta\alpha}$, равна

$$\nabla_\alpha F^{\alpha\beta} = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}F^{\alpha\beta}\right) + F^{\alpha\beta}\left(T_\alpha + \frac{1}{2}Q_\alpha\right) + \frac{1}{2}F^{\alpha\gamma}T_{\alpha\gamma}{}^\beta. \tag{7.55}$$

В римановой геометрии это тождество упрощается:

$$\widetilde{\nabla}_\alpha F^{\alpha\beta} = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}F^{\alpha\beta}\right). \tag{7.56}$$

Приведем также формулу для антисимметричной ковариантной производной от 1-формы:

$$\nabla_\alpha X_\beta - \nabla_\beta X_\alpha = \partial_\alpha X_\beta - \partial_\beta X_\alpha - T_{\alpha\beta}{}^\gamma X_\gamma. \tag{7.57}$$

Из этого равенства следует, что антисимметризация обычной производной от 1-формы дает тензорное поле типа $(0,2)$. Это процедура положена в основу определения внешней производной от произвольной формы (см. раздел 4.3).

Наличие ковариантной производной ∇_α и метрики $g_{\alpha\beta}$ позволяет строить ковариантные дифференциальные операторы, действующие на произвольные тензорные поля на многообразии. Для их построения достаточно взять произвольный дифференциальный оператор в (псевдо-)евклидовом пространстве, заменить частные производные ∂_α на ковариантные ∇_α и (псевдо-)евклидову метрику $\eta_{\alpha\beta}$ на метрику многообразия $g_{\alpha\beta}$.

Инвариантный дифференциальный оператор второго порядка, действующий на произвольные тензорные поля на многообразии имеет вид

$$\triangle := g^{\alpha\beta}\nabla_\alpha\nabla_\beta = \nabla_\alpha(g^{\alpha\beta}\nabla_\beta) - Q_\alpha{}^{\alpha\beta}\nabla_\beta. \tag{7.58}$$

Он часто называется *оператором Лапласа–Бельтрами* независимо от сигнатуры метрики. В римановой геометрии оператор Лапласа–Бельтрами устроен проще:

$$\tilde{\triangle} := g^{\alpha\beta}\widetilde{\nabla}_\alpha\widetilde{\nabla}_\beta = \widetilde{\nabla}_\alpha(g^{\alpha\beta}\widetilde{\nabla}_\beta). \tag{7.59}$$

Замечание. Если метрика имеет лоренцеву сигнатуру, то мы будем обозначать оператор Лапласа–Бельтрами квадратом $\Box$, а не треугольником $\triangle$, по аналогии с оператором Даламбера. В этом случае уравнение $\Box K_{\beta_1\dots\beta_s}{}^{\alpha_1\dots\alpha_r} = 0$ для каждой компоненты тензорного поля $K \in \mathcal{T}^r_s(\mathbb{M})$ будет гиперболического типа. $\Box$

Действие оператора Лапласа–Бельтрами на скалярное поле можно записать в виде

$$\triangle f = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}g^{\alpha\beta}\partial_\beta f\right) + \left(T^\beta + \frac{1}{2}Q^\beta - Q_\alpha{}^{\alpha\beta}\right)\partial_\beta f. \tag{7.60}$$

В римановой геометрии правая часть этого равенства записывается через обычную производную от тензорной плотности:

$$\tilde{\triangle} f = \frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}g^{\alpha\beta}\partial_\beta f\right). \tag{7.61}$$

Приведем также формулу интегрирования по частям. Используя (7.46) и (7.48), проверяется, что

$$\partial_\alpha\left(\sqrt{|g|}X^{\alpha\text{A}}Y_\text{A}\right) = \sqrt{|g|}\nabla_\alpha X^{\alpha\text{A}}Y_\text{A} - \sqrt{|g|}\left(T_\alpha + \frac{1}{2}Q_\alpha\right)X^{\alpha\text{A}}Y_\text{A} + \sqrt{|g|}X^{\alpha\text{A}}\nabla_\alpha Y_\text{A}.$$

Здесь индекс A обозначает произвольную совокупность ковариантных и контравариантных индексов, по которым подразумевается суммирование. Интегрируя это соотношение по многообразию и пренебрегая граничными членами, получим

$$\int d^nx\,\sqrt{|g|}X^{\alpha\text{A}}\nabla_\alpha Y_\text{A} = -\int d^nx\,\sqrt{|g|}\left[\nabla_\alpha X^{\alpha\text{A}}Y_\text{A} - \left(T_\alpha + \frac{1}{2}Q_\alpha\right)X^{\alpha\text{A}}Y_\text{A}\right]. \tag{7.62}$$

В римановой геометрии эта формула имеет тот же вид, что и обычное интегрирование по частям:

$$\int d^nx\,\sqrt{|g|}X^{\alpha\text{A}}\widetilde{\nabla}_\alpha Y_\text{A} = -\int d^nx\,\sqrt{|g|}\left[\widetilde{\nabla}_\alpha X^{\alpha\text{A}}Y_\text{A}\right]. \tag{7.63}$$

Отсюда следует, что с точностью до граничных членов имеет место равенство

$$\int d^nx\,\sqrt{|g|}X^\text{A}\tilde{\triangle}Y_\text{A} = \int d^nx\,\sqrt{|g|}\tilde{\triangle}X^\text{A}Y_\text{A}. \tag{7.64}$$

7.7 Локальное определение тензора кривизны

Помимо кручения аффинная связность $\Gamma_{\alpha\beta}{}^\gamma$ на многообразии $\mathbb{M}$ задает еще один важный геометрический объект – *тензор кривизны* аффинной связности (6.35) или *тензор Римана–Кристоффеля*. В локальной системе координат он имеет следующие компоненты:

$$R_{\alpha\beta\gamma}{}^\delta := \partial_\alpha\Gamma_{\beta\gamma}{}^\delta - \Gamma_{\alpha\gamma}{}^\epsilon\Gamma_{\beta\epsilon}{}^\delta - (\alpha \leftrightarrow \beta), \tag{7.65}$$

где скобки $(\alpha \leftrightarrow \beta)$ обозначают предыдущие слагаемые с переставленными индексами α и β.

Замечание. Тензор кривизны никакого отношения к метрике не имеет и определяется только связностью. □

Тензор кривизны играет очень важную роль в дифференциальной геометрии и возникает в различных контекстах. Покажем, что тензор кривизны позволяет сформулировать критерий локальной тривиальности аффинной связности. Пусть на многообразии задана аффинная связность с нулевым кручением, $T_{\alpha\beta}{}^{\gamma} = 0$. При этом неметричность, если задана также метрика, может быть отлична от нуля. Рассмотрим соотношения (6.31) как уравнения на функции перехода $x^{\alpha}(x')$. Потребуем, чтобы в новой системе координат компоненты связности обращались в нуль в некоторой окрестности. Тогда функции перехода должны удовлетворять уравнению

$$\frac{\partial^2 x^{\alpha}}{\partial x^{\beta'}\partial x^{\gamma'}} = -\frac{\partial x^{\beta}}{\partial x^{\beta'}}\frac{\partial x^{\gamma}}{\partial x^{\gamma'}}\Gamma_{\beta\gamma}{}^{\alpha}. \tag{7.66}$$

Дифференцируя это соотношение по $x^{\delta'}$ и исключая вторые производные от функций перехода с помощью (7.66), получим равенство

$$\frac{\partial^3 x^{\alpha}}{\partial x^{\beta'}\partial x^{\gamma'}\partial x^{\delta'}} = \frac{\partial x^{\beta}}{\partial x^{\beta'}}\frac{\partial x^{\gamma}}{\partial x^{\gamma'}}\frac{\partial x^{\delta}}{\partial x^{\delta'}}\Big(- \partial_{\delta}\Gamma_{\beta\gamma}{}^{\alpha} + \Gamma_{\beta\delta}{}^{\epsilon}\Gamma_{\epsilon\gamma}{}^{\alpha} + \Gamma_{\delta\gamma}{}^{\epsilon}\Gamma_{\beta\epsilon}{}^{\alpha}\Big).$$

Условия интегрируемости уравнений (7.66) получаются из этих уравнений антисимметризацией выражения в круглых скобках по индексам $[\beta\delta]$ или $[\gamma\delta]$ и приравниванием результата нулю. Обе антисимметризации приводят к единственному условию: равенству нулю тензора кривизны (7.65).

Замечание. Если кручение на многообразии равно нулю, то равенство нулю тензора кривизны в некоторой области является необходимым и достаточным условием обращения в нуль симметричной части компонент аффинной связности.

□

Определение. Назовем аффинную связность на многообразии $\mathbb{M}$ *локально тривиальной*, если для любой точки $x \in \mathbb{M}$ найдется окрестность $\mathbb{U}_x \ni x$ и такая система координат, что компоненты связности $\Gamma_{\alpha\beta}{}^{\gamma}$ на $\mathbb{U}_x$ равны нулю. □

Таким образом, найден критерий локальной тривиальности аффинной связности.

Теорема 7.7.1. *Для локальной тривиальности аффинной связности на многообразии $\mathbb{M}$ необходимо и достаточно, чтобы ее кручение и тензор кривизны равнялись нулю на $\mathbb{M}$.*

Замечание. Тензор кручения является тензором, и его компоненты нельзя обратить в нуль никаким преобразованием координат. Поэтому равенство нулю тензора кривизны необходимо для локальной тривиальности аффинной связности. □

Определение. Аффинная связность, для которой тензоры кривизны и кручения равны нулю, называется *интегрируемой*. □

Пусть на многообразии $\mathbb{M}$ помимо связности $\Gamma_{\alpha\beta}{}^{\gamma}$ задана также метрика $g_{\alpha\beta}$. Тогда определен тензор неметричности (7.14). При ненулевом тензоре неметричности локальная тривиальность аффинной связности не означает, что многообразие является локально (псевдо-)евклидовым. Действительно, при $\Gamma_{\alpha\beta}{}^{\gamma} = 0$ и $T_{\alpha\beta}{}^{\gamma} = 0$ уравнение (7.14) дает соотношение между метрикой и неметричностью

$$\partial_{\alpha} g_{\beta\gamma} = -Q_{\alpha\beta\gamma}. \tag{7.67}$$

В частности, метрику можно задать произвольно, и она будет определять неметричность. В римановой геометрии $T_{\alpha\beta}{}^{\gamma} = 0$ и $Q_{\alpha\beta\gamma} = 0$, и равенство нулю тензора кривизны, $\widetilde{R}_{\alpha\beta\gamma}{}^{\delta} = 0$, означает, что найдется такая система координат, в которой $\partial_\alpha g_{\gamma\delta} = 0$, т.е. метрика имеет постоянные компоненты в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$. С помощью последующего линейного преобразования координат ее всегда можно привести к диагональному виду, когда на диагонали будут стоять ± 1. Это дает критерий локальной (псевдо-)евклидовости. Сформулируем сразу глобальное утверждение.

Теорема 7.7.2. *Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия. Тогда, если тензоры кручения, неметричности и кривизны равны нулю на всем многообразии, то $\mathbb{M}$ изометрично либо (псевдо-)евклидову пространству $\mathbb{R}^n$, либо факторпространству $\mathbb{R}^n/\mathbb{G}$, где $\mathbb{G}$ – группа преобразований, действующая свободно и собственно разрывно.*

Доказательство. См., например, [**?**]. □

7.8 Свойства тензора кривизны

Помимо тождеств Бианки тензоры кручения и кривизны удовлетворяют ряду других дифференциальных соотношений. Выведем ряд полезных тождеств, исходя из определения тензора кривизны (7.65). Прямые вычисления с учетом выражения аффинной связности через метрику, кручение и неметричность показывают, что антисимметризация тензора кривизны по первым трем индексам определяется только тензором кручения и его ковариантными производными:

$$\begin{aligned} R_{\alpha\beta\gamma}{}^{\delta} + R_{\beta\gamma\alpha}{}^{\delta} + R_{\gamma\alpha\beta}{}^{\delta} &= \nabla_\alpha T_{\beta\gamma}{}^{\delta} + \nabla_\beta T_{\gamma\alpha}{}^{\delta} + \nabla_\gamma T_{\alpha\beta}{}^{\delta} \\ &\quad + T_{\alpha\beta}{}^{\epsilon} T_{\epsilon\gamma}{}^{\delta} + T_{\beta\gamma}{}^{\epsilon} T_{\epsilon\alpha}{}^{\delta} + T_{\gamma\alpha}{}^{\epsilon} T_{\epsilon\beta}{}^{\delta}. \end{aligned} \tag{7.68}$$

Это означает, что в римановой геометрии, а также в аффинной геометрии без кручения тензор кривизны удовлетворяет тождеству

$$R_{\alpha\beta\gamma}{}^{\delta} + R_{\beta\gamma\alpha}{}^{\delta} + R_{\gamma\alpha\beta}{}^{\delta} = 0. \tag{7.69}$$

Обратное утверждение в общем случае неверно. То есть из равенства (7.69) не следует, что кручение равно нулю.

Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия, т.е. метрика и связность. Приведем явное выражение для тензора кривизны со всеми опущенными индексами. С учетом разложения связности (7.15) и тождества (7.23) получаем

$$R_{\alpha\beta\gamma\delta} := R_{\alpha\beta\gamma}{}^{\epsilon} g_{\epsilon\delta} = \partial_\alpha \Gamma_{\beta\gamma\delta} + \Gamma_{\alpha\gamma}{}^{\epsilon} \Gamma_{\beta\delta\epsilon} - \Gamma_{\alpha\gamma}{}^{\epsilon} (T_{\beta\delta\epsilon} + Q_{\beta\delta\epsilon}) - (\alpha \leftrightarrow \beta). \tag{7.70}$$

Если у тензора кривизны со всеми опущенными индексами произвести симметризацию по последней паре индексов, то получим тождество

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\beta\delta\gamma} = \nabla_\alpha Q_{\beta\gamma\delta} - \nabla_\beta Q_{\alpha\gamma\delta} + T_{\alpha\beta}{}^{\epsilon} Q_{\epsilon\gamma\delta}. \tag{7.71}$$

Свертывая это тождество по последней паре индексов, получаем равенство

$$\begin{aligned} 2R_{\alpha\beta\gamma}{}^{\gamma} = 2(\partial_\alpha \Gamma_\beta - \partial_\beta \Gamma_\alpha) &= \nabla_\alpha Q_\beta - \nabla_\beta Q_\alpha + T_{\alpha\beta}{}^{\gamma} Q_\gamma \\ &= \partial_\alpha Q_\beta - \partial_\beta Q_\alpha, \end{aligned} \tag{7.72}$$

где $Q_\alpha = Q_{\alpha\beta}{}^\beta$. Отсюда следует, что в геометрии Римана–Картана–Вейля свертка тензора кривизны по последним двум индексам дает напряженность для формы Вейля.

Рассмотрим симметрии тензора кривизны относительно перестановок индексов. В общем случае аффинной геометрии единственная симметрия тензора кривизны со всеми опущенными индексами – это антисимметрия по первой паре индексов:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}, \tag{7.73}$$

что сразу вытекает из определения (7.65). Отсюда следует, что в аффинной геометрии число линейно независимых компонент тензора кривизны равно

$$[R_{\alpha\beta\gamma\delta}] = \frac{n^3(n-1)}{2}.$$

В геометрии Римана–Картана тензор кривизны антисимметричен также и по второй паре индексов, что является следствием (7.71),

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma}. \tag{7.74}$$

Поэтому число его линейно независимых компонент меньше:

$$[R_{\alpha\beta\gamma\delta}] = \frac{n^2(n-1)^2}{4}.$$

В римановой геометрии тензор кривизны обладает дополнительной симметрией: его антисимметризация по первым трем индексам тождественно обращается в нуль

$$\widetilde{R}_{[\alpha\beta\gamma]\delta} = 0. \tag{7.75}$$

что является следствием (7.68). Следовательно, число его линейно независимых компонент равно

$$[\widetilde{R}_{\alpha\beta\gamma\delta}] = \frac{n^2(n^2-1)}{12}.$$

Из свойств (7.73), (7.74) и (7.75) следует, что в римановой геометрии тензор кривизны симметричен также относительно перестановки первой пары индексов со второй,

$$\widetilde{R}_{\alpha\beta\gamma\delta} = \widetilde{R}_{\gamma\delta\alpha\beta}. \tag{7.76}$$

Замечание. Обратные утверждения, связывающие геометрию с симметрией тензора кривизны неверны. Например, тензор кривизны может быть антисимметричен по второй паре индексов и в то же время тензор неметричности может быть нетривиальным. □

Определение. По заданному тензору кривизны путем свертки пары индексов можно построить *тензор Риччи*

$$R_{\alpha\gamma} = R_{\alpha\beta\gamma}{}^\beta. \quad \square \tag{7.77}$$

Замечание. Тензор кривизны и тензор Риччи строятся только по аффинной связности, без использования метрики. □

В общем случае тензор Риччи не обладает никакой симметрией по своим индексам. В римановой геометрии он симметричен относительно перестановки индексов:

$$\widetilde{R}_{\alpha\gamma} = \widetilde{R}_{\gamma\alpha}.$$

Определение. Свертывая тензор Риччи с обратной метрикой, получаем *скалярную кривизну* многообразия

$$R = R_{\alpha\beta} g^{\alpha\beta}. \tag{7.78}$$

Скалярная кривизна зависит и от связности, и от метрики. Функция

$$K := -\frac{1}{2} R, \tag{7.79}$$

где R –скалярная кривизна, называется *гауссовой кривизной* многообразия. □

Пример 7.8.1. Пусть двумерная поверхность задана в трехмерном евклидовом пространстве $\mathbb{R}^3$ уравнением $z = f(x, y)$, причем

$$f(0,0) = 0, \qquad f_x(0,0) = 0, \qquad f_y(0,0) = 0,$$

где $f_x = \partial_x f$ и $f_y = \partial_y f$. Тогда скалярная кривизна в начале координат, соответствующая индуцированной метрике,

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 = \\ &= (1 + f_x^2) dx^2 + 2 f_x f_y dx dy + (1 + f_y^2) dy^2, \end{aligned}$$

пропорциональна определителю гессиана

$$R = -2 \,\mathsf{det} \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = -2K,$$

где K – гауссова кривизна поверхности. Поворотом евклидова пространства в плоскости x, y гессиан всегда можно привести к диагональному виду

$$\begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix},$$

где, скажем, $f_{xx}(0,0) \geq f_{yy}(0,0)$. Если $f_{xx} > f_{yy}$, то оси x, y называются *главными направлениями* поверхности в данной точке. При $f_{xx} = f_{yy}$, главные направления не определены. Если гауссова кривизна положительна, $K > 0$, то поверхность в начале координат имеет локальный экстремум и лежит по одну сторону плоскости $z = 0$. Если кривизна меньше нуля, $K < 0$, то поверхность в начале координат имеет седловую точку. При $K = 0$ по крайней мере одна из вторых частных производных обращается в нуль, и мы имеем касание более высокого порядка. □

Замечание. Многие авторы определяют тензор кривизны многообразия таким образом, чтобы скалярная кривизна совпала с гауссовой кривизной. Мы используем определения тензора кривизны, тензора Риччи и скалярной кривизны, которые широко используется в дифференциальной геометрии и приложениях. Это не приводит к недоразумениям. Однако, надо помнить, что в наших обозначениях скалярная и гауссова кривизна двумерной сферы радиуса r равны: $R = -2K = -2/r^2$ и $K = 1/r^2$ (см. вычисления в разделе **??**). □

В аффинной геометрии компоненты связности $\Gamma_{\alpha\beta}{}^{\gamma}$ инвариантны относительно преобразования

$$g_{\alpha\beta} \mapsto -g_{\alpha\beta}, \qquad T_{\alpha\beta}{}^{\gamma} \mapsto T_{\alpha\beta}{}^{\gamma}, \qquad Q_{\alpha\beta}{}^{\gamma} \mapsto Q_{\alpha\beta}{}^{\gamma}. \tag{7.80}$$

Поэтому тензор кривизны и тензор Риччи также инвариантны. При этом скалярная кривизна (7.78) меняет знак.

Рассмотрим на многообразии $\mathbb{M}$ произвольное векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ и 1-форму $A = dx^\alpha A_\alpha \in \Lambda_1(\mathbb{M})$. Ковариантные производные от их компонент имеют вид (7.2), (7.3). Прямые вычисления показывают, что коммутатор двух ковариантных производных определяется тензором кривизны и кручения:

$$[\nabla_\alpha, \nabla_\beta] X^\gamma = R_{\alpha\beta\delta}{}^\gamma X^\delta - T_{\alpha\beta}{}^\delta \nabla_\delta X^\gamma, \tag{7.81}$$

$$[\nabla_\alpha, \nabla_\beta] X_\gamma = -R_{\alpha\beta\gamma}{}^\delta X_\delta - T_{\alpha\beta}{}^\delta \nabla_\delta X_\gamma. \tag{7.82}$$

и отличается только знаком перед слагаемым с кривизной. Коммутатор ковариантных производных от скалярного поля $\varphi(x)$ проще:

$$[\nabla_\alpha, \nabla_\beta] \varphi = -T_{\alpha\beta}{}^\gamma \nabla_\gamma \varphi \tag{7.83}$$

и определяется только тензором кручения.

Формулы (7.81), (7.82) обобщаются на тензоры произвольного ранга.

Пример 7.8.2. Коммутатор ковариантных производных от компонент тензора $K \in \mathcal{T}_1^1(\mathbb{M})$ типа $(1,1)$ имеет вид

$$[\nabla_\alpha, \nabla_\beta] K^\gamma{}_\delta = R_{\alpha\beta\epsilon}{}^\gamma K^\epsilon{}_\delta - R_{\alpha\beta\delta}{}^\epsilon K^\gamma{}_\epsilon - T_{\alpha\beta}{}^\epsilon \nabla_\epsilon K^\gamma{}_\delta.$$

Для тензоров более высокого ранга будем иметь по одному слагаемому с кривизной со знаком плюс и минус соответственно для каждого контравариантного и ковариантного индекса и одно общее слагаемое с тензором кручения. □

Формулы для коммутатора ковариантных производных можно использовать для доказательства некоторых тождеств. Докажем несколько формул, справедливых в римановой геометрии. Пусть $F^{\alpha\beta}$ – произвольный антисимметричный тензор второго ранга. Тогда в римановой геометрии справедливо равенство

$$\widetilde{\nabla}_\alpha \widetilde{\nabla}_\beta F^{\alpha\beta} = 0.$$

Действительно,

$$[\widetilde{\nabla}_\alpha, \widetilde{\nabla}_\beta] F^{\alpha\beta} = \widetilde{R}_{\alpha\beta\gamma}{}^\alpha F^{\gamma\beta} + \widetilde{R}_{\alpha\beta\gamma}{}^\beta F^{\alpha\gamma} = 2\widetilde{R}_{\alpha\beta} F^{\alpha\beta} = 0,$$

поскольку тензор Риччи в этом случае симметричен.

Предложение 7.8.1. *В (псевдо-)римановой геометрии имеет место следующая формула*

$$\widetilde{\nabla}_\alpha \widetilde{\nabla}_\beta \widetilde{R}^{\alpha\beta\gamma\delta} = 0.$$

Доказательство. Справедливы следующие равенства:

$$\begin{aligned} 2\widetilde{\nabla}_\alpha \widetilde{\nabla}_\beta \widetilde{R}^{\alpha\beta\gamma\delta} &= [\widetilde{\nabla}_\alpha, \widetilde{\nabla}_\beta] R^{\alpha\beta\gamma\delta} = \\ &= \widetilde{R}_{\alpha\beta\epsilon}{}^\alpha \widetilde{R}^{\epsilon\beta\gamma\delta} + \widetilde{R}_{\alpha\beta\epsilon}{}^\beta \widetilde{R}^{\alpha\epsilon\gamma\delta} + \widetilde{R}_{\alpha\beta\epsilon}{}^\gamma \widetilde{R}^{\alpha\beta\epsilon\delta} + \widetilde{R}_{\alpha\beta\epsilon}{}^\delta \widetilde{R}^{\alpha\beta\gamma\epsilon}. \end{aligned}$$

Первые два слагаемых обращаются в нуль в силу симметрии тензора Риччи. Последние два слагаемых сокращаются:

$$\widetilde{R}_{\alpha\beta\epsilon}{}^\gamma \widetilde{R}^{\alpha\beta\epsilon\delta} - \widetilde{R}_{\alpha\beta\epsilon}{}^\delta \widetilde{R}^{\alpha\beta\epsilon\gamma} = 0$$

в силу симметрии слагаемых по индексам γ, δ. □

В заключение настоящего раздела приведем важную формулу, связывающую скалярную кривизну $R(g,\Gamma)$ в геометрии Римана–Картана со скалярной кривизной $\tilde{R}(g)$ в соответствующей римановой геометрии, вычисленной только по метрике при нулевом кручении,

$$R+\frac{1}{4}T_{\alpha\beta\gamma}T^{\alpha\beta\gamma}-\frac{1}{2}T_{\alpha\beta\gamma}T^{\gamma\alpha\beta}-T_\alpha T^\alpha-\frac{2}{\sqrt{|g|}}\partial_\alpha(\sqrt{|g|}T^\alpha)=\tilde{R}, \tag{7.84}$$

где $T_\alpha = T_{\beta\alpha}{}^\beta$ – след тензора кручения. Это тождество проверяется путем прямых вычислений и справедливо в пространстве произвольной размерности $n \geq 2$ и для метрики произвольной сигнатуры. Приведенное тождество можно переписать в явно ковариантном виде

$$R+\frac{1}{4}T_{\alpha\beta\gamma}T^{\alpha\beta\gamma}-\frac{1}{2}T_{\alpha\beta\gamma}T^{\gamma\alpha\beta}-T_\alpha T^\alpha-2\tilde{\nabla}_\alpha T^\alpha=\tilde{R},$$

где мы воспользовались тождеством (7.54). Важность этой формулы заключается в том, что, поскольку обе скалярные кривизны приводят к уравнениям движения не выше второго порядка (в переменных Картана), то при выборе лагранжиана можно ограничиться одной скалярной кривизной и квадратами тензоров кручения.

7.9 Неголономный базис

Аффинная геометрия на многообразии $\mathbb{M}$ задается метрикой и аффинной связностью или, что эквивалентно, метрикой, кручением и неметричностью. При таком описании каждое преобразование координат сопровождаются соответствующим преобразованием компонент тензорных полей относительно координатного базиса. Существует также другой способ описания геометрии, когда компоненты тензорных полей рассматриваются относительно репера, который не меняется при преобразовании координат. В этом случае на компоненты тензорных полей действует группа локальных преобразований $\mathbb{GL}(n,\mathbb{R})$, что соответствует вращению репера. В результате геометрия на многообразии $\mathbb{M}$ будет задана репером $e_\alpha{}^a$ и линейной или $\mathbb{GL}(n,\mathbb{R})$-связностью $\omega_{\alpha a}{}^b$ (см. раздел 6.4).

Определение. Переменные репер $e_\alpha{}^a(x)$ и $\mathbb{GL}(n,\mathbb{R})$-связность $\omega_{\alpha a}{}^b(x)$, задающие на многообразии $\mathbb{M}$ аффинную геометрию, называются *переменными Картана*. В четырехмерном пространстве-времени репер называется *тетрадой*. В двумерном и трехмерном пространстве репер называется соответственно *диадой* и *триадой*. □

Замечание. В моделях математической физики переменные Картана, как правило, упрощают вычисления и необходимы при рассмотрении спинорных полей на многообразии $\mathbb{M}$. □

Напомним, что координатный базис касательных пространств $\mathbb{T}_x(\mathbb{M})$ во всех точках многообразия $x \in \mathbb{M}$ мы обозначаем ∂_α, и он называется *голономным*. Важным свойством координатных базисных векторов ∂_α является их коммутативность:

$$[\partial_\alpha, \partial_\beta] = 0.$$

Предположим, что в каждой точке многообразия $x \in \mathbb{M}$ задан произвольный базис касательного пространства $e_a(x)$ (репер) и дуальный к нему базис 1-форм $e^a(x)$ (корепер), $a = 1,\ldots,n$. Дуальность означает, что значение 1-форм e^a на векторных полях e_b равно символу Кронекера: $e^a(e_b) = \delta^a_b$.

Замечание. Как уже отмечалось, репер может существовать глобально не для всех многообразий. Например, его не существует на неориентируемых многообразиях. В таких случаях все, сказанное ниже, имеет локальный характер. Тем не менее полученные формулы важны для вычислений, которые, как правило, проводятся в какой либо системе координат. □

Репер и корепер можно разложить по координатному базису:

$$e_a = e^\alpha{}_a \partial_\alpha, \qquad e^a = dx^\alpha e_\alpha{}^a,$$

где $e^\alpha{}_a$ и $e_\alpha{}^a$ – взаимно обратные невырожденные матрицы:

$$e^\alpha{}_a e_\alpha{}^b = \delta^b_a, \qquad e^\alpha{}_a e_\beta{}^a = \delta^\alpha_\beta.$$

По-предположению, матрицы $e^\alpha{}_a$ и $e_\alpha{}^a$ невырождены и гладко зависят от точки многообразия.

В общем случае репер представляет собой *неголономный базис* касательного пространства, т.е. не существует такой системы координат $x^a = x^a(x)$, что

$$e^\alpha{}_a = \partial_a x^\alpha. \tag{7.85}$$

Репер определен с точностью до локальных $\mathbb{GL}(n,\mathbb{R})$ преобразований, действующих на латинские индексы. Его важнейшей характеристикой являются *компоненты неголономности* $c_{ab}{}^c$, которые определяются коммутатором базисных векторных полей:

$$[e_a, e_b] = c_{ab}{}^c e_c \tag{7.86}$$

и антисимметричны по нижним индексам,

$$c_{ab}{}^c(x) = -c_{ba}{}^c(x).$$

Из тождеств Якоби для алгебры Ли векторных полей,

$$\Big[e_a[e_b, e_c]\Big] + \Big[e_b[e_c, e_a]\Big] + \Big[e_c[e_a, e_b]\Big] = 0,$$

следуют тождества для компонент неголономности:

$$\partial_a c_{bc}{}^d + \partial_b c_{ca}{}^d + \partial_c c_{ab}{}^d + c_{ab}{}^e c_{ce}{}^d + c_{bc}{}^e c_{ae}{}^d + c_{ca}{}^e c_{be}{}^d = 0, \tag{7.87}$$

где $\partial_a := e^\alpha{}_a \partial_\alpha$.

Из определения (7.86) следует явное выражение для компонент неголономности через компоненты репера и их производные

$$c_{ab}{}^c = \Big(e^\alpha{}_a \partial_\alpha e^\beta{}_b - e^\alpha{}_b \partial_\alpha e^\beta{}_a\Big) e_\beta{}^c. \tag{7.88}$$

Умножив это соотношение на обратные матрицы $e_\alpha{}^a$, получим эквивалентную формулу

$$c_{\alpha\beta}{}^c = e_\alpha{}^a e_\beta{}^b c_{ab}{}^c = -\partial_\alpha e_\beta{}^c + \partial_\beta e_\alpha{}^c, \tag{7.89}$$

которую можно переписать в виде

$$de^c = \frac{1}{2} e^a \wedge e^b \, c_{ba}{}^c, \tag{7.90}$$

где использовано определение внешнего умножения и дифференцирования форм (см. раздел 4).

Многие формулы содержат след компонент неголономности, который определяется следующим образом:

$$c_a = c_{ba}{}^b = \partial_\alpha e^\alpha{}_a + \frac{\partial_a \sqrt{|g|}}{\sqrt{|g|}}. \tag{7.91}$$

Компоненты неголономности ковариантны относительно преобразования координат $x^\alpha \to x^{\alpha'}(x)$, но не являются компонентами вектора относительно локальных $\mathbb{GL}(n,\mathbb{R})$ преобразований. Нетрудно проверить, что равенство нулю компонент неголономности является необходимым и достаточным условием локальной разрешимости системы уравнений (7.85). Это означает, что, если компоненты неголономности равны нулю в некоторой области, то для любой точки из этой области существует окрестность, в которой можно выбрать такую систему координат, что базис станет голономным $e_a = \partial_a$.

Использование неголономного базиса вместо координатного бывает значительно удобнее и часто используется в приложениях. Поэтому получим основные формулы дифференциальной геометрии в неголономном базисе.

Произвольное векторное поле можно разложить как по координатному, так и по некоординатному (неголономному) базису

$$X = X^\alpha \partial_\alpha = X^a e_a,$$

где $X^\alpha = e^\alpha{}_a X^a$ и $X^a = X^\alpha e_\alpha{}^a$. Предположим, что переход от греческих индексов к латинским и наоборот у компонент тензорных полей произвольного ранга всегда осуществляется с помощью компонент репера и корепера. При этом все симметрии относительно перестановок индексов, конечно, сохраняются. Компоненты метрики в неголономном базисе имеют вид

$$g_{ab} = e^\alpha{}_a e^\beta{}_b g_{\alpha\beta}. \tag{7.92}$$

В общем случае компоненты метрики $g_{ab}(x)$ зависят от точки многообразия. Метрика g_{ab} всегда имеет ту же сигнатуру, что и метрика $g_{\alpha\beta}$, т.к. матрица $e^\alpha{}_a$ невырождена. Подъем и опускание греческих и латинских индексов всегда осуществляется с помощью метрик $g_{\alpha\beta}$ и g_{ab}, соответственно.

Как правило, репер используют в тех случаях, когда матрица g_{ab} является диагональной и постоянной, а на диагонали расположены плюс и минус единицы:

$$g_{ab} = \eta_{ab} = \mathsf{diag}\,(\underbrace{+\cdots+}_{p}\underbrace{-\cdots-}_{q}).$$

Локально такой репер существует, поскольку уравнение(7.92) при одинаковых сигнатурах метрик g_{ab} и $g_{\alpha\beta}$ всегда разрешимо относительно репера. Такой репер называется *ортонормальным* и определен с точностью до $\mathbb{O}(p,q)$ вращений. Ортонормальный базис часто бывает более удобным, т.к. метрика в этом базисе постоянна.

Для римановой метрики множество реперов делится на два класса: с положительным и отрицательным определителем. Для многообразий с метрикой лоренцевой сигнатуры множество реперов можно разбить на четыре класса (см. раздел **??**), по числу несвязных компонент группы Лоренца.

Компоненты тензоров второго и более высокого рангов могут содержать одновременно и греческие и латинские индексы. По построению, ковариантная производная от компонент такого тензора содержит по одному слагаемому с аффинной связностью для каждого греческого индекса и одному слагаемому с линейной связностью для каждого латинского индекса. Из определения (локальной формы) линейной связности (6.51) следует взаимно однозначная связь между линейной и аффинной связностью:

$$\omega_{\alpha a}{}^{b} = \Gamma_{\alpha\beta}{}^{\gamma} e^{\beta}{}_{a} e_{\gamma}{}^{b} - \partial_{\alpha} e_{\beta}{}^{b} e^{\beta}{}_{a}. \tag{7.93}$$

Эту формулу можно переписать в виде равенства нулю ковариантной производной от компонент корепера:

$$\nabla_{\alpha} e_{\beta}{}^{a} = \partial_{\alpha} e_{\beta}{}^{a} - \Gamma_{\alpha\beta}{}^{\gamma} e_{\gamma}{}^{a} + e_{\beta}{}^{b} \omega_{\alpha b}{}^{a} = 0. \tag{7.94}$$

Отсюда следует, что ковариантная производная от компонент репера также обращается в нуль

$$\nabla_{\alpha} e^{\beta}{}_{a} = \partial_{\alpha} e^{\beta}{}_{a} + \Gamma_{\alpha\gamma}{}^{\beta} e^{\gamma}{}_{a} - \omega_{\alpha a}{}^{b} e^{\beta}{}_{b} = 0.$$

Тогда, используя правило Лейбница, можно свободно переходить от греческих индексов к латинским и наоборот под знаком ковариантного дифференцирования:

$$\nabla_{\alpha} X^{a} = \nabla_{\alpha}(X^{\beta} e_{\beta}{}^{a}) = (\nabla_{\alpha} X^{\beta}) e_{\beta}{}^{a},$$
$$\nabla_{\alpha} X_{a} = \nabla_{\alpha}(X_{\beta} e^{\beta}{}_{a}) = (\nabla_{\alpha} X_{\beta}) e^{\beta}{}_{a},$$

где

$$\begin{aligned} \nabla_{\alpha} X^{a} &= \partial_{\alpha} X^{a} + \omega_{\alpha b}{}^{a} X^{b}, \\ \nabla_{\alpha} X_{a} &= \partial_{\alpha} X_{a} - \omega_{\alpha a}{}^{b} X_{b} \end{aligned} \tag{7.95}$$

– ковариантные производные от компонент векторного поля относительно неголономного базиса. Если аффинная связность не является метрической, то операция подъема и опускания индексов с помощью метрик $g_{\alpha\beta}$ и g_{ab} не коммутирует с ковариантной производной.

Замечание. Формулу (7.93) можно рассматривать, как калибровочное преобразование (вращение) $\Gamma \to \omega$ в касательном пространстве, которое совпадает с преобразованием калибровочных полей Янга–Миллса (**??**). При этом репер $e_{\alpha}{}^{a} \in \mathbb{GL}(n, \mathbb{R})$ играет роль матрицы преобразования (локального вращения), а координаты многообразия не затрагиваются. □

В координатном базисе преобразование координат сопровождается преобразование компонент тензорных полей. Введение репера позволяет отделить преобразование координат от преобразований в касательном пространстве. Это достигается путем введения n^2 новых полей $e_{\alpha}{}^{a}(x)$. В результате появляется дополнительная возможность совершать локальные $\mathbb{GL}(n, \mathbb{R})$ преобразования, зависящие также от n^2 функций, в касательном пространстве, не затрагивая координат многообразия. Очевидно, что всегда можно совершить такое преобразование, что в результате репер совпадет с координатным базисом $e_{\alpha}{}^{a} = \delta^{a}_{\alpha}$. В этом случае линейная связность совпадет с аффинной $\omega_{\alpha a}{}^{b} = \Gamma_{\alpha a}{}^{b}$, а выражения для кривизны (7.96) и кручения (7.97) перейдут в уже знакомые формулы аффинной геометрии, т.к. компоненты неголономности обратятся в нуль: $c_{ab}{}^{c} = 0$.

Формулы для кривизны (6.56) и кручения (6.55) содержат два греческих индекса. Эти индексы также можно преобразовать в неголономные. Простые вычисления

приводят к следующим компонентам тензора кривизны и кручения в неголономном базисе

$$R_{abc}{}^{d} = \partial_a \omega_{bc}{}^{d} - \partial_b \omega_{ac}{}^{d} - \omega_{ac}{}^{e}\omega_{be}{}^{d} + \omega_{bc}{}^{e}\omega_{ae}{}^{d} - c_{ab}{}^{e}\omega_{ec}{}^{d}, \tag{7.96}$$

$$T_{ab}{}^{c} = \omega_{ab}{}^{c} - \omega_{ba}{}^{c} - c_{ab}{}^{c}, \tag{7.97}$$

где $\omega_{ab}{}^{c} := e^{\alpha}{}_{a}\omega_{\alpha b}{}^{c}$ и $\partial_a := e^{\alpha}{}_{a}\partial_\alpha$. Эти формулы также часто используются в приложениях, особенно тогда, когда компоненты линейной связности ω_{abc} являются постоянными относительно некоторого неголономного базиса. В разделе 9 мы используем их для вычисления тензора кручения и кривизны групп Ли.

Если на многообразии задана метрика, то ковариантные компоненты тензора кривизны в неголономном базисе равны

$$R_{abcd} = \partial_a \omega_{bcd} - \partial_b \omega_{acd} - \omega_{bc}{}^{e}\omega_{ade} + \omega_{ac}{}^{e}\omega_{bde} - c_{ab}{}^{e}\omega_{ecd}. \tag{7.98}$$

В этом случае линейную связность, аналогично аффинной, можно выразить через репер, кручение и неметричность. Из уравнения (7.15) и определения (7.93) следует выражение для линейной связности со всеми неголономными индексами:

$$\begin{aligned}\omega_{abc} = {} & \frac{1}{2}(\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}) + \frac{1}{2}(c_{abc} - c_{bca} + c_{cab}) \\ & + \frac{1}{2}(T_{abc} - T_{bca} + T_{cab}) + \frac{1}{2}(Q_{abc} + Q_{bca} - Q_{cab}),\end{aligned} \tag{7.99}$$

где $c_{abc} = c_{ab}{}^{d} g_{dc}$.

Если (псевдо-) риманово многообразие допускает векторное поле Киллинга $K^a = K^\alpha e_\alpha{}^{a}$, то уравнение Киллинга в неголономном базисе принимает вид

$$\widetilde{\nabla}_a K_b + \widetilde{\nabla}_b K_a = 0, \tag{7.100}$$

где

$$\widetilde{\nabla}_a K_b = e^{\alpha}{}_{a}\partial_\alpha K_b - \widetilde{\omega}_{ab}{}^{c} K_c, \qquad K_b = K^a g_{ab}. \tag{7.101}$$

Как и ранее знак тильды означает, что связность построена по метрике при нулевом кручении и неметричности. Эта формула следует из того, что переход между индексами можно проводить под знаком ковариантного дифференцирования.

В римановой геометрии кручение и неметричность равны нулю. Рассмотрим ортонормальные реперы, для которых $g_{ab} = \delta_{ab}$ или η_{ab}, если сигнатура отличается от евклидовой. Такие реперы определены с точностью до локальных $\mathbb{O}(n)$ вращений (или $\mathbb{O}(p,q)$ вращений, $p+q=n$, для неевклидовой сигнатуры). Тогда из формулы (7.99) следует выражение для соответствующей линейной связности ($\mathbb{O}(n)$- или $\mathbb{O}(p,q)$-связности) через компоненты неголономности:

$$\widetilde{\omega}_{abc} = \frac{1}{2}(c_{abc} - c_{bca} + c_{cab}). \tag{7.102}$$

7.10 Тождества Бианки

Тождества Бианки играют большую роль в приложениях и поэтому в настоящем разделе мы рассмотрим их в компонентах. Громоздкие, но прямые вычисления поз-

воляют записать тождества (6.44) и (6.45) после спуска на базу в переменных Картана:

$$\nabla_\alpha R_{\beta\gamma a}{}^b + \nabla_\beta R_{\gamma\alpha a}{}^b + \nabla_\gamma R_{\alpha\beta a}{}^b = T_{\alpha\beta}{}^\delta R_{\gamma\delta a}{}^b + T_{\beta\gamma}{}^\delta R_{\alpha\delta a}{}^b + T_{\gamma\alpha}{}^\delta R_{\beta\delta a}{}^b, \tag{7.103}$$

$$\begin{aligned}\nabla_\alpha T_{\beta\gamma}{}^a + \nabla_\beta T_{\gamma\alpha}{}^a + \nabla_\gamma T_{\alpha\beta}{}^a &= T_{\alpha\beta}{}^\delta T_{\gamma\delta}{}^a + T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^a + T_{\gamma\alpha}{}^\delta T_{\beta\delta}{}^a \\ &\quad + R_{\alpha\beta\gamma}{}^a + R_{\beta\gamma\alpha}{}^a + R_{\gamma\alpha\beta}{}^a.\end{aligned} \tag{7.104}$$

В аффинной геометрии их можно свернуть, соответственно, с репером $e^\gamma{}_b$ и $e^\gamma{}_a$:

$$\begin{aligned}\nabla_\alpha R_{\beta a} - \nabla_\beta R_{\alpha a} + \nabla_\gamma R_{\alpha\beta a}{}^\gamma &= -T_{\alpha\beta}{}^\delta R_{\delta a} - T_{\alpha\gamma}{}^\delta R_{\alpha\delta a}{}^\gamma + T_{\beta\gamma}{}^\delta R_{\alpha\delta a}{}^\gamma, \\ \nabla_\alpha T_\beta - \nabla_\beta T_\alpha - \nabla_\gamma T_{\alpha\beta}{}^\gamma &= T_{\alpha\gamma}{}^\delta T_{\beta\delta}{}^\gamma - T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^\gamma - T_{\alpha\beta}{}^\gamma T_\gamma + R_{\alpha\beta} - R_{\beta\alpha} - R_{\alpha\beta\gamma}{}^\gamma.\end{aligned}$$

В геометрии Римана–Картана, когда выполнено условие метричности, последнее слагаемое в последнем тождестве обращается в нуль. Тогда свернутые тождества Бианки можно переписать в виде

$$\nabla_\alpha R_{\beta\gamma} - \nabla_\beta R_{\alpha\gamma} + \nabla_\delta R_{\alpha\beta\gamma}{}^\delta = -T_{\alpha\beta}{}^\delta R_{\delta\gamma} - T_{\alpha\delta}{}^\epsilon R_{\beta\epsilon\gamma}{}^\delta + T_{\beta\delta}{}^\epsilon R_{\alpha\epsilon\gamma}{}^\delta, \tag{7.105}$$

$$\nabla_\alpha T_\beta - \nabla_\beta T_\alpha - \nabla_\gamma T_{\alpha\beta}{}^\gamma = T_{\alpha\gamma}{}^\delta T_{\beta\delta}{}^\gamma - T_{\beta\gamma}{}^\delta T_{\alpha\delta}{}^\gamma - T_{\alpha\beta}{}^\gamma T_\gamma + R_{\alpha\beta} - R_{\beta\alpha}, \tag{7.106}$$

где мы перешли ко всем греческим индексам.

Из последнего тождества (7.106) следует, что антисимметричная часть тензора Риччи в римановой геометрии (при нулевом кручении) всегда равна нулю: $\widetilde{R}_{\alpha\beta} = \widetilde{R}_{\beta\alpha}$.

Далее, свертка (7.105) с $g^{\alpha\gamma}$ (при нулевом тензоре неметричности) приводит к равенству

$$2\nabla_\beta R_\alpha{}^\beta - \nabla_\alpha R = 2T_{\alpha\beta\gamma}R^{\gamma\beta} - R_{\alpha\beta\gamma\delta}T^{\gamma\delta\beta}. \tag{7.107}$$

В римановой геометрии это тождество упрощается:

$$\widetilde{\nabla}_\beta G^{\beta\alpha} = 0, \tag{7.108}$$

где

$$G_{\alpha\beta} := \widetilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta}\widetilde{R}$$

– тензор Эйнштейна, построенный из тензора Риччи и скалярной кривизны при нулевом кручении и тензоре неметричности. Расширить ???

Тождество (7.106) антисимметрично по индексам $\alpha\beta$ и поэтому свертка с $g^{\alpha\beta}$ приводит к тождеству $0 = 0$.

Глава 8

Криволинейные координаты в $\mathbb{R}^3$

Формализм дифференциальной геометрии, развитый в предыдущих разделах, становится полезным и естественным даже в евклидовом пространстве, если вычисления проводятся в криволинейных системах координат. В настоящем главе мы продемонстрируем это на примере сферической и цилиндрической систем координат в трехмерном евклидовом пространстве $\mathbb{R}^3$, которые наиболее часто используются в приложениях. В частности, понятие метрики и символов Кристоффеля чрезвычайно полезно при нахождении явного вида ковариантных дифференциальных операторов, которые часто используются в приложениях.

8.1 Сферические координаты

Рассмотрим *сферические координаты* r, θ, φ в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x, y, z, рис. 8.1. Функции перехода от сферических

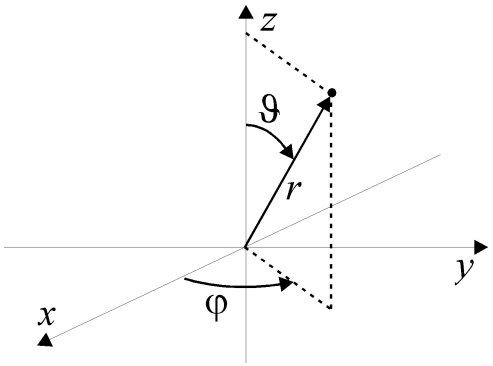


Рис. 8.1: Сферические координаты r, θ, φ в трехмерном евклидовом пространстве $\mathbb{R}^3$.

координат к декартовым имеют хорошо известный вид

$$\begin{aligned} x &= r \sin\theta \cos\varphi, \\ y &= r \sin\theta \sin\varphi, \\ z &= r \cos\theta, \end{aligned} \tag{8.1}$$

и определены при всех значениях $r,\ \theta,\ \varphi$. Углы θ и φ называются соответственно *азимутальным* и *полярным*. Якобиан этого преобразования легко вычислить

$$J = r^2 \sin\theta. \tag{8.2}$$

Отсюда следует, что преобразование координат (8.1) вырождено при $r = 0$ или $\theta = 0$, π, т.е. на оси z. Действительно, обратные преобразования,

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \theta &= \operatorname{arctg} \frac{\sqrt{x^2 + y^2}}{z}, \\ \varphi &= \operatorname{arctg} \frac{y}{x}, \end{aligned} \tag{8.3}$$

однозначно определены всюду, кроме оси z. При этом все евклидово пространство, из которого удалена полуплоскость $y = 0$, $x \geq 0$, включающая ось z, взаимно однозначно отображается на открытую область

$$0 < r < \infty, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi, \tag{8.4}$$

Если полуплоскость не удалять, то точки с координатами φ и $\varphi + 2\pi$ необходимо отождествить.

Переход к сферическим координатам сохраняет ориентацию, поскольку якобиан перехода (8.2) положителен в области (8.4).

Из формул (8.1) следуют правила преобразования дифференциалов:

$$\begin{aligned} dx &= \sin\theta \cos\varphi dr + r\cos\theta\cos\varphi d\theta - r\sin\theta\sin\varphi d\varphi, \\ dy &= \sin\theta\sin\varphi dr + r\cos\theta\sin\varphi d\theta + r\sin\theta\cos\varphi d\varphi, \\ dz &= \cos\theta dr - r\sin\theta d\theta. \end{aligned} \tag{8.5}$$

Подставляя эти выражения в евклидову метрику

$$ds^2 = dx^2 + dy^2 + dz^2, \tag{8.6}$$

получим следующее выражение для метрики в сферической системе координат

$$ds^2 = dr^2 + r^2 d\Omega, \tag{8.7}$$

где

$$d\Omega = d\theta^2 + \sin^2\theta d\varphi^2 \tag{8.8}$$

– дифференциал телесного угла. Отсюда следует, что компоненты евклидовой метрики в сферической системе координат

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \tag{8.9}$$

являются функциями, а не константами. Отметим, что определитель метрики равен квадрату якобиана преобразования координат, $\det g_{\alpha\beta} = J^2$.

Помимо инвариантной квадратичной формы декартовых дифференциалов (8.6) существует еще одна линейно с ней независимая квадратичная форма

$$(xdx + ydy + zdz)^2,$$

инвариантная относительно $\mathbb{SO}(3)$ вращений. В сферических координатах также имеются две независимые инвариантные квадратичные формы, которые даются, например, выражениями dr^2 и $d\Omega$. Между этими формами существует связь:

$$\begin{aligned} dr^2 &= \frac{1}{r^2}(xdx + ydy + zdz)^2, \\ d\Omega &= \frac{1}{r^2}(dx^2 + dy^2 + dz^2) - \frac{1}{r^4}(xdx + ydy + zdz)^2. \end{aligned} \tag{8.10}$$

Обсудим, какие структуры существуют в евклидовом пространстве $\mathbb{R}^3$, инвариантные относительно $\mathbb{SO}(3)$ вращений. Общая теория инвариантных структур будет рассмотрена позже в разделе 10.3.

Из декартовых координат можно составить только одну функционально независимую сферически симметричную комбинацию – это радиус $r = \sqrt{x^2+y^2+z^2}$. Нетрудно показать, что любая сферически симметричная функция или плотность f может зависеть только от радиуса, $f = f(r)$.

Сферически симметричное векторное поле в сферической системе координат имеет только одну нетривиальную компоненту, направленную по радиусу,

$$X = \{X^r(r) := f(r), 0, 0\}.$$

В декартовой системе координат оно имеет вид

$$X = \{xf, yf, xf\}$$

или

$$X^i = x^i f(r),$$

где $\{x^i\} = \{x, y, z\}$, $i = 1, 2, 3$, а $f(r)$ – некоторая функция только от радиуса.

Инвариантный симметричный тензор второго ранга удобно записывать в декартовой системе координат, выделив явно след и бесследовую часть (неприводимые компоненты):

$$X^{ij} = \delta^{ij} f_1(r) + \left(\delta^{ij} - 3\frac{x^i x^j}{r^2}\right) f_2(r),$$

где δ^{ij} – евклидова метрика. Первое слагаемое в этом разложении пропорционально следу тензора $X^{ij}\delta_{ij} = 3f_1$, а второе слагаемое имеет нулевой след. То есть сферически инвариантный тензор второго ранга взаимно однозначно определяется двумя функциями только от радиуса: $f_1(r)$ и $f_2(r)$.

Антисимметричный тензор второго ранга $Y^{ij} = -Y^{ji}$, инвариантный относительно полной группы вращений $\mathbb{O}(3)$, определяется одной псевдоскалярной функцией $f^*(r)$ от радиуса:

$$Y^{ij} = \varepsilon^{ijk} x_k f^*(r),$$

где ε^{ijk} – полностью антисимметричный псевдотензор третьего ранга. При отражении одной из декартовых осей координат ε^{ijk} и f^* меняют знаки, оставляя антисимметричный тензор без изменения.

Дифференциал телесного угла (8.8) задает метрику на сфере единичного радиуса с центром в начале координат. С топологической точки зрения евклидово пространство $\mathbb{R}^3$ с выколотым началом координат является прямым произведением сферы $\mathbb{S}^2$ и положительной полуоси $\mathbb{R}_+$:

$$\mathbb{R}^3 \backslash \{0\} = \mathbb{S}^2 \times \mathbb{R}_+.$$

Приведем явные формулы преобразования частных производных (координатного базиса векторных полей), которые часто используются в приложениях:

$$\begin{aligned}
\partial_r &= \sin\theta\cos\varphi\partial_x + \sin\theta\sin\varphi\partial_y + \cos\theta\partial_z,\\
\partial_\theta &= r\cos\theta\cos\varphi\partial_x + r\cos\theta\sin\varphi\partial_y - r\sin\theta\partial_z,\\
\partial_\varphi &= -r\sin\theta\sin\varphi\partial_x + r\sin\theta\cos\varphi\partial_y,
\end{aligned}$$

и их обратные:

$$\begin{aligned}
\partial_x &= \sin\theta\cos\varphi\partial_r + \frac{\cos\theta\cos\varphi}{r}\partial_\theta - \frac{\sin\varphi}{r\sin\theta}\partial_\varphi,\\
\partial_y &= \sin\theta\sin\varphi\partial_r + \frac{\cos\theta\sin\varphi}{r}\partial_\theta + \frac{\cos\varphi}{r\sin\theta}\partial_\varphi,\\
\partial_z &= \cos\theta\partial_r - \frac{\sin\theta}{r}\partial_\theta.
\end{aligned}$$

В приложениях тензорные поля часто рассматриваются в ортонормированном базисе векторных полей и 1-форм:

$$\begin{aligned}
e_r &:= \partial_r, \qquad & e_\theta &:= \frac{1}{r}\partial_\theta, \qquad & e_\varphi &:= \frac{1}{r\sin\theta}\partial_\varphi,\\
e^r &:= dr, \qquad & e^\theta &:= rd\theta, \qquad & e^\varphi &:= r\sin\theta d\varphi,
\end{aligned} \tag{8.11}$$

что следует непосредственно из вида метрики (8.7). Этому базису соответствуют следующие компоненты ортонормированного репера и его обратного:

$$\begin{aligned}
{e_r}^{\hat r} &= 1, \qquad & {e_\theta}^{\hat\theta} &= r, \qquad & {e_\varphi}^{\hat\varphi} &= r\sin\theta,\\
{e^r}_{\hat r} &= 1, \qquad & {e^\theta}_{\hat\theta} &= \frac{1}{r}, \qquad & {e^\varphi}_{\hat\varphi} &= \frac{1}{r\sin\theta}.
\end{aligned} \tag{8.12}$$

При этом все недиагональные компоненты репера равны нулю. Здесь и далее шляпками обозначены индексы относительно этого репера. Векторное поле можно разложить как по координатному, так и по ортонормированному базису:

$$X = X^r\partial_r + X^\theta\partial_\theta + X^\varphi\partial_\varphi = \hat X^r e_r + \hat X^\theta e_\theta + \hat X^\varphi e_\varphi,$$

причем переход от компонент X^α к $\hat X^\alpha$ осуществляется с помощью репера. Отметим, что, поскольку репер ортонормирован, то компоненты векторных полей и соответствующих 1-форм совпадают:

$$\hat X_r = \hat X^r, \quad \hat X_\theta = \hat X^\theta, \quad \hat X_\varphi = \hat X^\varphi.$$

Получим явные формулы для градиента, ротора, дивергенции и лапласиана в сферической системе координат, которые используются в задачах со сферической симметрией. Чтобы получить явное выражение для инвариантных дифференциальных операторов в сферической системе координат нам понадобятся символы Кристоффеля (7.22). Они имеют девять ненулевых компонент:

$$\begin{aligned}
{\Gamma_{r\theta}}^\theta &= {\Gamma_{\theta r}}^\theta = \frac{1}{r}, \qquad & {\Gamma_{\theta\varphi}}^\varphi &= {\Gamma_{\varphi\theta}}^\varphi = \frac{\cos\theta}{\sin\theta},\\
{\Gamma_{r\varphi}}^\varphi &= {\Gamma_{\varphi r}}^\varphi = \frac{1}{r}, \qquad & {\Gamma_{\varphi\varphi}}^r &= -r\sin^2\theta,\\
{\Gamma_{\theta\theta}}^r &= -r, \qquad & {\Gamma_{\varphi\varphi}}^\theta &= -\sin\theta\cos\theta.
\end{aligned} \tag{8.13}$$

Из условия равенства нулю ковариантной производной репера (8.12)

$$\nabla_\alpha {e_\beta}^a = \partial_\alpha {e_\beta}^a - {\Gamma_{\alpha\beta}}^\gamma {e_\gamma}^a + \omega_{\alpha b}{}^a {e_\beta}^b = 0, \tag{8.14}$$

где $\omega_{\alpha b}{}^{a}$ – $\mathbb{SO}(3)$-связность (без кручения), находим, что у $\mathbb{SO}(3)$-связности только 6 компонент отличны от нуля:

$$\begin{aligned}
\omega_{\theta\hat{r}}{}^{\hat{\theta}} &= -\omega_{\theta\hat{\theta}}{}^{\hat{r}} = 1,\\
\omega_{\varphi\hat{r}}{}^{\hat{\varphi}} &= -\omega_{\varphi\hat{\varphi}}{}^{\hat{r}} = \sin\theta,\\
\omega_{\varphi\hat{\theta}}{}^{\hat{\varphi}} &= -\omega_{\varphi\hat{\varphi}}{}^{\hat{\theta}} = \cos\theta.
\end{aligned} \tag{8.15}$$

Градиент скалярного поля f дает 1-форму:

$$df = e^r\partial_r f + e^\theta\frac{1}{r}\partial_\theta f + e^\varphi\frac{1}{r\sin\theta}\partial_\varphi f. \tag{8.16}$$

Для получения этой формулы в дифференциале $df = dr\partial_r f + d\theta\partial_\theta f + d\varphi\partial_\varphi f$ дифференциалы dr, $d\theta$ и $d\varphi$ необходимо заменить на ортонормированный базис e^r, e^θ и e^φ с помощью соотношений (8.11).

Дивергенция вектора (7.54) дает функцию

$$\nabla_\alpha X^\alpha = \frac{1}{r^2}\partial_r(r^2\hat{X}^r) + \frac{1}{r\sin\theta}\partial_\theta(\sin\theta\hat{X}^\theta) + \frac{1}{r\sin\theta}\partial_\varphi\hat{X}^\varphi. \tag{8.17}$$

Эта и последующие формулы получаются подстановкой символов Кристоффеля (8.13) в ковариантные производные и выражением компонент вектора в координатном базисе через его компоненты в ортонормированном базисе.

В инвариантном виде ротор 1-формы $A = dx^\alpha A_\alpha$ определяется оператором Ходжа (4.49) и внешним дифференцированием (4.33):

$$\mathrm{rot}\,A = *dA = e^r\hat{B}_r + e^\theta\hat{B}_\theta + e^\varphi\hat{B}_\varphi, \tag{8.18}$$

где компоненты разложения задаются формулами:

$$\begin{aligned}
\hat{B}_r = \hat{B}^r &= \frac{1}{r\sin\theta}\partial_\theta(\sin\theta\hat{A}_\varphi) - \frac{1}{r\sin\theta}\partial_\varphi\hat{A}_\theta,\\
\hat{B}_\theta = \hat{B}^\theta &= \frac{1}{r\sin\theta}\partial_\varphi\hat{A}_r - \frac{1}{r}\partial_r(r\hat{A}_\varphi),\\
\hat{B}_\varphi = \hat{B}^\varphi &= \frac{1}{r}\partial_r(r\hat{A}_\theta) - \frac{1}{r}\partial_\theta\hat{A}_r.
\end{aligned} \tag{8.19}$$

Лапласиан функции (7.61) дает также функцию:

$$\triangle f = \frac{1}{r^2}\partial_r(r^2\partial_r f) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta f) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2 f. \tag{8.20}$$

Лапласиан 1-формы снова дает 1-форму: $\triangle\hat{A}_\alpha = g^{\beta\gamma}\widetilde{\nabla}_\beta\widetilde{\nabla}_\gamma\hat{A}_\alpha$. После громоздких вычислений можно получить явные выражения для компонент:

$$\begin{aligned}
\triangle\hat{A}_r &= \frac{1}{r^2}\partial_r(r^2\partial_r\hat{A}_r) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta\hat{A}_r) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2\hat{A}_r - \frac{2}{r^2}\hat{A}_r\\
&\quad - \frac{2}{r^2\sin\theta}\partial_\theta(\sin\theta\hat{A}_\theta) - \frac{2}{r^2\sin\theta}\partial_\varphi\hat{A}_\varphi,\\
\triangle\hat{A}_\theta &= \frac{1}{r}\partial_r^2(r\hat{A}_\theta) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta\hat{A}_\theta) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2\hat{A}_\theta - \frac{1}{r^2\sin^2\theta}\hat{A}_\theta\\
&\quad + \frac{2}{r^2}\partial_\theta\hat{A}_r - \frac{2\cos\theta}{r^2\sin^2\theta}\partial_\varphi\hat{A}_\varphi,\\
\triangle\hat{A}_\varphi &= \frac{1}{r}\partial_r^2(r\hat{A}_\varphi) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta\hat{A}_\varphi) + \frac{1}{r^2\sin^2}\partial_\varphi^2\hat{A}_\varphi - \frac{1}{r^2\sin^2\theta}\hat{A}_\varphi\\
&\quad + \frac{2\cos\theta}{r^2\sin^2\theta}\partial_\varphi\hat{A}_\theta + \frac{2}{r^2\sin\theta}\partial_\varphi\hat{A}_r.
\end{aligned} \tag{8.21}$$

Отметим появление дополнительных слагаемых по сравнению с лапласианом функции (8.20).

Приведенные выше формулы для дифференциальных операторов можно также получить путем элементарных, но громоздких геометрических построений. Такой подход эквивалентен введенному ранее понятию ковариантной производной. Это показывает, что дифференциальная геометрия предоставляет естественный и конструктивный подход к получению явного вида дифференциальных операторов в различных криволинейных системах координат. Разумеется, геометрия пространства при этом остается евклидовой.

8.2 Цилиндрические координаты

Рассмотрим *цилиндрические координаты* r, φ, z в трехмерном евклидовом пространстве $\mathbb{R}^3$, рис. 8.2. При переходе к цилиндрическим координатам мы вводим полярные координаты на плоскости x, y, оставляя координату z без изменений. Функции перехода и их обратные имеют вид (за третьей координатой z можно не следить):

$$\begin{aligned} x &= r\cos\varphi, & r &= \sqrt{x^2+y^2}, \\ y &= r\sin\varphi, & \varphi &= \operatorname{arctg}(y/x), \end{aligned}$$

Угол φ, как и в случае сферических координат, называется *полярным*. Якобиан преобразования координат равен $J = r$ и является вырожденным в начале полярных координат, что соответствует оси z. Область определения цилиндрических координат задается неравенствами:

$$0 < r < \infty, \quad 0 < \varphi < 2\pi, \quad -\infty < z < \infty,$$

что соответствует евклидову пространству с удаленной полуплоскостью $y = 0$, $x \geq 0$.

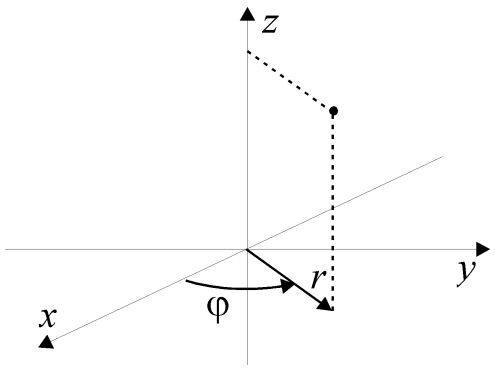


Рис. 8.2: Цилиндрические координаты r, φ, z в трехмерном евклидовом пространстве $\mathbb{R}^3$.

Дифференциалы декартовых и полярных координат (координатные базисы 1-форм) связаны преобразованием:

$$\begin{aligned} dx &= \cos\varphi dr - r\sin\varphi d\varphi, & dr &= \cos\varphi dx + \sin\varphi dy, \\ dy &= \sin\varphi dr + r\cos\varphi d\varphi, & d\varphi &= -\frac{\sin\varphi}{r}dx + \frac{\cos\varphi}{r}dy. \end{aligned} \tag{8.22}$$

Частные производные в декартовых и полярных координатах (координатные базисы векторных полей) связаны простыми соотношениями:

$$\begin{aligned} \partial_x &= \cos\varphi\partial_r - \frac{\sin\varphi}{r}\partial_\varphi, & \partial_r &= \cos\varphi\partial_x + \sin\varphi\partial_y, \\ \partial_y &= \sin\varphi\partial_r + \frac{\cos\varphi}{r}\partial_\varphi, & \partial_\varphi &= -r\sin\varphi\partial_x + r\cos\varphi\partial_y. \end{aligned} \tag{8.23}$$

Прямая проверка показывает, что справедлива формула

$$(xdx + ydy)^2 = r^2dr^2.$$

Соотношения (8.22) приводят к следующему выражению для евклидовой метрики (8.6) в цилиндрических координатах:

$$ds^2 = dr^2 + r^2d\varphi^2 + dz^2.$$

В матричных обозначениях метрика и ее обратная имеют вид

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad g^{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Из символов Кристоффеля только три отличны от нуля:

$$\Gamma_{r\varphi}{}^{\varphi} = \Gamma_{\varphi r}{}^{\varphi} = \frac{1}{r}, \qquad \Gamma_{\varphi\varphi}{}^{r} = -r.$$

Ортонормированный базис векторных полей и 1-форм имеет вид (для полярных координат)

$$\begin{aligned} e_r &= \partial_r, & e_\varphi &= \frac{1}{r}\partial_\varphi, \\ e^r &= dr, & e^\varphi &= rd\varphi. \end{aligned}$$

Этому базису соответствует ортонормированный репер

$$\begin{aligned} e_r{}^{\hat{r}} &= 1, & e_\varphi{}^{\hat{\varphi}} &= r, \\ e^r{}_{\hat{r}} &= 1, & e^\varphi{}_{\hat{\varphi}} &= \frac{1}{r}. \end{aligned}$$

При этом все недиагональные компоненты равны нулю.

Из условия (8.14) находим компоненты $\mathbb{SO}(2)$-связности. Несложные вычисления показывают, что только две компоненты отличны от нуля:

$$\omega_{\varphi\hat{r}}{}^{\hat{\varphi}} = -\omega_{\varphi\hat{\varphi}}{}^{\hat{r}} = 1.$$

Вычисления, аналогичные случаю сферических координат, приводят к следующим выражениям для простейших дифференциальных операторов в цилиндриче-

ских координатах

$$df = e^r \partial_r f + e^\varphi \frac{1}{r} \partial_\varphi f + e^z \partial_z f, \tag{8.24}$$

$$\nabla_\alpha X^\alpha = \frac{1}{r} \partial_r (r \hat{X}^r) + \frac{1}{r} \partial_\varphi \hat{X}^\varphi + \partial_z \hat{X}^z, \tag{8.25}$$

$$\mathsf{rot}\, A = e^r \left[\frac{1}{r} \partial_\varphi \hat{A}_z - \partial_z \hat{A}_\varphi \right] + e^\varphi [\partial_z \hat{A}_r - \partial_r \hat{A}_z] + e^z \frac{1}{r} [\partial_r (r \hat{A}_\varphi) - \partial_\varphi \hat{A}_r], \tag{8.26}$$

$$\triangle f = \frac{1}{r} \partial_r (r \partial_r f) + \frac{1}{r^2} \partial_\varphi^2 f + \partial_z^2 f, \tag{8.27}$$

$$\triangle \hat{A}_r = \frac{1}{r} \partial_r (r \partial_r \hat{A}_r) + \frac{1}{r^2} \partial_\varphi^2 \hat{A}_r + \partial_z^2 \hat{A}_r - \frac{1}{r^2} \hat{A}_r - \frac{2}{r^2} \partial_\varphi \hat{A}_\varphi, \tag{8.28}$$

$$\triangle \hat{A}_\varphi = \frac{1}{r} \partial_r (r \partial_r \hat{A}_\varphi) + \frac{1}{r^2} \partial_\varphi^2 \hat{A}_\varphi + \partial_z^2 \hat{A}_\varphi - \frac{1}{r^2} \hat{A}_\varphi + \frac{2}{r^2} \partial_\varphi \hat{A}_r, \tag{8.29}$$

$$\triangle \hat{A}_z = \frac{1}{r} \partial_r (r \partial_r \hat{A}_z) + \frac{1}{r^2} \partial_\varphi^2 \hat{A}_z + \partial_z^2 \hat{A}_z. \tag{8.30}$$

Напомним, что шляпки над символами означают, что компоненты (ко-)векторов рассматриваются относительно ортонормированного базиса.

Глава 9

Группы Ли

Группы Ли образуют один из наиболее важных классов многообразий и имеют широкие приложения в математической физике. В настоящем разделе мы дадим формальные определения, опишем основные свойства, а также рассмотрим группы Ли с дифференциально геометрической точки зрения. В частности, многообразия полупростых групп Ли будут рассмотрены, как римановы пространства и пространства абсолютного параллелизма, когда групповая операция отождествляется с параллельным переносом.

9.1 Группы Ли и локальные группы Ли

Определение. Множество элементов $a, b, c, \ldots \in \mathbb{G}$ называется *группой Ли*, если оно является одновременно и группой, и гладким многообразием, при этом требуется, чтобы была определена гладкая групповая операция

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \;\mapsto\; ab \in \mathbb{G}, \qquad \forall a, b \in \mathbb{G},$$

со следующими свойствами:

1) Ассоциативность: $(ab)c = a(bc), \quad \forall a, b, c \in \mathbb{G}$.
2) Существование единицы $e \in \mathbb{G}$: $ea = ae = a, \quad \forall a \in \mathbb{G}$.
3) Для каждого элемента $a \in \mathbb{G}$ существует обратный элемент $a^{-1} \in \mathbb{G}$ такой, что $a^{-1}a = aa^{-1} = e$, причем отображение $\mathbb{G} \ni a \mapsto a^{-1} \in \mathbb{G}$ гладкое. □

Замечание. В определении оба условия гладкости отображений:

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \;\mapsto\; ab \in \mathbb{G}, \tag{9.1}$$

$$\mathbb{G} \ni a \;\mapsto\; a^{-1} \in \mathbb{G}, \tag{9.2}$$

можно объединить в эквивалентное условие гладкости одного отображения:

$$\mathbb{G} \times \mathbb{G} \ni (a, b) \;\mapsto\; ab^{-1} \in \mathbb{G}. \tag{9.3}$$

То, что (9.3) следует из (9.1) и (9.2), очевидно. Обратно. Из (9.3) при $a = e$ следует гладкость отображения (9.2). Гладкость отображения (9.1) следует из (9.3) после гладкого отображения $b^{-1} \mapsto b$. □

Перефразируем определение: группой Ли является группа, снабженная гладкой структурой многообразия. При этом не всякая группа допускает существование такой структуры.

Каждому элементу группы $b \in \mathbb{G}$ поставим в соответствие отображение

$$\mathbb{G} \ni a \mapsto ab \in \mathbb{G}, \quad \forall b \in \mathbb{G}. \tag{9.4}$$

Тогда из определения группы Ли следует, что это отображение взаимно однозначно и гладко, т.е. является диффеоморфизмом $\mathbb{G} \xrightarrow{b} \mathbb{G}$

Замечание. Если на произвольной группе задать дискретную топологию, то она будет группой Ли, рассматриваемой, как 0-мерное многообразие (если отбросить требование счетности базы топологии). В дальнейшем, чтобы исключить подобные ситуации, при рассмотрении групп Ли мы не будем рассматривать 0-мерные многообразия. С этой точки зрения множество элементов группы Ли не может быть счетным, т.к. множество точек области в $\mathbb{R}^n$ несчетно. □

Пример 9.1.1. Группа перестановок конечного числа элементов не является группой Ли, т.к. содержит конечное число элементов. □

Можно сказать также, что группой Ли является многообразие, на котором задана групповая операция. При этом не на всяком многообразии можно определить гладкую бинарную операцию, удовлетворяющую групповым аксиомам.

Пример 9.1.2. Двумерная сфера $\mathbb{S}^2$ не может быть наделена групповой структурой. Действительно, зафиксируем произвольный отличный от нуля вектор в касательном пространстве к единице группы $\mathbb{T}_e(\mathbb{G})$. С помощью дифференциала отображения (9.4) этот вектор можно разнести по всему групповому многообразию. В результате получим гладкое векторное поле на $\mathbb{G}$, которое всюду отлично от нуля, т.к. отображение (9.4) является диффеоморфизмом. Это противоречит теореме 11.2.1, утверждающей, что на двумерной сфере не существует непрерывного отличного от нуля векторного поля. □

Поскольку группы Ли являются многообразиями, то на них без изменения переносится большинство характеристик многообразий.

Определение. *Размерностью* группы Ли называется ее размерность как многообразия. Группа Ли называется *компактной*, если групповое многообразие компактно. □

Из непрерывности групповой операции следует, что, если подмножество $\mathbb{U} \subset \mathbb{G}$ – открыто или замкнуто в $\mathbb{G}$, то таковым же будет подмножество $\mathbb{U}^{-1}$, состоящее из обратных элементов, а также подмножества $a\mathbb{U}$ и $\mathbb{U}a$ для всех $a \in \mathbb{G}$. Отсюда вытекает, что для открытого подмножества $\mathbb{U}$ подмножества $\mathbb{V}\mathbb{U}$ и $\mathbb{U}\mathbb{V}$ также открыты в $\mathbb{G}$ каково бы ни было подмножество $\mathbb{V} \in \mathbb{G}$.

Замечание. В определении группы Ли мы потребовали гладкость т.е. бесконечную дифференцируемость групповой операции и, соответственно, гладкость группы Ли, как многообразия. На самом деле ситуация следующая. Согласно теореме Глисона–Монтгомери–Циппина [?, ?] на всякой непрерывной группе (класса $\mathcal{C}^0$) можно ввести структуру аналитического многообразия (класса $\mathcal{C}^\omega$), совместимую с групповой структурой. Эта теорема дает положительное решение пятой проблемы Гильберта: следует ли из существования в группе $\mathbb{G}$ каких-нибудь координат существование в ней дифференцируемых координат. Таким образом, не ограничивая общности, можно считать, что группы Ли являются вещественно аналитическими многообразиями, а групповая операция – вещественно аналитическое отображение. □

Таким образом, группы Ли представляют собой специальный класс групп, для которых групповая операция может быть записана в координатной форме. Это важное свойство позволяет применить методы математического анализа для исследования свойств групп Ли.

Приведем несколько примеров групп Ли.

Пример 9.1.3. Прямая $\mathbb{R}$, на которой определены сдвиги (сложение) $a + b = c$, где $a, b, c \in \mathbb{R}$, является одномерной связной односвязной некомпактной абелевой группой Ли. □

Пример 9.1.4. Проколатая прямая $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ является группой Ли по отношению к умножению, $xy = z$, где $x, y, z \in \mathbb{R}_0$. Эта группа одномерна абелева несвязна и состоит из двух компонент. Каждая компонента некомпактна. Связная компонента единицы $\mathbb{R}_+$ (положительные числа) изоморфна группе вещественных чисел по сложению, $\mathbb{R}_+ \simeq \mathbb{R}$. Изоморфизм задается показательной функцией с произвольным положительным основанием, отличным от единицы. Например, $x = \mathsf{e}^a$. □

Пример 9.1.5. Евклидово пространство $\mathbb{R}^n$ (как и произвольное векторное пространство) является n-мерной связной односвязной некомпактной абелевой группой Ли относительно сложения векторов. □

Пример 9.1.6. Единичная окружность $\mathbb{S}^1$, точками которой являются комплексные числа $z = e^{i\varphi}$ является абелевой компактной группой Ли по отношению к умножению. Она обозначается через $\mathbb{U}(1)$ (унитарные матрицы размера 1×1) и изоморфна группе собственных двумерных вращений $\mathbb{SO}(2)$. Эта группа связна, но не односвязна. Ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$. □

Пример 9.1.7. Трехмерная сфера $\mathbb{S}^3$ в пространстве кватернионов, точками которой являются кватернионы q, для которых $|q| = 1$, может быть снабжена групповой структурой (см. раздел 2.8). Это – трехмерная связная односвязная неабелева компактная группа Ли, которая изоморфна группе унитарных матриц размерности 2×2 с единичным определителем $\mathbb{SU}(2)$. □

Можно показать, что сферы только двух размерностей $\mathbb{S}^1$ и $\mathbb{S}^3$ исчерпывают все сферы на которых можно задать структуру группы Ли.

В рассмотренных примерах есть две одномерные абелевы группы Ли: $\mathbb{R}$ и $\mathbb{S}^1$. Любая одномерная группа Ли является абелевой, и изоморфна либо вещественной прямой $\mathbb{R}$, либо окружности $\mathbb{S}^1$. Этот результат имеет широкие применения, т.к. каждая группа Ли имеет одномерные подгруппы.

В случае одномерных групп Ли можно описать все возможные гомоморфизмы между ними:

$$\begin{aligned}
&\mathbb{R} \ni x \;\mapsto\; ax \in \mathbb{R}, && a \in \mathbb{R};\\
&\mathbb{R} \ni x \;\mapsto\; \mathsf{e}^{iax} \in \mathbb{S}^1, && a \in \mathbb{R};\\
&\mathbb{S}^1 \ni z \;\mapsto\; 0 \in \mathbb{R};\\
&\mathbb{S}^1 \ni z \;\mapsto\; z^n \in \mathbb{S}^1, && n \in \mathbb{Z};
\end{aligned}$$

где x – точка вещественной прямой, z – комплексное число, равное по модулю единице, а числа a и n параметризуют гомоморфизмы.

Прямое произведение групп Ли $\mathbb{G}_1 \times \mathbb{G}_2$ является группой Ли с дифференцируемой структурой прямого произведения многообразий и групповой структурой прямого произведения групп.

Пример 9.1.8. Тор $\mathbb{T}^n = \underbrace{\mathbb{S}^1 \times \ldots \times \mathbb{S}^1}_{n} = \underbrace{\mathbb{U}(1) \times \ldots \times \mathbb{U}(1)}_{n}$ для всех n является n-мерной абелевой компактной группой Ли как прямое произведение окружностей. Эта группа связна, но не односвязна. □

С топологической точки зрения группы Ли $\mathbb{G}$ могут состоять из нескольких несвязных компонент[1]. При этом каждая компонента имеет одинаковую размерность, равную размерности группы. Связная компонента единицы $\mathbb{G}_0$ является открытой подгруппой в $\mathbb{G}$ и представляет собой нормальный делитель. Остальные компоненты представляют собой смежные классы по этой подгруппе, и фактор группа $\mathbb{G}/\mathbb{G}_0$ состоит не более, чем из счетного числа элементов. Если группу Ли рассматривать, как группу преобразований, то каждая связная компонента представляет собой орбиту произвольного элемента из этой компоненты относительно действия элементов группы из связной компоненты единицы $\mathbb{G}_0$.

Поскольку группа Ли является многообразием, то групповое умножение можно записать в координатах. Рассмотрим связную компоненту единицы $\mathbb{G}_0$ группы Ли размерности N. Пусть в окрестности единицы задана некоторая система координат $\{a^{\mathrm{A}}\}$, $\mathrm{A} = 1, \ldots, \mathrm{N}$. Не ограничивая общности, будем считать, что начало координат $a^{\mathrm{A}} = 0$ совпадает с единицей группы. Тогда групповое умножение

$$c = ab = f(a, b), \qquad a, b, c \in \mathbb{G}_0, \tag{9.5}$$

и групповые аксиомы (см. с. 2.2) можно записать в координатном виде:

$$\begin{array}{lll}
1) & c^{\mathrm{A}} = f^{\mathrm{A}}(a, b) & \text{– закон композиции,} \\
2) & f^{\mathrm{A}}\Big(a, f(b, c)\Big) = f^{\mathrm{A}}\Big(f(a, b), c\Big) & \text{– ассоциативность,} \\
3) & a^{\mathrm{A}} = f^{\mathrm{A}}(a, 0) = f^{\mathrm{A}}(0, a) & \text{– существование единицы,} \\
4) & f^{\mathrm{A}}(a, a^{-1}) = f^{\mathrm{A}}(a^{-1}, a) = 0 & \text{– существование обратного элемента.}
\end{array} \tag{9.6}$$

Набор N функций f^{A} от 2N переменных, удовлетворяющих условиям (9.6), называется *функцией композиции* для группы Ли $\mathbb{G}$.

Связная компонента группы Ли $\mathbb{G}_0$ может оказаться нетривиальным многообразием и не покрываться одной картой. Поэтому введем новое понятие.

Определение. *Локальной группой Ли* называется пара $(\mathbb{U}, f)$, где $\mathbb{U} \subset \mathbb{R}^{\mathrm{N}}$ – область евклидова пространства, содержащая начало координат, и f – гладкое отображение $\mathbb{U} \times \mathbb{U} \to \mathbb{R}^{\mathrm{N}}$ такое, что выполнены условия (9.6). При этом мы предполагаем, что все равенства выполнены в той области, где определены все выражения, входящие в данное равенство. □

Замечание. В определении локальной группы Ли мы не предполагаем замкнутости области $\mathbb{U}$ относительно группового умножения, т.е. возможно существование таких точек в $\mathbb{U}$, что их произведение не лежит в $\mathbb{U}$. Конечно, для таких точек про ассоциативность в определении ничего не говорится. □

Каждая группа Ли $\mathbb{G}$ порождает бесконечное множество локальных групп Ли следующим образом. Пусть $(\mathbb{V}, \varphi)$ – карта на $\mathbb{G}$, содержащая единицу, которая отображается в начало координат евклидова пространства $\mathbb{R}^{\mathrm{N}}$. Возьмем такую окрестность $\mathbb{U} \subset \mathbb{V}$, что выполнены включения $\mathbb{U} \cdot \mathbb{U} \subset \mathbb{V}$ и $\mathbb{U}^{-1} \subset \mathbb{V}$, где под произведением

[1] Поскольку мы предполагаем, что многообразие обладает счетной базой, то число несвязных компонент у групп Ли может быть не более, чем счетно.

$\mathbb{U}\cdot\mathbb{U}$ мы понимаем множество всех произведений ab, где $a,b\in\mathbb{U}$. Существование такой окрестности следует из непрерывности отображений $(a,b)\mapsto ab$ и $a\mapsto a^{-1}$. Тогда каждая пара $(\mathbb{U},f)$, где мы отождествили область $\mathbb{U}\subset\mathbb{G}$ с ее образом $\varphi(\mathbb{U})\in\mathbb{R}^{\mathrm{N}}$, а f – функция композиции на $\mathbb{G}$, записанная в координатах, является локальной группой Ли.

Определение. Две локальные группы Ли $(\mathbb{U}_1,f_1)$ и $(\mathbb{U}_2,f_2)$ называются *изоморфными*, если существуют такие окрестности нуля $\mathbb{U}'_1\subset\mathbb{U}_1$ и $\mathbb{U}'_2\subset\mathbb{U}_2$ и такой диффеоморфизм $h:\ \mathbb{U}'_1\to\mathbb{U}'_2$, что диаграмма

$$\begin{array}{ccc}
\mathbb{U}'_1\times\mathbb{U}'_1 & \xrightarrow{f_1} & \mathbb{R}^{\mathrm{N}}\supset\mathbb{U}'_1\\
\Big\downarrow{\scriptstyle h\times h} & & \Big\downarrow{\scriptstyle h}\\
\mathbb{U}'_2\times\mathbb{U}'_2 & \xrightarrow{f_2} & \mathbb{R}^{\mathrm{N}}\supset\mathbb{U}'_2
\end{array}$$

коммутативна, т.е. $h\Big(f_1(a,b)\Big)=f_2\Big(h(a),h(b)\Big)$ при $a,b\in\mathbb{U}'_1$, там, где обе части равенства определены. □

Ясно, что все локальные группы Ли, полученные описанным выше способом из одной группы Ли, изоморфны между собой. Поэтому введем

Определение. Две группы Ли $\mathbb{G}_1$ и $\mathbb{G}_2$ называются *локально изоморфными*, если порождаемые ими локальные группы Ли изоморфны. □

Конечно, две изоморфные группы Ли являются также локально изоморфными. Нетривиальность данного выше определения заключается в том, что существуют локально изоморфные группы Ли, которые в то же время не изоморфны между собой.

Пример 9.1.9. Пусть $\mathbb{G}_1=\mathbb{R}$ и $\mathbb{G}_2=\mathbb{S}^1$. В качестве $\mathbb{U}'_1\subset\mathbb{R}$ возьмем интервал $(-\pi,\pi)$, а в качестве $\mathbb{U}'_2\subset\mathbb{S}^1$ – дополнение к точке $-1\in\mathbb{S}^1$. Отображение

$$h:\ \ \mathbb{R}\supset\mathbb{U}'_1\ni x\ \mapsto\ \mathsf{e}^{ix}\in\mathbb{U}'_2\subset\mathbb{S}^1$$

устанавливает диффеоморфизм между $\mathbb{U}'_1$ и $\mathbb{U}'_2$, перестановочный с умножением. Таким образом, группы Ли $\mathbb{R}$ и $\mathbb{S}^1$ локально изоморфны. В то же время они не изоморфны между собой. □

Существует определенная связь между группами Ли, локальными группами Ли и алгебрами Ли, которые будут определены в следующем разделе 9.2. Сейчас мы приведем несколько утверждений об этой связи.

Теорема 9.1.1 (Э. Картан). *Каждая алгебра Ли $\mathfrak{g}$ является алгеброй Ли некоторой группы Ли $\mathbb{G}$.*

Важность понятия локальной группы Ли заключается в следующем утверждении, которое является следствием теоремы Картана. ???

Теорема 9.1.2. *Каждая алгебра Ли является алгеброй Ли некоторой локальной группы Ли. Локальные группы Ли изоморфны тогда и только тогда, когда изоморфны их алгебры Ли.*

Доказательство. Приведено в следующем разделе. □

Другими словами, каждая алгебра Ли определяет локальную группу Ли с точностью до изоморфизма, но не как ни группу Ли в целом.

При описании связи между группами Ли, локальными группами Ли и алгебрами Ли используется

Теорема 9.1.3 (о монодромии). *Пусть $\mathbb{G}$ – связная односвязная группа Ли, и $\mathbb{H}$ – произвольная группа Ли. Тогда всякий локальный гомоморфизм $\mathbb{G}$ в $\mathbb{H}$ (т.е. гомоморфизм соответствующих локальных групп), однозначно продолжается до глобального гомоморфизма $\mathbb{G}$ в $\mathbb{H}$.*

Доказательство. См., например, [**?**]. □

Теорема 9.1.4. *Каждой алгебре Ли $\mathfrak{g}$ соответствует единственная, с точностью до изоморфизма, связная односвязная группа Ли $\mathbb{G}$ (универсальная накрывающая), для которой $\mathfrak{g}$ является алгеброй Ли. Все связные группы Ли, для которых $\mathfrak{g}$ является алгеброй Ли имеют вид $\mathbb{G}/\mathbb{D}$, где $\mathbb{D}$ – дискретный нормальный делитель, лежащий в центре группы Ли $\mathbb{G}$.*

Доказательство. См., например, [**?**] □

Напомним, что центром группы $\mathbb{G}$ называется множество всех элементов группы, которые перестановочны со всеми элементами из $\mathbb{G}$. Слово дискретный в сформулированной теореме означает, что множество $\mathbb{D} \subset \mathbb{G}$ состоит из отдельных точек.

Структура группы Ли накладывает жесткие условия на функцию композиции f^{A}. Как уже было отмечено, функция композиции $f^{\mathrm{A}}(a,b)$ вещественно аналитична (т.е. разлагается в сходящиеся степенные ряды) во всех точках группового многообразия по всем $2\mathrm{N}$ переменным.

Пример 9.1.10. Рассмотрим группу невырожденных $n \times n$ матриц $\mathbb{GL}(n,\mathbb{R})$. Произвольную матрицу $A \in \mathbb{GL}(n,\mathbb{R})$ можно представить в виде

$$A = \mathbb{1} + \left(A_a{}^b\right), \qquad a,b = 1,\ldots,n,$$

где $\mathbb{1}$ – единичная матрица, и $A_a{}^b$ – набор n^2 чисел, которые мы примем за координаты матрицы A. Полученное таким образом отображение всей группы $\mathbb{GL}(n,\mathbb{R})$ на область евклидова пространства $\mathbb{R}^{n^2}$ переводит единицу группы в начало координат. Групповое многообразие при этом представляет собой открытую область в $\mathbb{R}^{n^2}$, которая является дополнением замкнутого множества точек, определяемого уравнением $\det A = 0$. Функция композиции для группы $\mathbb{GL}(n,\mathbb{R})$ принимает вид

$$f_a{}^b(A,B) = A_a{}^b + B_a{}^b + A_a{}^c B_c{}^b.$$

Таким образом, функция композиции является аналитической функцией. Аналогично, множество невырожденных $n \times n$ комплексных матриц $\mathbb{GL}(n,\mathbb{C})$ образует аналитическую группу Ли размерности $2n^2$. □

Из условия существования единицы (9.6) следуют равенства:

$$\left.\frac{\partial f^{\mathrm{A}}}{\partial a^{\mathrm{B}}}\right|_{b=0} = \delta^{\mathrm{A}}_{\mathrm{B}}, \qquad \left.\frac{\partial f^{\mathrm{A}}}{\partial b^{\mathrm{B}}}\right|_{a=0} = \delta^{\mathrm{A}}_{\mathrm{B}}.$$

Из этих уравнений и дифференцируемости функции композиции вытекает, что вблизи единицы группы функция композиции разрешима относительно координат $a^{\rm A}$ и $b^{\rm A}$ соответствующих элементов группы cb^{-1} и $a^{-1}c$, где $c = f(a,b)$. При этом класс дифференцируемости функций $a^{\rm A}(c,b)$ и $b^{\rm A}(a,c)$ такой же, как и у функции композиции.

Разложим функцию композиции в ряд Тейлора в окрестности единицы группы с точностью до членов третьего порядка

$$f^{\rm A}(a,b) = a^{\rm A} + b^{\rm A} + a^{\rm C}b^{\rm B}\hat{f}_{\rm BC}{}^{\rm A} + a^{\rm D}a^{\rm C}b^{\rm B}g_{\rm BCD}{}^{\rm A} + a^{\rm D}b^{\rm C}b^{\rm B}h_{\rm BCD}{}^{\rm A} + \dots \tag{9.7}$$

В этом разложении $\hat{f}_{\rm BC}{}^{\rm A}$, $g_{\rm BCD}{}^{\rm D} = g_{\rm BDC}{}^{\rm A}$, $h_{\rm BCD}{}^{\rm D} = h_{\rm CBD}{}^{\rm A}$ – некоторые константы. Из существования единичного элемента следует отсутствие нулевого члена разложения и вид линейных слагаемых. Слагаемые более высоких порядков обязательно должны содержать произведение $a^{\rm A}b^{\rm B}$, т.к. в противном случае условие 3) в определении функции композиции не может быть выполнено. Подстановка разложения (9.7) в условие ассоциативности накладывает условие на компоненты разложения:

$$\hat{f}_{\rm BC}{}^{\rm E}\hat{f}_{\rm ED}{}^{\rm A} + 2h_{\rm BCD}{}^{\rm A} = \hat{f}_{\rm CD}{}^{\rm E}\hat{f}_{\rm BE}{}^{\rm A} + 2g_{\rm BCD}{}^{\rm A}. \tag{9.8}$$

При антисимметризации этого выражения по индексам B, C, D слагаемые $h_{\rm BCD}{}^{\rm A}$ и $g_{\rm BCD}{}^{\rm A}$ выпадают из-за симметрии по паре индексов. В результате получаем ограничение на *структурные константы* группы Ли,

$$f_{\rm AB}{}^{\rm C} = -f_{\rm BA}{}^{\rm C} := -\hat{f}_{\rm AB}{}^{\rm C} + \hat{f}_{\rm BA}{}^{\rm C}, \tag{9.9}$$

известные как *тождества Якоби*

$$f_{\rm AB}{}^{\rm D}f_{\rm CD}{}^{\rm E} + f_{\rm BC}{}^{\rm D}f_{\rm AD}{}^{\rm E} + f_{\rm CA}{}^{\rm D}f_{\rm BD}{}^{\rm E} = 0, \tag{9.10}$$

где слагаемые отличаются циклической перестановкой индексов A, B, C. При свертке тождеств Якоби по индексам C, E последние два слагаемых сокращаются, и получаем тождество

$$f_{\rm AB}{}^{\rm D}f_{\rm DC}{}^{\rm C} = 0.$$

Для абелевых групп функция композиции симметрична $f^{\rm A}(a,b) = f^{\rm A}(b,a)$, и структурные константы тождественно обращаются в нуль.

Координаты обратного элемента $a^{-1{\rm A}}(a)$ зависят от элемента a. Разложив это выражение по $a^{\rm A}$, и подставив в определение обратного элемента, получим следующее разложение для координат обратного элемента с точностью до членов третьего порядка

$$a^{-1{\rm A}} = -a^{\rm A} + a^{\rm C}a^{\rm B}\hat{f}_{\rm BC}{}^{\rm A} - \frac{1}{2}a^{\rm D}a^{\rm C}a^{\rm B}\hat{f}_{\rm BC}{}^{\rm E}(\hat{f}_{\rm DE}{}^{\rm A} + \hat{f}_{\rm ED}{}^{\rm A}) + \dots . \tag{9.11}$$

При получении этого выражения было использовано соотношение (9.8).

9.2 Действие группы слева

Рассмотрим группу Ли как группу преобразований группового многообразия, действующую слева. При этом каждому элементу $a \in \mathbb{G}$ ставится в соответствие отображение

$$\mathbb{G} \ni b \;\mapsto\; l_a(b) := ab \in \mathbb{G}.$$

Это отображение не является автоморфизмом группы $\mathbb{G}$, т.к. нетривиально действует на единичный элемент. Из единственности единичного элемента в группе следует, что эта группа преобразований действует свободно и транзитивно (см. раздел 10).

Функция $F(a) \in \mathcal{C}^\infty(\mathbb{G})$, заданная на групповом многообразии, будет инвариантна относительно действия группы, если выполнено условие $F(ab) = F(b)$, $\forall a \in \mathbb{G}$, и, следовательно, также для всех $b \in \mathbb{G}$. Очевидно, что инвариантные функции на группе тождественно равны константе. Другими словами, значение инвариантной функции, например, в единице $F(e)$ разносится по всему групповому многообразию с помощью действия группы слева. При этом группа Ли $\mathbb{G}$ может состоять из нескольких компонент.

Перейдем к рассмотрению левоинвариантных векторных полей на групповом многообразии. Поскольку на группе задано отображение l_a, которое является диффеоморфизмом, то в касательном пространстве определен соответствующий дифференциал отображения l_{a*}:

$$X\Big(l_a(b)\Big) = l_{a*}X(b), \tag{9.12}$$

который переводит вектор $X(b)$ в точке b в некоторый вектор $X\Big(l_a(b)\Big)$ в точке ab.

Определение. Векторное поле $X \in \mathcal{X}(\mathbb{G})$ на группе Ли $\mathbb{G}$ называется *левоинвариантным*, если выполнено условие (9.12) для всех $a \in \mathbb{G}$ и, следовательно, для всех $b \in \mathbb{G}$. □

Левоинвариантные векторные поля определены глобально, т.е. на всем групповом многообразии.

Проиллюстрируем действие дифференциала отображения для соответствующей локальной группы Ли, на которой задана функция композиции в явном виде. В компонентах дифференциал отображения записывается в виде

$$X^{\text{A}}(b) \;\mapsto\; X^{\text{A}}(c) = X^{\text{B}}(b)U_{\text{B}}{}^{\text{A}}(a,b), \qquad c = ab,$$

где $X^{\text{A}}(b)$ – компоненты касательного вектора в точке b относительно координатного базиса ∂_{A}, и дифференциал отображения задается матрицей

$$U_{\text{B}}{}^{\text{A}}(a,b) = \frac{\partial f^{\text{A}}(a,b)}{\partial b^{\text{B}}},$$

зависящей от координат двух точек $a, b \in \mathbb{G}$. Поскольку групповая операция, $(a,b) \mapsto f(a,b)$, является диффеоморфизмом группового многообразия на себя при фиксированном a или b, то матрица $U_{\text{A}}{}^{\text{B}}$ невырождена для всех $a, b \in \mathbb{G}$, $\det U_{\text{A}}{}^{\text{B}} \neq 0$.

Пусть $X_0 = \{X_0^{\text{A}}\}$ – произвольный вектор из касательного пространства к единице группы. Тогда этот вектор можно разнести по всему групповому многообразию с помощью дифференциала левого действия элементов группы. По построению, это будет левоинвариантное векторное поле $X(a) = l_{a*}X_0$. Оно определено глобально.

Компоненты левоинвариантного векторного поля вблизи единицы группы имеют вид

$$X^{\text{A}}(a) = X_0^{\text{B}} L_{\text{B}}{}^{\text{A}}(a),$$

где

$$L_{\text{B}}{}^{\text{A}}(a) := U_{\text{B}}{}^{\text{A}}(a,0) = \left.\frac{\partial f^{\text{A}}(a,b)}{\partial b^{\text{B}}}\right|_{b=0}.$$

Тем самым, каждому вектору из касательного пространства к единице однозначно ставится в соответствие левоинвариантное векторное поле, и, наоборот, каждое левоинвариантное векторное поле однозначно определяет некоторый вектор в касательном пространстве к единице группы.

Матрица $U_{\text{B}}{}^{\text{A}}(a,b)$ удовлетворяет простому тождеству. Дифференцируя равенство $f^{\text{A}}(0,b) = b^{\text{A}}$ по b^{B}, получаем тождество

$$U_{\text{B}}{}^{\text{A}}(0,b) = \delta^{\text{A}}_{\text{B}}, \tag{9.13}$$

которое нам понадобится в дальнейшем.

Множество векторных полей на группе Ли $\mathbb{G}$, как и на любом другом многообразии, образует бесконечномерную алгебру Ли $\mathcal{X}(\mathbb{G})$ (см. раздел 3.6.7) и $\mathcal{C}^\infty(\mathbb{G})$-модуль. Поскольку дифференциал отображения сохраняет скобку Ли (коммутатор) (3.73), то коммутатор двух левоинвариантных векторных полей будет левоинвариантным векторным полем. Следовательно, множество всех левоинвариантных векторных полей образует подалгебру Ли в $\mathcal{X}(\mathbb{G})$. Поскольку, как векторное пространство, множество левоинвариантных векторных полей изоморфно касательному пространству в единице группы $\mathbb{T}_e(\mathbb{G})$, то множество левоинвариантных векторных полей представляет собой конечномерную алгебру Ли такой же размерности N, как и сама группа.

Замечание. Алгебра Ли левоинвариантных векторных полей не выдерживает умножения на гладкие функции $f(a) \in \mathcal{C}^\infty(\mathbb{G})$, и поэтому представляет собой линейное пространство, а не $\mathcal{C}^\infty(\mathbb{G})$-модуль. □

Структура алгебры Ли левоинвариантных векторных полей естественным образом переносится на касательное пространство к единице группы:

$$[X_0, Y_0] := [l_{a*}X_0, l_{a*}Y_0]_e.$$

Словами. Берем два произвольных вектора X_0 и Y_0 из касательного пространства к единице группы, разносим их с помощью действия группы слева по всему многообразию, вычисляем коммутатор получившихся левоинвариантных векторных полей и выбираем в касательном пространстве к единице группы группы вектор, который соответствует коммутатору.

Множество левоинвариантных векторных полей образует N-мерное векторное пространство над полем вещественных чисел, которое изоморфно касательному пространству в начале координат (или в любой другой точке группового многообразия). В качестве базиса левоинвариантных векторных полей удобно выбрать N левоинвариантных векторных полей (*образующих алгебры Ли*)

$$L_{\text{A}}(a) := l_{a*}\partial_{\text{A}}|_e = L_{\text{A}}{}^{\text{B}}\partial_{\text{B}}|_a, \tag{9.14}$$

где мы разложили левоинвариантный базис по координатному базису $\partial_{\text{A}}|_a$ в точке $a \in \mathbb{G}$. Тем самым коэффициенты разложения $L_{\text{A}}{}^{\text{B}}(a)$ определены глобально. В построенном базисе произвольное левоинвариантное векторное поле X имеет постоянные компоненты, равные его компонентам в начале координат относительно координатного базиса:

$$X = X_0^{\text{A}} L_{\text{A}}.$$

Поскольку левоинвариантные векторные поля образуют алгебру Ли, то коммутатор двух базисных левоинвариантных векторных полей L_{A} будет левоинвариантным

векторным полем и, следовательно, его можно разложить по базису L_{A} с некоторыми постоянными коэффициентами:

$$[L_{\text{A}}, L_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} L_{\text{C}}. \tag{9.15}$$

То, что в правой части этого равенства стоят структурные константы $f_{\text{AB}}{}^{\text{C}}$, которые были определены ранее через функцию композиции (9.9) не случайно.

Докажем это, рассмотрев окрестность единицы группы и соответствующую локальную группу Ли. Вблизи единицы группы левоинвариантный базис в компонентах имеет вид

$$L_{\text{A}}(a) = L_{\text{A}}{}^{\text{B}}(a)\partial_{\text{B}},$$

где ∂_{B} – координатный базис касательных пространств, что оправдывает выбранные обозначения в (9.14). Из разложения (9.7) следует разложение для компонент левоинвариантного базиса:

$$L_{\text{A}}{}^{\text{B}} = \delta_{\text{A}}^{\text{B}} + a^{\text{C}} \hat{f}_{\text{AC}}{}^{\text{B}} + a^{\text{D}} a^{\text{C}} g_{\text{ACD}}{}^{\text{B}} + \dots . \tag{9.16}$$

Подстановка этого разложения в (9.15) доказывает, что в правой части стоят структурные константы (9.9), потому что коэффициенты разложения $f_{\text{AB}}{}^{\text{C}}$ постоянны. При этом возникают ограничения на коэффициенты разложения при высших степенях a^{A}, на которых мы останавливаться не будем.

Замечание. Для всех $a \in \mathbb{G}$ выполнено равенство $\det L_{\text{A}}{}^{\text{B}}(a) \neq 0$, потому что групповая операция задает диффеоморфизм группы $\mathbb{G}$ на себя. При малых a^{A} это следует также из (9.16). □

Продолжим рассмотрение локальной группы Ли и дадим независимый вывод формулы (9.15), исходя только из свойств функции композиции.

Предложение 9.2.1. *Базис алгебры Ли (9.14) удовлетворяет коммутационным соотношениям (9.15) с некоторым набором структурных констант, удовлетворяющих тождеству Якоби.*

Доказательство. Рассмотрев тождество $b = a^{-1}ab$ в касательном пространстве, получим равенство

$$L_{\text{A}}^{-1\,\text{B}}(a) = U_{\text{A}}{}^{\text{B}}(a^{-1}, a).$$

Далее, распишем условие ассоциативности $(ab)c = a(bc)$ в касательном пространстве:

$$U_{\text{A}}{}^{\text{B}}\Big(f(a,b), c\Big) = U_{\text{A}}{}^{\text{C}}(b,c) U_{\text{C}}{}^{\text{B}}\Big(a, f(b,c)\Big).$$

Полагая $c = 0$, получим

$$\frac{\partial f^{\text{C}}(a,b)}{\partial b^{\text{A}}} = U_{\text{A}}{}^{\text{C}}(a,b) = L_{\text{A}}^{-1\,\text{D}}(b) L_{\text{D}}{}^{\text{C}}(f), \tag{9.17}$$

где $f = f(a,b)$. Рассмотрим эту систему уравнений, как уравнения на функцию композиции. Система уравнений (9.17) разрешима тогда и только тогда, когда выполнены условия совместности. Чтобы получить эти условия, продифференцируем уравнение (9.17) по b^{B} и антисимметризуем по индексам A и B. После несложных алгебраических преобразований получим уравнение

$$[L_{\text{A}}(b), L_{\text{B}}(b)]^{\text{D}} L_{\text{D}}^{-1\,\text{C}}(b) = [L_{\text{A}}(f), L_{\text{B}}(f)]^{\text{D}} L_{\text{D}}^{-1\,\text{C}}(f),$$

где квадратные скобки обозначают коммутатор векторных полей. Поскольку левая и правая часть этого равенства рассматриваются в различных точках группового многообразия, и точки могут выбираться произвольно, то они должны быть равны константе. Эти константы проще всего определить, рассмотрев линейное приближение для левоинвариантных векторных полей (9.16). В результате получим уравнение для образующих алгебры Ли (9.15).

Тождества Якоби для структурных констант следуют из тождеств Якоби для векторных полей. □

Таким образом, мы доказали, что множество левоинвариантных векторных полей образует N-мерную алгебру Ли над полем вещественных чисел, которая является подалгеброй бесконечномерной алгебры всех векторных полей $\mathcal{X}(\mathbb{G})$.

Определение. Алгебра левоинвариантных векторных полей называется *алгеброй Ли* $\mathfrak{g}$ группы Ли $\mathbb{G}$. Базисные векторы (образующие) L_{A} алгебры Ли $\mathfrak{g}$ называются *генераторами группы Ли.* □

Алгебра Ли определяется с точностью до изоморфизма (выбора базиса левоинвариантных векторных полей) только окрестностью единицы группы. Поэтому, если две группы Ли локально изоморфны, то их алгебры Ли также изоморфны. В то же время сами группы Ли могут не быть изоморфными.

Пример 9.2.1. Группы $\mathbb{SU}(2)$, $\mathbb{SO}(3)$ и $\mathbb{O}(3)$ имеют изоморфные алгебры Ли. При этом группа $\mathbb{SU}(2)$ связна и односвязна и является универсальной накрывающей группы $\mathbb{SO}(3)$. Группа вращений $\mathbb{SO}(3)$ связна, но не односвязна. Группа $\mathbb{O}(3)$ состоит из двух неодносвязных компонент. □

Левоинвариантные векторные поля L_{A} образуют базис алгебры Ли. Этот базис определен с точностью до выбора базиса в касательном пространстве к единице группы, т.е. с точностью до действия группы $\mathbb{GL}(n,\mathbb{R})$. При этом вид структурных констант зависит от выбора базиса. Поэтому, если структурные константы алгебры Ли отличны от нуля и фиксированы, то базис, в общем случае, менять нельзя.

Выше мы показали, как групповая операция определяет алгебру Ли левоинвариантных векторных полей. Проведем обратное построение и покажем, что структурные константы (9.9) со свойством (9.10) позволяют восстановить функцию композиции $f^{\mathrm{A}}(a,b)$ в окрестности единицы. Это построение дает

Доказательство теоремы 9.1.2. Перепишем соотношение (9.15) в компонентах:

$$L_{\mathrm{A}}{}^{\mathrm{D}}\partial_{\mathrm{D}}L_{\mathrm{B}}{}^{\mathrm{C}} - L_{\mathrm{B}}{}^{\mathrm{D}}\partial_{\mathrm{D}}L_{\mathrm{A}}{}^{\mathrm{C}} = f_{\mathrm{AB}}{}^{\mathrm{D}}L_{\mathrm{D}}{}^{\mathrm{C}}. \tag{9.18}$$

Свертка этого выражения с тремя обратными матрицами L^{-1} по индексам A, B, C приводит к соотношению

$$\partial_{\mathrm{A}}L^{-1\,\mathrm{C}}_{\mathrm{B}} - \partial_{\mathrm{B}}L^{-1\,\mathrm{C}}_{\mathrm{A}} = -L^{-1\,\mathrm{E}}_{\mathrm{A}}L^{-1\,\mathrm{D}}_{\mathrm{B}}f_{\mathrm{DE}}{}^{\mathrm{C}}. \tag{9.19}$$

Это равенство является дифференциальным уравнением на матрицу $L^{-1}(a)$. Левая часть этого уравнения антисимметрична по индексам A и B и, следовательно, представляет собой 2-форму на $\mathbb{G}_0$. Для того, чтобы найти необходимые и достаточные условия интегрируемости этой системы уравнений, ее необходимо продифференцировать, скажем, по a^{D}, и антисимметризовать по индексам A, B и D. После несложных вычислений, получим, что необходимым и достаточным условием интегрируемости

уравнений (9.19) являются тождества Якоби для структурных констант. Таким образом, матрицы $L^{-1}(a)$ и, следовательно, $L(a)$ определяются структурными константами с точностью до некоторой константы, которая фиксируется условием $L_{\mathrm{A}}{}^{\mathrm{B}}(0) = \delta_{\mathrm{A}}^{\mathrm{B}}$. После этого функция композиции однозначно находится из уравнения (9.17) с фиксированными условиями в нуле. Условия интегрируемости этих уравнений, как было показано выше, опять-таки сводятся к тождествам Якоби. Таким образом, если задан произвольный набор структурных констант $f_{\mathrm{AB}}{}^{\mathrm{C}} = -f_{\mathrm{BA}}{}^{\mathrm{C}}$, удовлетворяющих тождествам Якоби (9.10), то функция композиции однозначно определяется в окрестности единицы группы и, следовательно, групповая структура задается в некоторой окрестности единицы группы $\mathbb{G}$. Таким образом, доказано, что произвольная алгебра Ли определяет локальную группу Ли с точностью до изоморфизма.

В обратную сторону: каждая локальная группа Ли определяет алгебру Ли с точностью до изоморфизма. Это сразу следует из определения алгебры Ли, как множества левоинвариантных векторных полей. □

Замечание. Явный вид функции композиции известен только в самых простейших случаях, например, для абелевой группы сложения векторов евклидова пространства $f^{\mathrm{A}} = a^{\mathrm{A}} + b^{\mathrm{A}}$. В примере 9.1.10 функция композиции была построена для группы $\mathbb{GL}(n,\mathbb{R})$. В разделе 9.7 будет явно построена функция композиции в простейшем случае двумерной неабелевой группы Ли (группа аффинных преобразований прямой). В более сложных случаях решить явно уравнения (9.19) и (9.17) не удается. Но это обычно и не нужно. В приложениях, как правило, работают с матричным представлением алгебры Ли и структурными константами. □

Определение. Пусть $X = X_0^{\mathrm{A}} L_{\mathrm{A}}{}^{\mathrm{B}}(a)\partial_{\mathrm{B}}$ – произвольное левоинвариантное векторное поле. Форма $\omega = dx^{\mathrm{A}}\omega_{\mathrm{A}}(a)$ называется *левоинвариантной,* если ее значение на левоинвариантном векторном поле равно константе,

$$\omega(X) = X_0^{\mathrm{A}} L_{\mathrm{A}}{}^{\mathrm{B}}\omega_{\mathrm{B}} = \mathsf{const}. \qquad \square$$

Из определения следует, что компоненты инвариантной 1-формы в координатном базисе имеют вид

$$\omega_{\mathrm{A}}(a) = L_{\mathrm{A}}^{-1\mathrm{B}}(a)\omega_{0\mathrm{B}}, \tag{9.20}$$

где $\omega_{0\mathrm{B}}$ – компоненты левоинвариантной 1-формы в начале координат. При работе с левоинвариантными 1-формами удобно ввести левоинвариантный базис, который дуален к левоинвариантным векторным полям: $\omega^{\mathrm{A}}(L_{\mathrm{B}}) = \delta_{\mathrm{B}}^{\mathrm{A}}$. Он имеет следующий вид в координатном базисе:

$$\omega^{\mathrm{A}}(a) = dx^{\mathrm{B}} L_{\mathrm{B}}^{-1\mathrm{A}}(a),$$

Тогда произвольная левоинвариантная 1-форма будет иметь в левоинвариантном базисе постоянные компоненты $\omega = \omega^{\mathrm{A}}\omega_{0\mathrm{A}}$.

Предложение 9.2.2 (Формула Маурера–Картана). *Пусть ω – произвольная левоинвариантная 1-форма и X, Y – два произвольных левоинвариантных векторных поля. Тогда значение внешней производной $d\omega$ на векторных полях X, Y пропорционально значению* 1*-формы ω на коммутаторе этих полей:*

$$d\omega(X,Y) = -\frac{1}{2}\omega\Big([X,Y]\Big) \tag{9.21}$$

Это тождество известно, как формула Маурера–Картана. *Ее можно переписать для левоинвариантного базиса 1-форм:*

$$d\omega^{\mathrm{A}} = -\frac{1}{2}\omega^{\mathrm{B}} \wedge \omega^{\mathrm{C}} f_{\mathrm{BC}}{}^{\mathrm{A}}. \tag{9.22}$$

Доказательство. Прямая проверка с использованием тождества (9.19). □

При рассмотрении главных расслоений нам понадобится рассматривать формы, заданные на некотором многообразии $\mathbb{M}$, со значениями в алгебре Ли $\mathfrak{g}$. Введем для таких форм понятие коммутатора. Пусть L_{A}, $\mathrm{A} = 1, \ldots, \mathrm{N}$, – базис алгебры Ли. Пусть A и B – соответственно r и s формы на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$. Тогда они имеют вид

$$A = A^{\mathrm{A}} L_{\mathrm{A}}, \quad B = B^{\mathrm{A}} L_{\mathrm{A}},$$

где $A^{\mathrm{A}} \in \Lambda_r(\mathbb{M})$ и $B^{\mathrm{A}} \in \Lambda_s(\mathbb{M})$ для всех значений индекса $\forall \mathrm{A} = 1, \ldots, \mathrm{N}$. Определим коммутатор заданных форм следующим равенством

$$[A, B] := A^{\mathrm{A}} \wedge B^{\mathrm{B}} [L_{\mathrm{A}}, L_{\mathrm{B}}] = A^{\mathrm{A}} \wedge B^{\mathrm{B}} f_{\mathrm{AB}}{}^{\mathrm{C}} L_{\mathrm{C}}. \tag{9.23}$$

Выражение в правой части равенства представляет собой $r + s$ форму на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$.

Если на многообразии $\mathbb{M}$ заданы три формы A, B и C степеней r, s и t, соответственно, со значениями в алгебре Ли $\mathfrak{g}$, то из тождеств Якоби вытекает равенство

$$(-1)^{rt}\Big[[A, B], C\Big] + (-1)^{rs}\Big[[B, C], A\Big] + (-1)^{st}\Big[[C, A], B\Big]. \tag{9.24}$$

Пример 9.2.2. Пусть A и B – 1-формы на $\mathbb{M}$ со значениями в алгебре Ли $\mathfrak{g}$ и $X, Y \in \mathcal{X}(\mathbb{M})$ – два векторных поля на $\mathbb{M}$. Тогда значение коммутатора форм на векторных полях равно

$$\begin{aligned}[A, B](X, Y) &= A^{\mathrm{A}} \wedge B^{\mathrm{A}}(X, Y)[L_{\mathrm{A}}, L_{\mathrm{B}}] = \\ &= \frac{1}{2}\Big(A^{\mathrm{A}}(X) B^{\mathrm{B}}(Y) - A^{\mathrm{A}}(Y) B^{\mathrm{B}}(X)\Big) = \frac{1}{2}[A(X), B(Y)] - \frac{1}{2}[A(Y), B(X)].\end{aligned}$$

В частности,

$$[A, A](X, Y) = [A(X), A(Y)].$$ □

Предложение 9.2.3. *Если A и B – соответственно r и s формы со значениями в алгебре Ли $\mathfrak{g}$, то справедливы следующие формулы:*

$$\begin{aligned}[A, B] &= (-1)^{rs+1}[B, A], \\ d[A, B] &= [dA, B] + (-1)^r [A, dB].\end{aligned}$$

Доказательство. Прямая проверка. □

В дальнейшем особую роль будут играть формы, заданные на группе Ли $\widetilde{\mathbb{G}}$, со значениями в ее алгебре Ли $\mathfrak{g}$. То есть в качестве многообразия $\mathbb{M}$ мы рассматриваем саму группу $\widetilde{\mathbb{G}}$, которую для отличия мы отметили знаком тильды. Другими словами, рассмотрим два экземпляра одной и той же группы Ли: $\widetilde{\mathbb{G}}$ и $\mathbb{G}$. Пусть $\omega^{*\mathrm{A}}$ – левоинвариантный базис кокасательного расслоения к $\widetilde{\mathbb{G}}$. Реализуем алгебру Ли $\mathfrak{g}$, как алгебру левоинвариантных векторных полей на $\mathbb{G}$ с базисом L_{A}.

Определение. Левоинвариантная 1-форма на группе Ли $\widetilde{\mathbb{G}}$ со значениями в алгебре Ли $\mathfrak{g}$,

$$\theta = \omega^{*\mathrm{A}} L_{\mathrm{A}}, \tag{9.25}$$

называется *канонической*. □

Если $X(\tilde{a}) \in \mathcal{X}(\widetilde{\mathbb{G}})$ – векторное поле на $\widetilde{\mathbb{G}}$, то его можно разложить по левоинвариантному базису, $X = X^{\mathrm{A}}(\tilde{a})L^*_{\mathrm{A}}$. Тогда значение канонической формы на этом поле равно

$$\theta(X)(\tilde{a}) = X^{\mathrm{A}}(\tilde{a})L_{\mathrm{A}}.$$

То есть каждой точке $\tilde{a} \in \widetilde{\mathbb{G}}$ ставится в соответствие левоинвариантное векторное поле на $\mathbb{G}$. Эта форма будет использована при изучении связностей на главных расслоениях в разделе 14.1. Формулу Маурера–Картана (9.22) можно переписать для канонической формы

$$d\theta = -\frac{1}{2}[\theta, \theta], \tag{9.26}$$

где справа стоит коммутатор в алгебре Ли $\mathfrak{g}$ и внешние произведения 1-форм $\omega^{*\mathrm{A}}$.

Предложение 9.2.4. *Любая группа Ли является $\mathbb{G}$ является ориентируемым многообразием.*

Доказательство. Левоинвариантные 1-формы ω^{A} определены на всем групповом многообразии и линейно независимы. Поэтому N-форма $\omega^1 \wedge \ldots \wedge \omega^{\mathrm{N}}$ нигде не обращается в нуль. Следовательно, согласно теореме 11.4.3, любая группа Ли ориентируема. □

9.3 Действие группы справа

Все, сказанное относительно действия группы на себя слева, естественным образом переносится на действие группы справа. Ниже мы рассмотрим отличия, которые при этом возникают. Для простоты, мы рассмотрим локальную группу Ли, т.е. окрестность единицы группы.

Рассмотрим группу Ли $\mathbb{G}$, как группу преобразований, действующую справа:

$$\mathbb{G} \ni b \;\mapsto\; r_a(b) := ba \in \mathbb{G}.$$

Пусть

$$V_{\mathrm{B}}{}^{\mathrm{A}}(a, b) = \frac{\partial f^{\mathrm{A}}(a, b)}{\partial a^{\mathrm{B}}}.$$

Дифференцируя тождество $f^{\mathrm{A}}(a, 0) = a^{\mathrm{A}}$ по a^{B}, получаем равенство

$$V_{\mathrm{B}}{}^{\mathrm{A}}(a, 0) = \delta^{\mathrm{A}}_{\mathrm{B}},$$

которое нам понадобится в дальнейшем.

Аналогично случаю левого действия группы, построим правоинвариантный базис векторных полей

$$R_{\mathrm{A}}(a) = R_{\mathrm{A}}{}^{\mathrm{B}}(a)\partial_{\mathrm{B}}, \tag{9.27}$$

где

$$R_{\mathrm{A}}{}^{\mathrm{B}}(a) := V_{\mathrm{A}}{}^{\mathrm{B}}(0, a) = \left.\frac{\partial f^{\mathrm{B}}(b, a)}{\partial b^{\mathrm{A}}}\right|_{b=0}. \tag{9.28}$$

В этом базисе произвольное правоинвариантное векторное поле имеет постоянные компоненты. Как и для левоинвариантных векторных полей, $\det R_{\mathrm{A}}{}^{\mathrm{B}}(a) \neq 0$ для всех $a \in \mathbb{G}$.

Как и в случае левого действия группы, можно доказать, что правоинвариантные векторные поля образуют N-мерную алгебру Ли над полем вещественных чисел. При этом коммутатор векторов правоинвариантного базиса,

$$[R_{\mathrm{A}}, R_{\mathrm{B}}] = -f_{\mathrm{AB}}{}^{\mathrm{C}} R_{\mathrm{C}}, \tag{9.29}$$

отличается знаком структурных констант от коммутатора левоинвариантного базиса. Алгебры Ли лево- и правоинвариантных векторных полей изоморфны, поскольку переходят друг в друга после отображения $R_{\mathrm{A}} \mapsto -R_{\mathrm{A}}$.

Формула для коммутатора правоинвариантных векторных полей (9.29) справедлива, как легко видеть, глобально на всей группе Ли.

Из разложения функции композиции (9.7) следует разложение для компонент правоинвариантных векторных полей относительно координатного базиса

$$R_{\mathrm{B}}{}^{\mathrm{A}}(a) = \delta^{\mathrm{A}}_{\mathrm{B}} + a^{\mathrm{C}} \hat{f}_{\mathrm{CB}}{}^{\mathrm{A}} + a^{\mathrm{D}} a^{\mathrm{C}} h_{\mathrm{CDB}}{}^{\mathrm{A}} + \dots . \tag{9.30}$$

Из тождества $b = baa^{-1}$, записанного в касательном пространстве, следует

$$R^{-1}_{\mathrm{B}}{}^{\mathrm{A}}(a) = V_{\mathrm{B}}{}^{\mathrm{A}}(a, a^{-1}). \tag{9.31}$$

Рассмотрим функцию $F(a) \in \mathcal{C}^\infty(\mathbb{G})$ на группе Ли $\mathbb{G}$. Подействуем на точки многообразия бесконечно малым элементом $\epsilon \in \mathbb{G}$ справа. Тогда точка a перейдет в точку $a' = a + da$, где

$$da^{\mathrm{A}} = f^{\mathrm{A}}(a, \epsilon) - a^{\mathrm{A}} = \epsilon^{\mathrm{B}} L_{\mathrm{B}}{}^{\mathrm{A}}(a).$$

Поскольку для функции $F'(a') = F(a)$, то изменение формы функции будет равно:

$$\delta F(a) = F'(a) - F(a) = -\epsilon^{\mathrm{A}} L_{\mathrm{A}}{}^{\mathrm{B}} \partial_{\mathrm{B}} F(a).$$

Таким образом, левоинвариантные векторные поля, с точностью до знака, являются генераторами действия группы Ли справа. Аналогично показывается, что генераторами групповых преобразований слева являются правоинвариантные векторные поля.

Поскольку преобразования, вызванные действием группы слева и справа коммутируют, то отсюда следует, что лево- и правоинвариантные векторные поля коммутируют между собой,

$$[L_{\mathrm{A}}, R_{\mathrm{B}}] = 0. \tag{9.32}$$

Это равенство справедливо, конечно, глобально.

Из формулы для производной Ли от векторного поля (3.121) и равенства (9.32) следует, что производная Ли от правоинвариантного векторного поля вдоль левоинвариантного векторного поля равна нулю и наоборот.

В дальнейшем под алгеброй Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$ всегда понимается множество левоинвариантных векторных полей, которые генерируют действие группы справа.

9.4 Присоединенное представление

Рассмотрим автоморфизм группы Ли,

$$\mathsf{ad}\,(a): \quad \mathbb{G} \ni \ b \mapsto \mathsf{ad}\,(a)b := aba^{-1} \in \mathbb{G}, \tag{9.33}$$

который сопоставляется каждому элементу группы $a \in \mathbb{G}$. Каждый автоморфизм группы Ли порождает автоморфизм ее алгебры Ли, потому что левоинвариантные

векторные поля получены с помощью группового действия. Поэтому автоморфизм $\mathsf{ad}\, a$ порождает автоморфизм алгебры Ли $\mathfrak{g}$, который также обозначается $\mathsf{ad}\, a \in \mathbb{GL}(\mathrm{N}, \mathbb{R})$.

Определение. Представление $a \to \mathsf{ad}\, a$, где

$$\mathsf{ad}\,(a): \quad \mathfrak{g} \ni X = X_0^{\mathrm{A}} L_{\mathrm{A}} \mapsto X' = X_0^{\mathrm{A}} S_{\mathrm{A}}{}^{\mathrm{B}}(a) L_{\mathrm{B}} \in \mathfrak{g}, \tag{9.34}$$

индуцированное групповым действием (9.33), называется *присоединенным представлением* группы Ли $\mathbb{G}$ в ее алгебре Ли $\mathfrak{g}$. □

Присоединенное представление каждому элементу группы $a \in \mathbb{G}$ ставит в соответствие $\mathrm{N} \times \mathrm{N}$ матрицу $S_{\mathrm{A}}{}^{\mathrm{B}}(a)$. Для краткости мы будем писать

$$\mathsf{ad}\,(a) = S_{\mathrm{A}}{}^{\mathrm{B}}(a).$$

Ясно, что единичному элементу группы соответствует единичная матрица

$$\mathsf{ad}\,(e) = S_{\mathrm{A}}{}^{\mathrm{B}}(e) = \delta_{\mathrm{A}}^{\mathrm{B}},$$

а обратному элементу группы – обратная матрица

$$\mathsf{ad}\,(a^{-1}) = S_{\mathrm{A}}{}^{\mathrm{B}}(a^{-1}) = S^{-1}{}_{\mathrm{A}}{}^{\mathrm{B}}(a).$$

В наших обозначениях матрица для произведения двух элементов группы равна произведению матриц для каждого элемента,

$$S_{\mathrm{A}}{}^{\mathrm{B}}(ab) = S_{\mathrm{A}}{}^{\mathrm{C}}(b) S_{\mathrm{C}}{}^{\mathrm{B}}(a),$$

взятых в обратном порядке.

Замечание. Для абелевых групп присоединенное представление всегда тривиально. □

Ядром отображения (9.33) является центр группы. Каждому элементу центра группы соответствует единичная матрица присоединенного представления.

Для локальных неабелевых групп Ли можно построить явный вид матриц присоединенного представления в терминах функции композиции. С этой целью рассмотрим действие преобразования (9.33) в касательном пространстве в последовательности $b \mapsto (ab)a^{-1}$:

$$X^{\mathrm{A}}(b) \mapsto X^{\mathrm{C}}(b) U_{\mathrm{C}}{}^{\mathrm{B}}(a,b) V_{\mathrm{B}}{}^{\mathrm{A}}\Big(f(a,b), a^{-1}\Big).$$

Полагая $b = 0$ и учитывая (9.31), получаем

$$X_0^{\mathrm{A}} \mapsto X_0^{\mathrm{B}} L_{\mathrm{B}}{}^{\mathrm{C}}(a) R_{\mathrm{C}}^{-1\,\mathrm{A}}(a) := X_0^{\mathrm{B}} S_{\mathrm{B}}{}^{\mathrm{A}}(a),$$

где введена матрица присоединенного представления $S_{\mathrm{B}}{}^{\mathrm{A}}(a) \in \mathbb{GL}(n, \mathbb{R})$, сопоставляемая каждому элементу $a \in \mathbb{G}$. Таким образом получаем явные выражение для матрицы присоединенного представления через производные от функции композиции:

$$S_{\mathrm{B}}{}^{\mathrm{A}}(a) = L_{\mathrm{B}}{}^{\mathrm{C}}(a) R_{\mathrm{C}}^{-1\,\mathrm{A}}(a) = R_{\mathrm{B}}{}^{\mathrm{C}}(a^{-1}) L_{\mathrm{C}}^{-1\,\mathrm{A}}(a^{-1}), \tag{9.35}$$

где последнее выражение получено в результате рассмотрения действия $\mathsf{ad}\, a$ в последовательности $a(ba^{-1})$. Из этого представления следует, что

$$S_{\mathrm{B}}{}^{\mathrm{A}}(a^{-1}) = S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(a).$$

Разложения (9.16) и (9.30) приводят к следующему разложению матрицы присоединенного представления вблизи начала координат

$$S_{\text{B}}{}^{\text{A}}(a) = \delta^{\text{A}}_{\text{B}} + a^{\text{C}} f_{\text{CB}}{}^{\text{A}} + \dots \,. \tag{9.36}$$

Третье слагаемое ? Отсюда следует, что структурные константы, взятые с обратным знаком, можно рассматривать, как генераторы присоединенного представления: $L_{(\text{A})\text{B}}{}^{\text{C}} = -f_{\text{AB}}{}^{\text{C}}$. При этом первый индекс A нумерует генераторы и, для отличия, взят в скобки. Второй и третий индексы структурных констант B, C рассматриваются, как матричные. Тождества Якоби (9.10) переписываются в виде

$$[L_{(\text{A})}, L_{(\text{B})}]_{\text{C}}{}^{\text{D}} = f_{\text{AB}}{}^{\text{E}} L_{(\text{E})\text{C}}{}^{\text{D}}, \tag{9.37}$$

где квадратные скобки обозначают коммутатор матриц.

Рассмотрим два последовательных автоморфизма группы Ли $\mathbb{G}$, соответствующих элементам b и a:

$$c \;\mapsto\; bcb^{-1} \;\mapsto\; abcb^{-1}a^{-1} = (ab)c(ab)^{-1}.$$

Соответствующее преобразование в алгебре Ли $\mathfrak{g}$ задается матрицей присоединенного представления для произведения $f(a,b)$

$$S_{\text{A}}{}^{\text{B}}\Big(f(a,b)\Big) = S_{\text{A}}{}^{\text{C}}(b) S_{\text{C}}{}^{\text{B}}(a), \tag{9.38}$$

которая равна произведению матриц для каждого из элементов, но взятых с обратным знаком. Пусть $b \ll 1$. Разложив это соотношение в ряд с использованием формулы (9.36) и разложения для функции композиции

$$f^{\text{A}}(a,b) = a^{\text{A}} + b^{\text{B}} L_{\text{B}}{}^{\text{A}}(a) + \dots,$$

в первом порядке по b получим правило дифференцирования матрицы присоединенного представления

$$L_{\text{C}} S_{\text{A}}{}^{\text{B}} = L_{\text{C}}{}^{\text{D}} \partial_{\text{D}} S_{\text{A}}{}^{\text{B}} = f_{\text{CA}}{}^{\text{D}} S_{\text{D}}{}^{\text{B}}. \tag{9.39}$$

Из этого равенства с учетом тождества $L_{\text{C}}(S_{\text{A}}{}^{\text{D}} S^{-1}{}_{\text{D}}{}^{\text{B}}) = 0$ получаем правило действия левоинвариантных векторных полей на обратные матрицы присоединенного представления

$$L_{\text{C}} S^{-1}{}_{\text{A}}^{\text{B}} = -S^{-1}{}_{\text{A}}^{\text{D}} f_{\text{CD}}{}^{\text{B}}. \tag{9.40}$$

Аналогично, раскладывая соотношение (9.38) в ряд по $a \ll 1$, получим

$$R_{\text{C}} S_{\text{A}}{}^{\text{B}} = R_{\text{C}}{}^{\text{D}} \partial_{\text{D}} S_{\text{A}}{}^{\text{B}} = S_{\text{A}}{}^{\text{D}} f_{\text{CD}}{}^{\text{B}}. \tag{9.41}$$

С учетом равенства $R_{\text{C}}(S_{\text{A}}{}^{\text{D}} S^{-1}{}_{\text{D}}{}^{\text{B}}) = 0$ это дает

$$R_{\text{C}} S^{-1}{}_{\text{A}}^{\text{B}} = -f_{\text{CA}}{}^{\text{D}} S^{-1}{}_{\text{D}}^{\text{B}}.$$

Замечание. Полученные формулы дифференцирования матриц присоединенного представления (9.39) и (9.41) нековариантны, поскольку частная производная от тензора не является тензором. Однако при рассмотрении алгебр Ли левоинвариантные векторные поля имеют постоянные компоненты относительно левоинвариантного базиса и поэтому достаточно ограничиться аффинными преобразованиями координат с постоянными коэффициентами. В этом случае частные производные преобразуются по тензорным правилам. □

Предложение 9.4.1. *Структурные константы группы Ли* $\mathbb{G}$ *инвариантны относительно присоединенного действия группы,*

$$f_{\mathrm{AB}}{}^{\mathrm{C}} = S_{\mathrm{A}}{}^{\mathrm{D}} S_{\mathrm{B}}{}^{\mathrm{E}} f_{\mathrm{DE}}{}^{\mathrm{F}} S^{-1}{}_{\mathrm{F}}{}^{\mathrm{C}}. \tag{9.42}$$

Доказательство. Подействуем присоединенным представлением (9.33) на групповое многообразие $\mathbb{G}$. Тогда базис левоинвариантных векторных полей преобразуется в соответствии с присоединенным представлением

$$L_{\mathrm{A}} \mapsto S_{\mathrm{A}}{}^{\mathrm{B}} L_{\mathrm{B}}.$$

При этом коммутационные соотношения (9.15) сохранят свой вид, поскольку действие дифференциала отображения сохраняет коммутатор векторных полей. Следовательно, для структурных констант справедливо соотношение

$$S_{\mathrm{A}}{}^{\mathrm{D}} S_{\mathrm{B}}{}^{\mathrm{E}} f_{\mathrm{DE}}{}^{\mathrm{C}} = f_{\mathrm{AB}}{}^{\mathrm{D}} S_{\mathrm{D}}{}^{\mathrm{C}}.$$

Умножив это равенство справа на S^{-1} получим (9.42). □

9.5 Группы Ли как римановы пространства

На групповом многообразии $\mathbb{G}$ можно определить метрику, превратив тем самым группу Ли в (псевдо-)риманово пространство. Среди всех возможных метрик важную роль играют метрики, инвариантные относительно групповых преобразований. Эти метрики описываются следующим образом. Зададим в начале координат произвольную симметричную невырожденную матрицу θ_{AB} и разнесем ее по групповому многообразию с помощью действия группы слева или справа. В результате получим соответственно лево- и правоинвариантные метрики. Компоненты этих метрик в лево- и правоинвариантном базисе имеют тот же вид θ_{AB}, что и начале координат. В координатном базисе $g = dx^{\mathrm{A}} \otimes dx^{\mathrm{B}} g_{\mathrm{AB}}$ и компоненты метрик нетривиально зависят от точки $a \in \mathbb{G}$:

$$\begin{aligned} g^L_{\mathrm{AB}}(a) &= L^{-1}{}_{\mathrm{A}}{}^{\mathrm{C}}(a) L^{-1}{}_{\mathrm{B}}{}^{\mathrm{D}}(a) \theta_{\mathrm{CD}}, \\ g^R_{\mathrm{AB}}(a) &= R^{-1}{}_{\mathrm{A}}{}^{\mathrm{C}}(a) R^{-1}{}_{\mathrm{B}}{}^{\mathrm{D}}(a) \theta_{\mathrm{CD}}, \end{aligned} \tag{9.43}$$

где индексами L и R отмечены соответственно лево- и правоинвариантная метрики. В общем случае лево- и правоинвариантные метрики различны и их компоненты в координатном базисе нетривиально зависят от точки группового многообразия.

Замечание. Для абелевых групп лево- и правоинвариантные метрики равны, т.к. совпадают действия группы слева с справа. □

Скалярное произведение левоинвариантных векторных полей: $X = X_0^{\mathrm{A}} L_{\mathrm{A}}$ и $Y = Y_0^{\mathrm{A}} L_{\mathrm{A}}$, определенное левоинвариантной метрикой, равно

$$g^L(X, Y) = X_0^{\mathrm{A}} Y_0^{\mathrm{B}} g^L(L_{\mathrm{A}}, L_{\mathrm{B}}) = X_0^{\mathrm{A}} Y_0^{\mathrm{B}} \theta_{\mathrm{AB}}$$

и не зависит от точки группы Ли $\mathbb{G}$. Аналогично, скалярное произведение правоинвариантных векторных полей, определенное правоинвариантной метрикой равно константе.

Условия левой и правой инвариантности метрик в инвариантном виде имеют следующий вид. Пусть $X, Y \in \mathcal{X}(\mathbb{G})$ – два произвольных векторных поля на группе Ли $\mathbb{G}$. Тогда лево- и правоинвариантные метрики определены соотношениями

$$l_a^* g^L(X,Y) = g^L(l_{a*}X, l_{a*}Y),$$
$$r_a^* g^R(X,Y) = g^R(r_{a*}X, r_{a*}Y),$$

где l_a^* и r_a^* – возвраты левого и правого действия группы.

Особую роль играют двусторонне инвариантные метрики $g^L = g^R$ на неабелевых группах Ли. Для таких метрик скалярное произведение левоинвариантных векторных полей инвариантно относительно действия группы справа. Двусторонне инвариантные метрики существуют далеко не на всех группах Ли. Условие двусторонней инвариантности немедленно следует, например, из (9.43),

$$\theta_{\text{AB}} = S_{\text{A}}{}^{\text{C}}(a) S_{\text{B}}{}^{\text{D}}(a) \theta_{\text{CD}}, \qquad \forall a \in \mathbb{G},$$

где $S_{\text{A}}{}^{\text{B}}$ – матрица присоединенного представления. Для малых a это уравнение сводится к условию

$$f_{\text{AB}}{}^{\text{D}} \theta_{\text{DC}} + f_{\text{AC}}{}^{\text{D}} \theta_{\text{DB}} = 0. \tag{9.44}$$

То есть, если двусторонне инвариантная метрика на группе Ли существует, то структурные константы со всеми опущенными индексами должны быть антисимметричны по всем трем индексам.

Замечание. Двусторонне инвариантные метрики на группе Ли играют ту же роль, что и метрика Лоренца в пространстве Минковского $\mathbb{R}^{1,n-1}$ или евклидова метрика в евклидовом пространстве $\mathbb{R}^n$. □

Построим двусторонне инвариантную метрику на неабелевой группе Ли в том случае, когда это возможно.

Определение. Свертка структурных констант алгебры Ли

$$\eta_{\text{AB}} := -f_{\text{AC}}{}^{\text{D}} f_{\text{BD}}{}^{\text{C}} = \mathsf{tr}\,(f_{\text{A}} f_{\text{B}}) \tag{9.45}$$

называется *формой Киллинга–Картана.* □

Здесь мы рассматриваем структурные константы $f_{\text{AC}}{}^{\text{D}}$ в виде набора матриц. При этом второй и третий индексы рассматриваются, как матричные, а первый – нумерует матрицы. Форма (9.45) двусторонне инвариантна, что следует из инвариантности структурных констант (9.42).

Предложение 9.5.1. *Для произвольной группы Ли $\mathbb{G}$ структурные константы со всеми опущенными индексами,*

$$f_{\text{ABC}} = f_{\text{AB}}{}^{\text{D}} \eta_{\text{DC}}$$

антисимметричны по всем трем индексам,

$$f_{[abc]} = 0.$$

Доказательство. Прямое следствие тождеств Якоби (9.10), которое сводится к простой проверке. □

Таким образом, если форма Киллинга–Картана является невырожденной, то она задает двусторонне инвариантную метрику на группе Ли. Форма Киллинга–Картана является невырожденной не для всякой группы Ли. Напомним, что связная группа Ли называется полупростой, если она не имеет нетривиальных инвариантных связных абелевых подгрупп.

Теорема 9.5.1 (Э. Картан). *Форма Киллинга–Картана для связной группы Ли является невырожденной тогда и только тогда, когда группа Ли является полупростой.*

Доказательство. См., например, [**?**]. □

Полупростые группы не исчерпывают весь класс групп Ли, на которых можно задать двусторонне инвариантную метрику. Например, на абелевых группах любая невырожденная матрица в единице, разнесенная по групповому многообразию, задает двусторонне инвариантную метрику.

Пример 9.5.1. Двумерная абелева группа Ли изоморфна либо евклидовой плоскости $\mathbb{R}^2$, либо цилиндру, либо тору. Пусть x, y – декартовы координаты на плоскости. Тогда лево- и правоинвариантные векторные поля совпадают и равны

$$L_1 = R_1 = \partial_x, \qquad L_2 = R_2 = \partial_y.$$

Двусторонне инвариантная метрика на $\mathbb{R}^2$ получается путем разнесения произвольной симметричной невырожденной 2×2 матрицы θ_{AB}, $\text{A}, \text{B} = 1, 2$, заданной в начале координат, с помощью действия группы. В результате получаем двусторонне инвариантную метрику, которая в декартовых координатах имеет вид θ_{AB} для всех точек $\mathbb{R}^2$. Евклидова плоскость с такой метрикой является однородным, но не изотропным пространством. При $\theta_{\text{AB}} = \delta_{\text{AB}}$ евклидова плоскость становится также изотропной, т.е. инвариантной также относительно действия группы вращений $\mathbb{O}(2)$. Цилиндр и тор в этом случае неинвариантны относительно вращений. □

Форма Киллинга–Картана (9.45) задает метрику на групповом многообразии полупростой группы Ли $\mathbb{G}$ либо в лево-, либо в правоинвариантном базисе. Для определенности будем считать, что метрика задана в левоинвариантном базисе. Тогда в координатном базисе она будет нетривиально зависеть от точки группового многообразия:

$$g_{\text{AB}}(a) = L^{-1}{}_{\text{A}}{}^{\text{C}}(a) L^{-1}{}_{\text{B}}{}^{\text{D}}(a) \eta_{\text{CD}}.$$

Двусторонняя инвариантность формы Киллинга–Картана приводит к существованию, по крайней мере, 2N векторов Киллинга L_{A} и R_{A}. Уравнения Киллинга в неголономном базисе (7.100) выполнены ввиду антисимметрии символов Кристоффеля (9.46) по первым двум индексам.

Вычислим геометрические характеристики группового многообразия неабелевой полупростой группы Ли, рассматривая его, как (псевдо-)риманово пространство с нулевым кручением и неметричностью. Вычисления можно проводить в любом базисе, но в данном случае удобнее рассматривать все характеристики многообразия в неголономном базисе, определяемом левоинвариантными векторными полями L_{A}, поскольку компоненты метрики в этом базисе постоянны. Соответствующие формулы для вычислений приведены в разделе 7.9.

Компоненты неголономности базиса определяются коммутатором левоинвариантных векторных полей (9.15) и совпадают со структурными константами:

$$c_{\mathrm{AB}}{}^{\mathrm{C}} = f_{\mathrm{AB}}{}^{\mathrm{C}}.$$

Символы Кристоффеля в неголономном базисе, вычисленные по формуле (7.99), равны

$$\widetilde{\omega}_{\mathrm{ABC}} = \frac{1}{2} f_{\mathrm{ABC}}, \tag{9.46}$$

где знак тильды означает, что кручение и неметричность равны нулю. В рассматриваемом случае символы Кристоффеля антисимметричны по первым двум индексам ввиду неголономности используемого базиса. Соответствующий тензор кривизны также задается структурными константами

$$\widetilde{R}_{\mathrm{ABC}}{}^{\mathrm{D}} = \frac{1}{4} f_{\mathrm{AB}}{}^{\mathrm{E}} f_{\mathrm{CE}}{}^{\mathrm{D}}.$$

Отсюда вытекают выражения для тензора Риччи и скалярной кривизны:

$$\widetilde{R}_{\mathrm{AB}} = \frac{1}{4}\eta_{\mathrm{AB}}, \qquad \widetilde{R} = \frac{\mathrm{N}}{4},$$

где N – размерность полупростой группы Ли.

Покажем, что полупростая группа Ли с формой Киллинга–Картана в качестве метрики, представляет собой пространство постоянной кривизны. Пусть структурные константы f_{ABC} задают полностью антисимметричный тензор третьего ранга в левоинвариантном базисе. Тогда ковариантная производная этого тензора со связностью (9.46),

$$\widetilde{\nabla}_{\mathrm{A}} f_{\mathrm{BCD}} = L_{\mathrm{A}} f_{\mathrm{BCD}} - \frac{1}{2} f_{\mathrm{AB}}{}^{\mathrm{E}} f_{\mathrm{ECD}} - \frac{1}{2} f_{\mathrm{AC}}{}^{\mathrm{E}} f_{\mathrm{BED}} - \frac{1}{2} f_{\mathrm{AD}}{}^{\mathrm{E}} f_{\mathrm{BCE}} = 0,$$

равна нулю в силу тождеств Бианки. Поскольку связность является метрической, а подъем и опускание индексов можно переставлять с оператором ковариантного дифференцирования, то ковариантная производная от тензора кривизны равна нулю,

$$\widetilde{\nabla}_{\mathrm{A}} \widetilde{R}_{\mathrm{BCD}}{}^{\mathrm{E}} = 0.$$

Это и означает, что группа Ли как (псевдо-)риманово пространство является пространством постоянной кривизны.

Тензор кривизны со всеми опущенными индексами в неголономном базисе имеет вид

$$\widetilde{R}_{\mathrm{ABCD}} = -\frac{1}{4} f_{\mathrm{AB}}{}^{\mathrm{E}} f_{\mathrm{CDE}}, \tag{9.47}$$

в котором явно прослеживаются все свойства симметрии относительно перестановки индексов.

Замечание. Тензор кривизны (9.47) нельзя записать в виде (**??**) через метрику. Действительно, равенство

$$f_{\mathrm{AB}}{}^{\mathrm{E}} f_{\mathrm{CDE}} = C(\eta_{\mathrm{AC}}\eta_{\mathrm{BD}} - \eta_{\mathrm{AD}}\eta_{\mathrm{BC}}), \qquad C = \mathsf{const},$$

не может быть выполнено ни при каких значениях константы C. Для доказательства достаточно свернуть это равенство сначала с η^{BD}, а затем с $f^{\mathrm{CD}}{}_{\mathrm{F}}$, что приведет к противоречию. □

Теорема 9.5.2. *Связная полупростая группа Ли компактна тогда и только тогда, когда ее форма Киллинга–Картана положительно определена.*

Доказательство. См., например, [**?**]. □

9.6 Группы Ли как пространства Римана–Картана

В предыдущем разделе связность на групповом многообразии определялась по метрике при нулевом кручении и неметричности. Ниже мы рассмотрим другой способ определения связности, который во многих отношениях является более естественным для групп Ли. А именно, отождествим групповое действие справа с параллельным переносом. Это означает, что результат параллельного переноса вектора из точки a в точку c совпадает с дифференциалом группового преобразования $c = f(a, b)$ под действием элемента $b = a^{-1}c$.

Формализуем данное определение связности. В рассматриваемом случае правоинвариантные векторные поля можно рассматривать, как результат параллельного переноса векторов из начала координат по всему групповому многообразию. В левоинвариантном базисе правоинвариантное векторное поле имеет вид

$$X = X_0^{\text{A}} R_{\text{A}}{}^{\text{B}}(a)\partial_{\text{B}} = X_0^{\text{A}} S_{\text{A}}{}^{\text{B}}(a^{-1}) L_{\text{B}},$$

где мы воспользовались представлением (9.35) для матрицы присоединенного представления. Несложные вычисления показывают, что разница векторов в соседних точках равна

$$\delta X^{\text{A}}(a) = X^{\text{A}}(a + da) - X^{\text{A}}(a) = -X^{\text{C}} da^{\text{B}} f_{\text{BC}}{}^{\text{A}}.$$

С другой стороны, при параллельном переносе компоненты вектора получают приращение, определяемое аффинной связностью:

$$\delta X^{\text{A}} = -X^{\text{C}} da^{\text{B}} \omega_{\text{BC}}{}^{\text{A}}.$$

Отсюда следует явное выражение для компонент аффинной связности в левоинвариантном базисе

$$\omega_{\text{AB}}{}^{\text{C}} = f_{\text{AB}}{}^{\text{C}}. \tag{9.48}$$

В правоинвариантном базисе $\delta X^{\text{A}} = 0$, и связность тождественно равна нулю.

Замечание. В данном определении связности наличие или отсутствие метрики на групповом многообразии не играет никакой роли. □

Связность (9.48) отличается от связности Леви-Чивита (9.46) множителем 1/2. Соответствующие тензоры кривизны и кручения, вычисленные по формулам для неголономного базиса (7.96) и (7.97), равны

$$R_{\text{ABC}}{}^{\text{D}} = 0, \tag{9.49}$$

$$T_{\text{AB}}{}^{\text{C}} = f_{\text{AB}}{}^{\text{C}}. \tag{9.50}$$

Поскольку тензор кривизны тождественно равен нулю, то групповое пространство является пространством абсолютного параллелизма. Так и должно быть, потому что результат параллельного переноса вектора из точки a в точку $c = ab$ однозначно определяется начальной и конечной точкой и не зависит от пути, вдоль которого осуществляется параллельный перенос.

Компоненты связности (9.48) в левоинвариантном базисе определены независимо от метрики и для произвольных, не обязательно полупростых, групп. На полупростых группах Ли эта связность является метрической:

$$\nabla_{\text{A}} \eta_{\text{BC}} = -\omega_{\text{AB}}{}^{\text{D}} \eta_{\text{DC}} - \omega_{\text{AC}}{}^{\text{D}} \eta_{\text{BD}},$$

где η_{BC} – двусторонне инвариантная форма Киллинга–Картана. Это значит, что групповое многообразие полупростых групп Ли со связностью, определенной компонентами (9.48), представляет собой пространство Римана–Картана.

Группа Ли, рассматриваемая как пространство Римана–Картана является пространством постоянной (нулевой) кривизны и постоянного кручения. Действительно, ковариантная производная от тензора кривизны, очевидно, равна нулю. Ковариантная производная от тензора кручения со связностью (9.48),

$$\nabla_{\text{A}} T_{\text{BC}}{}^{\text{D}} = L_{\text{A}} T_{\text{BC}}{}^{\text{D}} - f_{\text{AB}}{}^{\text{E}} T_{\text{EC}}{}^{\text{D}} - f_{\text{AC}}{}^{\text{E}} T_{\text{BE}}{}^{\text{D}} + f_{\text{AE}}{}^{\text{D}} T_{\text{BC}}{}^{\text{E}} = 0,$$

также равна нулю в силу тождеств Якоби.

Данное определение связности не зависит от задания метрики на групповом многообразии, т.к. связность была определена исключительно групповой операцией. Попостроению, интегральные кривые правоинвариантных векторных полей на группе Ли являются геодезическими линиями для этой связности. Действительно, правоинвариантное векторное поле является касательным к интегральной кривой и в то же время является результатом параллельного переноса со связностью (9.48).

Замечание. Для задания на группе Ли инвариантной аффинной связности общего вида можно задать в некоторой точке метрику, кручение и тензор неметричности, а затем все эти объекты разнести по групповому многообразию. □

9.7 Двумерная неабелева группа Ли

Проиллюстрируем общие свойства групп Ли на примере простейшей неабелевой группы Ли $\mathbb{G}$. Размерность этой группы минимальна и равна двум. Этот пример интересен, поскольку нетривиален и в то же время достаточно прост для явного построения всех геометрических конструкций.

Простейшая двумерная алгебра Ли $\mathfrak{g}$ с образующими $\{L_{\text{A}}\} = (L_x, L_y)$, $\text{A} = x, y$, задается коммутационными соотношениями:

$$\begin{aligned} &[L_x, L_x] = [L_y, L_y] = 0, \\ &[L_x, L_y] = \alpha L_y, \qquad \alpha \in \mathbb{R}. \end{aligned} \tag{9.51}$$

При $\alpha = 0$ эта группа является абелевой. При $\alpha \neq 0$ растяжкой координат, которая соответствует преобразованию $L_x \to \alpha L_x$, можно обратить структурную константу в единицу, $\alpha = 1$. Группа, соответствующая этой алгебре является подгруппой группы Лоренца $\mathbb{SO}(1,2)$.

Предложение 9.7.1. *Алгебра Ли (9.51) является единственной двумерной неабелевой алгеброй Ли с точностью до изоморфизма.*

Доказательство. В общем случае нетривиальное коммутационное соотношение для неабелевой алгебры Ли определяется двумя отличными от нуля постоянными α и β:

$$[L_x, L_y] = \beta L_x + \alpha L_y.$$

Сдвигом

$$L_y \mapsto L_x + \frac{\alpha}{\beta} L_y$$

это коммутационное соотношение приводится к виду (9.51). □

Зададим алгебру Ли (9.51) двумя левоинвариантными векторными полями (генераторами действия группы справа) на евклидовой плоскости $\mathbb{R}^2$ в декартовых координатах x, y:

$$L_x = \partial_x, \qquad L_y = \mathsf{e}^{\alpha x}\partial_y. \tag{9.52}$$

Нетрудно проверить, что эти векторные поля действительно удовлетворяют алгебре (9.51). При $\alpha = 0$ группа Ли является абелевой и совпадает с группой трансляций плоскости. Поэтому ограничимся неабелевым случаем $\alpha > 0$. Структурные константы группы Ли равны

$$f_{xy}{}^{x} = -f_{yx}{}^{x} = 0,$$
$$f_{xy}{}^{y} = -f_{yx}{}^{y} = \alpha.$$

Простые вычисления дают форму Киллинга–Картана (9.45)

$$\eta_{\text{AB}} = \begin{pmatrix} \alpha^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Эта форма вырождена, поскольку группа Ли $\mathbb{G}$ содержит инвариантную абелеву подгруппу с образующей L_y, и не является полупростой.

Постоим функцию композиции $f^{\text{A}}(a, b)$. Обозначим координаты элементов группы $a, b \in \mathbb{G}$ через $a = (x_a, y_a)$ и $b = (x_b, y_b)$. Тогда система дифференциальных уравнений на функцию композиции (9.17) примет вид

$$\frac{\partial f^x(a,b)}{\partial x_b} = 1, \qquad \frac{\partial f^y(a,b)}{\partial x_b} = 0,$$
$$\frac{\partial f^x(a,b)}{\partial y_b} = 0, \qquad \frac{\partial f^y(a,b)}{\partial y_b} = \mathsf{e}^{\alpha(f^x - x_b)}.$$

Два уравнения на f^x с учетом начального условия $f^x(a, 0) = x_a$ имеют единственное решение $f^x(a, b) = x_a + x_b$. Аналогично, оставшиеся уравнения определяют вторую компоненту функции композиции $f^y(a, b) = y_a + \mathsf{e}^{\alpha x_a} y_b$. Таким образом, дифференциальные уравнения (9.17) с начальными условиями $a^{\text{A}} = f^{\text{A}}(a, 0)$ однозначно определили функцию композиции по левоинвариантным векторным полям L_{A}:

$$\{f^{\text{A}}(a, b)\} = (x_a + x_b, y_a + \mathsf{e}^{\alpha x_a} y_b). \tag{9.53}$$

Поскольку функция композиции определена для всех $a, b \in \mathbb{R}^2$ и отображает $\mathbb{R}^2 \times \mathbb{R}^2$ на всю плоскость $\mathbb{R}^2$, то группа Ли $\mathbb{G}$, как многообразие, диффеоморфна евклидовой плоскости $\mathbb{R}^2$ и, следовательно, некомпактна.

Несложные вычисления показывают, что построенная функция композиции (9.53) действительно удовлетворяет всем групповым аксиомам (9.6) с обратным элементом:

$$a^{-1} = (-x_a, -\mathsf{e}^{-\alpha x_a} y_a).$$

Таким образом, мы построили неабелеву групповую структуру на евклидовой плоскости $\mathbb{R}^2$. Вспомним также, что плоскость $\mathbb{R}^2$ можно рассматривать, как абелеву группу Ли двумерных трансляций. Тем самым мы показали, что на одном и том же многообразии возможно задание различных групповых структур.

Знание функции композиции позволяет вычислить все объекты на групповом многообразии, которые задаются групповым действием. В частности, правоинвариантные векторные поля (9.27) имеют вид:

$$R_x = \partial_x + \alpha y \partial_y, \qquad R_y = \partial_y, \tag{9.54}$$

и удовлетворяют коммутационным соотношениям:

$$[R_x, R_x] = [R_y, R_y] = 0$$
$$[R_x, R_y] = -\alpha R_y,$$

которые отличаются знаком от коммутационных соотношений алгебры Ли левоинвариантных векторных полей (9.51). Нетрудно также проверить, что лево- и правоинвариантные векторные поля коммутируют, $[L_{\text{A}}, R_{\text{B}}] = 0$.

Построим интегральные кривые левоинвариантных векторных полей. Произвольное левоинвариантное векторное поле задается двумя постоянными $X_0, Y_0 \in \mathbb{R}$

$$L = X_0\partial_x + Y_0\,\mathsf{e}^{\alpha x}\partial_y.$$

Соответствующие интегральные кривые $\Big(x(t), y(t)\Big)$ определяются системой уравнений

$$\dot{x} = X_0, \qquad \dot{y} = Y_0\,\mathsf{e}^{\alpha x}.$$

При $X_0 \neq 0$, эти уравнения имеют общее решение, задающее интегральную кривую в параметрическом виде

$$x = X_0 t + x_0,$$
$$y = \frac{Y_0}{\alpha X_0}\,\mathsf{e}^{\alpha x_0}(\,\mathsf{e}^{\alpha X_0 t} - 1) + y_0,$$

где постоянные интегрирования x_0, y_0 являются координатами точки, через которую проходит интегральная кривая при $t = 0$. Отсюда следует соотношение, определяющее форму интегральной кривой

$$y = \frac{Y_0}{\alpha X_0}(\,\mathsf{e}^{\alpha x} - \,\mathsf{e}^{\alpha x_0}) + y_0.$$

При $X_0 = 0$ интегральные кривые параллельны оси y. Вид интегральных кривых, проходящих через начало координат под разными углами, которые определяются постоянными X_0, Y_0, показан на рис. **??**

Интегральные кривые определены для всех значений параметра $t \in \mathbb{R}$ и поэтому левоинвариантные векторные поля полны.

Зададим теперь геометрию, т.е. метрику и аффинную связность, на группе Ли $\mathbb{G}$. Конечно, интерес представляют те геометрические объекты, которые согласованы с групповой операцией. Начнем с метрики. Поскольку группа Ли $\mathbb{G}$ неабелева и не является полупростой, то на ней не существует двусторонне инвариантной метрики. Лучшее, что можно сделать, это построить право- или левоинвариантную метрику. С этой целью зададим матрицу θ_{AB} в начале координат и разнесем ее с помощью действия группы преобразований. Правоинвариантная метрика в координатном базисе (9.43) имеет вид

$$g_{\text{AB}} = g^R_{\text{AB}}(a) = R_{\text{A}}^{-1\text{C}}(a) R_{\text{B}}^{-1\text{D}}(a)\theta_{\text{CD}}, \tag{9.55}$$

где матрица θ_{AB} симметрична и невырождена, а в остальном совершенно произвольна. В частности, она может иметь лоренцеву сигнатуру, т.е. на группе Ли $\mathbb{G}$ можно построить как риманову, так и псевдориманову геометрию.

Рассмотрим случай, когда в начале координат задана единичная матрица, $\theta_{\text{AB}} = \delta_{\text{AB}}$. Тогда формула (9.55), в которой матрицы $R_{\text{A}}{}^{\text{B}}$ определены равенствами (9.54), задает следующую метрику в декартовых координатах на плоскости $\mathbb{R}^2$

$$g_{\text{AB}} = \begin{pmatrix} 1 + \alpha^2 y^2 & -\alpha y \\ -\alpha y & 1 \end{pmatrix}. \tag{9.56}$$

Эта метрика имеет единичный определитель, $\det g_{\alpha\beta} = 1$, и ее обратная имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} 1 & \alpha y \\ \alpha y & 1 + \alpha^2 y^2 \end{pmatrix}.$$

Прямые вычисления дают следующие выражения для символов Кристоффеля:

$$\begin{aligned} \widetilde{\Gamma}_{xx}{}^{x} &= -\alpha^3 y^2, & \widetilde{\Gamma}_{xx}{}^{y} &= -\alpha^2 y(1 + \alpha^2 y^2), \\ \widetilde{\Gamma}_{xy}{}^{x} &= \widetilde{\Gamma}_{yx}{}^{x} = \alpha^2 y, & \widetilde{\Gamma}_{xy}{}^{y} &= \widetilde{\Gamma}_{yx}{}^{y} = \alpha^3 y^2, \\ \widetilde{\Gamma}_{yy}{}^{x} &= -\alpha, & \widetilde{\Gamma}_{yy}{}^{y} &= -\alpha^2 y. \end{aligned}$$

У соответствующего тензора кривизны со всеми опущенными индексами только одна компонента отлична от нуля:

$$\widetilde{R}_{xyxy} = \alpha^2,$$

что можно переписать в ковариантном виде

$$\widetilde{R}_{\alpha\beta\gamma\delta} = \alpha^2(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}).$$

Тем самым групповое многообразие $\mathbb{G} \approx \mathbb{R}^2$ с метрикой (9.56) становится пространством постоянной кривизны (см. раздел **??**). Соответствующий тензор Риччи и скалярная кривизна нетривиальны:

$$\begin{aligned} \widetilde{R}_{xx} &= \alpha^2(1 + \alpha^2 y^2), \\ \widetilde{R}_{xy} &= \widetilde{R}_{yx} = -\alpha^3 y, \\ \widetilde{R}_{yy} &= \alpha^2, \\ \widetilde{R} &= -2K = 2\alpha^2, \end{aligned}$$

где $K = -\alpha^2$ – гауссова кривизна двумерной группы Ли $\mathbb{G}$. Как многообразие, это – поверхность постоянной отрицательной кривизны (одна компонента двуполостного гиперболоида), которая рассмотрена в разделе 25.2.

9.8 Гомоморфизмы групп Ли

Определение. Отображение $f : \mathbb{G} \to \mathbb{H}$ называется *гомоморфизмом групп Ли*, если f гладко и в то же время является гомоморфизмом групп. Если, кроме того, f является диффеоморфизмом, то отображение f называется *изоморфизмом групп Ли*. Изоморфизм группы Ли на себя называется ее *автоморфизмом*. □

Можно доказать, что из непрерывности гомоморфизма $f : \mathbb{G} \to \mathbb{H}$ следует его гладкость [**?**].

Определение. Гомоморфизм $\rho : \mathbb{G} \to \mathbb{H}$ называется *представлением* группы Ли $\mathbb{G}$, если $\mathbb{H} = \mathsf{aut}(\mathbb{V})$ для некоторого векторного пространства $\mathbb{V}$, или $\mathbb{H} = \mathbb{GL}(n, \mathbb{C})$, или $\mathbb{H} = \mathbb{GL}(n, \mathbb{R})$. Если гомоморфизм $\rho : \mathbb{G} \to f(\mathbb{G}) \subset \mathbb{H}$ является изоморфизмом, то представление называется *точным*. Представление ρ группы Ли $\mathbb{G}$ в векторном пространстве $\mathbb{V}$ мы обозначаем парой $(\rho, \mathbb{V})$. □

Замечание. В определении представления во всех трех случаях каждому элементу группы $\mathbb{G}$ ставится в соответствие матрица. Если элементами матриц являются вещественные или комплексные числа, то говорят, соответственно, о *вещественном* или *комплексном* представлении. При этом все матрицы имеют отличный от нуля определитель, т.к. у каждого элемента группы должен быть обратный элемент. Следовательно, каждой матрице представления соответствует некоторый автоморфизм векторного пространства. Для точного представления группы образ $f(\mathbb{G}) \subset \mathbb{H}$ совсем не обязательно совпадает со всем $\mathbb{H}$. □

Определение. Пусть $\mathfrak{g}$ и $\mathfrak{h}$ – алгебры Ли. Тогда отображение $\varphi : \mathfrak{g} \to \mathfrak{h}$ называется *гомоморфизмом алгебр Ли*, если оно линейно и сохраняет скобку Ли, т.е. $\varphi[X, Y] = [\varphi(X), \varphi(Y)] \in \mathfrak{h}$ для всех $X, Y \in \mathfrak{g}$. Если, кроме того, φ является взаимно однозначным отображением "на" (биекцией), то φ называется *изоморфизмом алгебр Ли*. Изоморфизм алгебры Ли на себя называется ее *автоморфизмом*. □

В определении гомоморфизма алгебр Ли мы не требуем гладкости отображения φ, т.к. алгебры Ли рассматриваются, как векторные пространства, а не многообразия. С другой стороны, на любом векторном пространстве можно ввести гладкую структуру многообразия, тогда из линейности отображения φ будет следовать его гладкость.

Определение. Гомоморфизм $\phi : \mathfrak{g} \to \mathfrak{h}$ называется *представлением алгебры Ли* $\mathfrak{g}$, если $\mathfrak{h} = \mathsf{end}(\mathbb{V})$ для некоторого векторного пространства $\mathbb{V}$, или $\mathfrak{h} = \mathfrak{gl}(n, \mathbb{C})$, или $\mathfrak{h} = \mathfrak{gl}(n, \mathbb{R})$. Если гомоморфизм $\phi : \mathfrak{g} \to \varphi(\mathfrak{g}) \subset \mathfrak{h}$ является изоморфизмом, то представление называется *точным*. □

Замечание. Как и в случае представления группы каждый элемент представления алгебры является матрицей. Однако эти матрицы могут быть вырождены и, соответственно, задавать только эндоморфизм векторного пространства. □

Пусть $f : \mathbb{G} \to \mathbb{H}$ – гомоморфизм групп Ли. Так как под действием отображения f единица группы $\mathbb{G}$ переходит в единицу группы $\mathbb{H}$, то дифференциал отображения отображает касательные пространства к единицам групп:

$$f_* : \quad \mathbb{T}_e(\mathbb{G}) \ \to \ \mathbb{T}_e(\mathbb{H}).$$

Замечание. Единицы разных групп мы обозначаем одним и тем же символом e, т.к. это не приводит к путанице. □

Поскольку касательные пространства к единицам изоморфны алгебрам Ли, то дифференциал отображения индуцирует отображение алгебр, которое мы также будем обозначать f_*. Таким образом,

$$f_* : \quad \mathfrak{g} \ \to \ \mathfrak{h}, \tag{9.57}$$

где образом $f_*(X)$ для $X \in \mathfrak{g}$ является единственное левоинвариантное векторное поле на $\mathbb{H}$, такое, что в единице группы $e \in \mathbb{H}$ оно определено соотношением

$$f_*(X)(e) = f_*(X(e)).$$

Предложение 9.8.1. *Пусть $\mathbb{G}$ и $\mathbb{H}$ – группы Ли с алгебрами Ли $\mathfrak{g}$ и $\mathfrak{h}$ соответственно, и пусть $f : \mathbb{G} \to \mathbb{H}$ – гомоморфизм групп Ли. Тогда отображение (9.57) является гомоморфизмом алгебр Ли.*

Доказательство. Дифференциал отображения линеен и сохраняет скобку Ли (3.73). □

Пусть $f : \mathbb{G} \to \mathbb{H}$ – гомоморфизм групп Ли. Тогда возврат отображения f^* отображает левоинвариантные 1-формы на $\mathbb{H}$ в левоинвариантные 1-формы на $\mathbb{G}$, как отображение сопряженного пространства к $\mathfrak{h}$ в сопряженное пространство к $\mathfrak{g}$.

Пример 9.8.1 (Обозначения). Запишем вектор из векторного пространства $\mathbb{V}$ в виде $X = X^i \boldsymbol{e}_i$, где $\boldsymbol{e}_i$, $i = 1, \ldots, \mathsf{dim}\,\mathbb{V}$, – некоторый базис. Если задано представление $(\rho, \mathbb{V})$ группы Ли $\mathbb{G}$, то мы пишем

$$\mathbb{G} \ni a : \quad \mathbb{V} \ni X = X^i \boldsymbol{e}_i \mapsto X' = X^i S_i{}^j \boldsymbol{e}_j \in \mathbb{V}.$$

То есть каждому элементу группы $a \in \mathbb{G}$ соответствует матрица $S_i{}^j(a)$. По определению, все элементы матрицы гладко зависят от a. Мы пишем

$$\rho : \quad \mathbb{G} \ni a \mapsto S_i{}^j(a) \in \mathsf{aut}\,\mathbb{V}.$$

Ясно, что единичному элементу группы соответствует единичная матрица,

$$S_i{}^j(e) = \delta_i^j,$$

а обратному элементу группы – обратная матрица,

$$S_i{}^j(a^{-1}) = S^{-1}{}_j{}^j(a).$$

Произведению двух элементов группы ab соответствует матрица

$$S_i{}^j(ab) = S_i{}^k(b) S_k{}^j(a),$$

которая равна произведению матриц для каждого элемента, взятых в обратном порядке. Порядок матриц является следствием принятого нами соглашения в теории групп преобразований, где элемент группы, по-определению, действует на точку многообразия справа (см. раздел 10).

Дифференциал отображения ρ_* отображает алгебры Ли:

$$\rho_* : \quad \mathfrak{g} \ni L_{\mathrm{A}} \mapsto L_{\mathrm{A}i}{}^j \in \mathsf{end}\,\mathbb{V}.$$

с матрицами представления $L_{\mathrm{A}i}{}^j$ генераторов L_{A}, которые удовлетворяют тем же коммутационным соотношениям, что и левоинвариантный базис векторных полей на группе Ли

$$[L_{\mathrm{A}}, L_{\mathrm{B}}]_i{}^j = f_{\mathrm{AB}}{}^{\mathrm{C}} L_{\mathrm{C}i}{}^j, \tag{9.58}$$

где $f_{\mathrm{AB}}{}^{\mathrm{C}}$ – структурные константы группы Ли $\mathbb{G}$, определенные в (9.15).

Можно доказать, что матрицы представления группы Ли $\mathbb{G}$ вблизи единицы группы в линейном приближении имеют вид

$$S_i{}^j(da) = \delta_i{}^j - da^{\mathrm{A}} L_{\mathrm{A}i}{}^j, \tag{9.59}$$

где $da \in \mathbb{G}$ – элемент группы, близкий к единице. Найдем правила дифференцирования матриц представления. С одной стороны, используя (9.59), имеем равенство

$$S_i{}^j\Big(f(a, da)\Big) = S_i{}^k(da) S_k{}^j(a) = \Big(\delta_i^k - da^{\mathrm{A}} L_{\mathrm{A}i}{}^k\Big) S_k{}^j(a). \tag{9.60}$$

С другой стороны, матрицы представления можно разложить в ряд Тейлора:

$$S_i{}^j\Big(f(a,da)\Big) = S_i{}^j(a) + da^{\text{A}} L_{\text{A}} S_i{}^j, \tag{9.61}$$

где левоинвариантные векторные поля L_{A} действуют на матрицы представления, как дифференцирования, и мы воспользовались равенством

$$f^{\text{A}}(a,da) \simeq a^{\text{A}} + da^{\text{B}} L_{\text{B}}{}^{\text{A}}.$$

Сравнивая формулы (9.60) и (9.61), получаем правило дифференцирования матриц представления

$$L_{\text{A}} S_i{}^j = -L_{\text{A}i}{}^k S_k{}^j. \tag{9.62}$$

Дифференцируя равенство $S^{-1}S = \mathbb{1}$, получаем правило дифференцирования обратных матриц представления

$$L_{\text{A}} S^{-1}{}_i^j = S^{-1}{}_i^k L_{\text{A}k}{}^j. \tag{9.63}$$

Каждому автоморфизму $\rho(a)$ векторного пространства $\mathbb{V}$ соответствует гомоморфизм алгебр $\rho_*(a)$. Поскольку гомоморфизм алгебр сохраняет скобку Ли, то матрицы представления генераторов группы инвариантны:

$$L_{\text{A}i}{}^j = S^{-1}{}_{\text{A}}^{\text{B}} S^{-1}{}_i^k L_{\text{B}k}{}^l S_l{}^j, \tag{9.64}$$

где $S_{\text{A}}{}^{\text{B}}$ – матрица присоединенного представления. □

Продолжим общее рассмотрение.

Определение. Пара $(f,\mathbb{H})$ называется *подгруппой Ли* группы Ли $\mathbb{G}$, если выполнены следующие условия:
1) $\mathbb{H}$ – группа Ли;
2) $(f,\mathbb{H})$ – подмногообразие в $\mathbb{G}$;
3) $f:\ \mathbb{H}\to\mathbb{G}$ – групповой гомоморфизм.
Пара $(f,\mathbb{H})$ называется *замкнутой подгруппой Ли* в $\mathbb{G}$, если к тому же $f(\mathbb{H})$ – замкнутое подмногообразие в $\mathbb{G}$. □

Определение замкнутого подмногообразия было дано в разделе 3.10. Если группа Ли $\mathbb{G}$ связна и $\mathbb{H}$ не изоморфна $\mathbb{G}$, то размерность замкнутой подгруппы Ли всегда меньше размерности самой группы.

Определение. Подпространство $\mathfrak{h}\subset\mathfrak{g}$ называется *подалгеброй Ли* алгебры Ли $\mathfrak{g}$, если для любых $X,Y\in\mathfrak{h}$ коммутатор $[X,Y]$ тоже лежит в $\mathfrak{h}$. Мы будем также употреблять краткую запись этого условия: $[\mathfrak{h},\mathfrak{h}]\subset\mathfrak{h}$. □

Ясно, что любая подалгебра сама является алгеброй.

Пусть $(f,\mathbb{H})$ – подгруппа Ли в $\mathbb{G}$, и $\mathfrak{h}$ и $\mathfrak{g}$ – соответствующие алгебры Ли. Тогда дифференциал отображения f_* задает изоморфизм $\mathfrak{h}$ с подалгеброй $f_*(\mathfrak{h})$ в алгебре $\mathfrak{g}$.

Гомоморфизмы групп можно описать следующим образом. Пусть $f:\ \mathbb{G}\to\mathbb{H}$ – сюръективное отображение. Если это отображение является взаимно однозначным, то мы имеем изоморфизм групп, $\mathbb{G}\simeq\mathbb{H}$. Если отображение f не является взаимно однозначным, то ядро этого отображения $\ker f$ представляет собой нормальную подгруппу в $\mathbb{G}$. Тогда фактор группа $\mathbb{G}/\ker f$ изоморфна $\mathbb{H}$. Если отображение f не является сюръективным, то группа $\mathbb{G}$ сюръективно отображается на какую то подгруппу $\mathbb{H}_1\subset\mathbb{H}$. Если отображение $\mathbb{G}$ на $\mathbb{H}_1$ взаимно однозначно, то отображение f является мономорфизмом. В противном случае, $\mathbb{G}/\ker f\simeq\mathbb{H}_1$.

Теорема 9.8.1. *Пусть $\mathbb{G}$ – связная группа Ли, и $\mathbb{U}$ – некоторая окрестность единицы. Тогда*

$$\mathbb{G} = \bigcup_{k=1}^{\infty} \mathbb{U}^k, \tag{9.65}$$

где множество элементов $\mathbb{U}^k$ состоит из всех возможных произведений k элементов из $\mathbb{U}$.

Доказательство. См., например, [?]. □

Замечание. Основываясь на теореме 9.8.1, иногда говорят, что произвольная окрестность единицы группы Ли (соответствующая локальная группа Ли) порождает всю связную компоненту единицы. Это не так. Знание локальной группы Ли недостаточно для построения всей связной компоненты единицы группы. Например, группы Ли $\mathbb{SO}(3)$ и $\mathbb{SU}(2)$ имеют одинаковые локальные группы Ли, однако не изоморфны. Дело в том, что в формуле (9.65) есть произведения большого числа элементов локальной группы Ли, которые лежат вне $\mathbb{U}$. То есть для построения связной компоненты единицы по формуле (9.65) необходимо знание правила умножения не только элементов локальной группы Ли, но и всех других элементов группы. □

Теорема 9.8.2. *Пусть $\mathbb{G}$ – группа Ли с алгеброй Ли $\mathfrak{g}$, и пусть $\tilde{\mathfrak{h}}$ – подалгебра в $\mathfrak{g}$. Тогда существует единственная связная подгруппа Ли $(f, \mathbb{H})$ группы $\mathbb{G}$ с алгеброй $\mathfrak{h}$, такая, что $f_*(\mathfrak{h}) = \tilde{\mathfrak{h}}$.*

Доказательство. Основано на теореме Фробениуса [?]. Левоинвариантные векторные поля из подалгебры $\mathfrak{h}$, согласно определению, находятся в инволюции. Из теоремы Фробениуса следует, что в каждой точке группы они задают интегральные подмногообразия. Максимальное интегральное подмногообразие, проходящее через единицу группы, и есть подгруппа Ли $(f, \mathbb{H})$. Если $\dim \tilde{\mathfrak{h}} < \dim \mathfrak{g}$, то подалгебра Ли $\tilde{\mathfrak{h}}$ обязательно является замкнутым подмножеством в $\mathbb{G}$. □

Следствие. Существует взаимно однозначное соответствие между связными подгруппами Ли группы Ли и подалгебрами ее алгебры Ли. □

Теорема 9.8.3. *Пусть $(f, \mathbb{H})$ – подгруппа Ли группы Ли $\mathbb{G}$. Тогда отображение f является регулярным вложением тогда и только тогда, когда $(f, \mathbb{H})$ – замкнутая подгруппа в $\mathbb{G}$.*

Доказательство. См., например, [?]. □

Теорема 9.8.4 (Адо). *Каждая алгебра Ли $\mathfrak{g}$ над полем комплексных чисел $\mathbb{C}$ имеет точное (инъективное) представление в алгебру Ли $\mathfrak{gl}(n, \mathbb{R})$ для некоторого n.*

Доказательство. См. [?]. □

Следствие. Для каждой алгебры Ли $\mathfrak{g}$ существует группа Ли $\mathbb{G}$ (и, в частности, односвязная группа Ли) с данной алгеброй Ли. □

Замечание. Если у алгебры Ли $\mathfrak{g}$ нулевой центр (т.е. единственный элемент, коммутирующий со всеми элементами алгебры – это нуль), то точное представление можно получить из присоединенного. Для этого просто определяем отображение

$$\mathsf{ad} \; : \quad \mathfrak{g} \ni X \; \mapsto \; \mathsf{ad}_X \in \mathsf{end}\, \mathfrak{g}$$

по формуле $\mathsf{ad}_X(Y) = [X, Y], \forall Y \in \mathfrak{g}$. После этого проверяем, что это – гомоморфизм алгебр Ли. Если у $\mathfrak{g}$ нулевой центр, то гомоморфизм инъективен, и мы получаем точное представление алгебры Ли $\mathfrak{g}$ в $\mathsf{end}\, \mathfrak{g}$. □

9.9 Экспоненциальное отображение для групп Ли

Определение. Гомоморфизм $f : \mathbb{R} \to \mathbb{G}$ вещественной прямой, рассматриваемой как группа Ли по сложению, в группу Ли называется *однопараметрической подгруппой* группы Ли $\mathbb{G}$. □

Рассмотрим отображение соответствующих алгебр Ли. Элемент алгебры Ли $\mathfrak{r}$ вещественной прямой $t \in \mathbb{R}$ имеет вид $s\partial_t \in \mathfrak{r}$, где $s \in \mathbb{R}$ и ∂_t – базисный вектор касательного пространства к $\mathbb{R}$ в нуле. Пусть $\mathfrak{g}$ – алгебра Ли группы Ли $\mathbb{G}$, которую в данном случае удобнее отождествить с касательным пространством к единице группы. Рассмотрим произвольный элемент алгебры Ли $X \in \mathfrak{g}$. Тогда отображение

$$\mathfrak{r} \ni s\partial_t \mapsto sX \in \mathfrak{g}$$

является гомоморфизмом алгебры Ли $\mathfrak{r}$ в алгебру Ли $\mathfrak{g}$. Так как вещественная прямая односвязна, то по утверждению теоремы 9.11.10 существует единственная однопараметрическая подгруппа, для которой мы введем специальное обозначение,

$$\mathsf{exp}_X : \quad \mathbb{R} \to \mathbb{G}, \tag{9.66}$$

такая, что

$$(\mathsf{exp}_X)_*(s\partial_t) = sX. \tag{9.67}$$

Определение. *Экспоненциальным отображением* алгебры Ли $\mathfrak{g}$ в группу Ли $\mathbb{G}$ называется отображение

$$\mathsf{exp} : \quad \mathfrak{g} \ni X \mapsto \mathsf{exp}\,(X) := \mathsf{exp}_X(1) \in \mathbb{G}. \quad \square \tag{9.68}$$

Происхождение такой терминологии будет ясно из дальнейшего, когда будет показано, что для линейной группы $\mathbb{GL}(n, \mathbb{R})$ экспоненциальное отображение задается показательной функцией на матрице.

Рассмотрим экспоненциальное отображение (9.68) с точки зрения интегральных кривых для векторных полей, введенных в разделе 3.6.5. Обозначим координаты в окрестности единицы группы через a^{A}, $\mathrm{A} = 1, \ldots, \mathrm{N}$, такие, что единице группы $e \in \mathbb{G}$ соответствует начало координат. Тогда отображение (9.66) в координатах задается набором функций $a^{\mathrm{A}}(t)$, которые определяются системой обыкновенных дифференциальных уравнений

$$\dot{a}^{\mathrm{A}} = X^{\mathrm{A}},$$

где X^{A} – компоненты векторного поля X, с начальными условиями $a^{\mathrm{A}}(0) = 0$. Другими словами, отображение $t \mapsto \mathsf{exp}_X(t)$ является единственной однопараметрической подгруппой в $\mathbb{G}$ с касательным вектором $X(0)$ в нуле.

Свойства экспоненциального отображения даются следующим утверждением.

Теорема 9.9.1. *Рассмотрим произвольный элемент алгебры Ли $X \in \mathfrak{g}$ на группе Ли $\mathbb{G}$. Тогда*

1) $\mathsf{exp}\,(tX) = \mathsf{exp}_X(t), \quad \forall t \in \mathbb{R}$;

2) $\mathsf{exp}\,[(t_1 + t_2)X] = \mathsf{exp}\,(t_1X)\,\mathsf{exp}\,(t_2X), \quad \forall t_1, t_2 \in \mathbb{R}$;

3) $\mathsf{exp}\,(-tX) = [\mathsf{exp}\,(tX)]^{-1}, \quad \forall t \in \mathbb{R}$;

4) отображение $\mathsf{exp} : \mathfrak{g} \to \mathbb{G}$ *гладкое, и дифференциал отображения* $\mathsf{exp}_* : \mathbb{T}_0(\mathfrak{g}) \to \mathbb{T}_e(\mathbb{G})$ *– тождественное отображение (при обычных отождествлениях), так что* exp *задает диффеоморфизм окрестности нуля в алгебре* $\mathfrak{g}$ *на окрестность единицы в группе* $\mathbb{G}$;

5) одномерное подмногообразие, полученное с помощью действия левых сдвигов $l_a \circ \exp_X \in \mathbb{G}$, – это единственная интегральная кривая левоинвариантного векторного поля X, проходящая при $t = 0$ через точку $a \in \mathbb{G}$;
6) однопараметрическая группа диффеоморфизмов X_t, порожденная левоинвариантным векторным полем X, задается правыми сдвигами

$$X_t = r_{\exp_X(t)}.$$

Доказательство. См., например, [**?**]. □

Из свойства 5), в частности, следует, что все левоинвариантные векторные поля полны.

Замечание. Согласно теореме 3.6.6, векторное поле на компактном многообразии, не обращающееся в нуль, является полным. В то же время существует много примеров неполных векторных полей на некомпактных многообразиях. В случае групп Ли каждое право- или левоинвариантное векторное поле полно, даже ясли группа Ли некомпактна. □

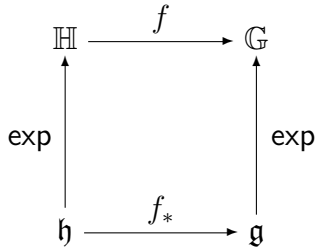
Для двух коммутирующих элементов алгебры Ли

$$[X, Y] = 0, \quad X, Y \in \mathfrak{g}$$

справедлива формула

$$\exp(X + Y) = \exp X \exp Y.$$

Теорема 9.9.2. *Пусть задан гомоморфизм групп Ли $f : \mathbb{H} \to \mathbb{G}$, тогда следующая диаграмма коммутативна*



Следствие. Любая однопараметрическая подгруппа в $\mathbb{G}$ имеет вид экспоненциального отображения $t \mapsto \exp_X(t)$ для некоторого $X \in \mathfrak{g}$. □

Если группа Ли $\mathbb{G}_0$ компактна и связна, то экспоненциальное отображение является сюръективным, т.е. отображает алгебру Ли на все многообразие группы. Как показывают явные формулы для групп $\mathbb{SO}(3)$ и $\mathbb{SU}(2)$ (см. раздел 2.8) экспоненциальное отображение не является взаимно однозначным: одному элементу группы Ли $\mathbb{G}_0$ в общем случае соответствует много различных элементов соответствующей алгебры Ли $\mathfrak{g}$. Поскольку n-мерная алгебра Ли всегда представляет собой вещественное n-мерное пространство $\mathbb{R}^n$, то при экспоненциальном отображении возникает параметризация связной компоненты единицы компактной группы Ли $\mathbb{G}_0$ точками евклидова пространства $\mathbb{R}^n$ с некоторым отношением эквивалентности. А именно, мы отождествляем те точки евклидова пространства, которые отображаются в один и тот же элемент группы. Экспоненциальное отображение является взаимно однозначным только в некоторой окрестности единицы группы $e \in \mathbb{G}_0$.

Если связная группа Ли некомпактна, то экспоненциальное отображение не является сюръективным. Пример приведен в [**?**].

9.10 Интегрирование на группах Ли

Рассмотрим группу Ли $\mathbb{G}$ размерности n. В разделе 9.2 было показано, что любая группа Ли является ориентируемым многообразием. Зафиксируем некоторую ориентацию на $\mathbb{G}$ и выберем ненулевую левоинвариантную дифференциальную n-форму ω на $\mathbb{G}$, согласованную с ориентацией. Для этого достаточно выбрать ненулевую n-форму правильной ориентации, например, в единице группы, и разнести ее по групповому многообразию с помощью действия группы слева. Напомним, что действие группы и слева, и справа сохраняет ориентацию выбранного базиса.

Пусть на группе Ли $\mathbb{G}$ задана функция f с компактным носителем. Тогда определен интеграл

$$\int_{\mathbb{G}} f\omega. \tag{9.69}$$

Этот интеграл зависит от формы ω, которая определена с точностью до постоянного положительного множителя. В случае компактных групп Ли форму ω можно однозначно фиксировать с помощью дополнительного условия

$$\int_{\mathbb{G}} \omega = 1.$$

Интеграл (9.69) от функции f является левоинвариантным. Это следует из следующей цепочки равенств:

$$\int_{\mathbb{G}} f = \int_{\mathbb{G}} f\omega = \int_{\mathbb{G}} l_a^*(f\omega) = \int_{\mathbb{G}} (f \circ l_a)\omega = \int_{\mathbb{G}} f \circ l_a,$$

где l_a – левое действие произвольного элемента $a \in \mathbb{G}$.

Поставим вопрос о том, когда интеграл (9.69) является одновременно и правоинвариантным, т.е. выполнено равенство

$$\int_{\mathbb{G}} f = \int_{\mathbb{G}} f \circ r_a.$$

Форма $r_a^*(\omega)$ по-прежнему левоинвариантна, потому что левые и правые сдвиги коммутируют между собой. Следовательно, форма $r_a^*(\omega)$ отличается от ω только постоянным положительным множителем. Это значит, что на $\mathbb{G}$ существует положительная вещественная функция $\lambda(a) > 0$ такая, что

$$r_a^*(\omega) = \lambda(a)\omega.$$

Нетрудно проверить, что

$$\lambda(ab) = \lambda(a)\lambda(b).$$

Тем самым мы показали, что λ задает гомоморфизм группы Ли $\mathbb{G}$ в мультипликативную группу положительных чисел. Функция λ называется *модулярной функцией*. Поскольку групповое действие является диффеоморфизмом, то

$$\int_{\mathbb{G}} f\omega = \int_{\mathbb{G}} (f \circ r_a)\lambda(a)\omega.$$

Отсюда следует, что левоинвариантный интеграл (9.69) является также правоинвариантным тогда и только тогда, когда $\lambda(a) = 1$ для всех $a \in \mathbb{G}$. Группа Ли $\mathbb{G}$, для которой $\lambda(a) = 1$ называется *унимодулярной*. Заметим, что каждая компактная группа Ли является унимодулярной, т.к.

$$1 = \int_{\mathbb{G}} \omega = \lambda(a) \int_{\mathbb{G}} \omega = \lambda(a).$$

Следовательно, интеграл от функции по компактной группе Ли является одновременно и право-, и левоинвариантным.

Существование лево- или правоинвариантной меры на группе Ли, *меры Хаара*, возможно только для локально компактных групп Ли.

Типичное применение интеграла по компактной группе Ли состоит в следующем.

Определение. Рассмотрим представление $\rho:\ \mathbb{G}\to \mathsf{aut}\,(\mathbb{V})$ группы Ли $\mathbb{G}$ в группе автоморфизмов вещественного или комплексного векторного пространства $\mathbb{V}$, наделенного скалярным произведением. Представление ρ называется *унитарным* ($\mathbb{V}$ комплексное) или *ортогональным* ($\mathbb{V}$ вещественное), если для всех $u,v\in\mathbb{V}$ и всех $a\in\mathbb{G}$ выполнено равенство

$$\Big(\rho(a)u,\rho(a)v\Big)=(u,v).\quad\square \tag{9.70}$$

Теорема 9.10.1. *Пусть $\mathbb{G}$ – компактная группа Ли, и $\mathbb{V}$ – комплексное (вещественное) векторное пространство. Тогда на $\mathbb{V}$ существует скалярное произведение, относительно которого представление ρ унитарно (ортогонально).*

Доказательство. Рассмотрим произвольное скалярной произведение $\langle,\rangle$ на $\mathbb{V}$. Определим новое скалярное произведение с помощью формулы

$$(u,v)=\int_{\mathbb{G}}\langle\rho(a)u,\rho(a)v\rangle,$$

где интегрирование ведется по $a\in\mathbb{G}$. Используя правую инвариантность интеграла, нетрудно убедиться в справедливости равенства (9.70). $\square$

9.11 Некоторые общие свойства групп и алгебр Ли

Напомним, что алгеброй называется множество элементов, которое является кольцом (со сложением и умножением) и одновременно векторным пространством над некоторым полем (в рассматриваемом случае над полем действительных чисел). В алгебре Ли $\mathfrak{g}$ умножение задается коммутатором, и поэтому в общем случае она является некоммутативной алгеброй. Алгебра Ли называется *абелевой*, если $[\mathfrak{g},\mathfrak{g}]=0$, т.е. любые два элемента алгебры $\mathfrak{g}$ коммутируют. Если алгебра Ли рассматривается, как кольцо, то в нем отсутствует единица. Действительно, каждая алгебра Ли содержит нетривиальную абелеву подалгебру $\mathfrak{h}\subset\mathfrak{g}$. Например, все элементы вида $X=aX_0$, где $a\in\mathbb{R}$, которые пропорциональны некоторому фиксированному элементу $X_0\in\mathfrak{g}$, образуют абелеву подалгебру. Поскольку $[\mathfrak{h},\mathfrak{h}]=0$, и нуль не является единицей по отношению к коммутатору: $[X,0]=0\neq X$ для отличных от нуля элементов $X\in\mathfrak{h}$, то единица в подалгебре $\mathfrak{h}$ и, следовательно, во всей алгебре $\mathfrak{g}$ отсутствует. Нулевой элемент алгебры Ли $0\in\mathfrak{g}$ является единичным элементом по отношению к сложению.

Пусть $\mathfrak{a}=\{A\}$ является левым идеалом в алгебре Ли $\mathfrak{g}$, т.е. $[\mathfrak{g},\mathfrak{a}]\subset\mathfrak{a}$. Тогда все элементы вида $-A$ также лежат в $\mathfrak{a}$, т.к. элементы идеала должны быть подгруппой по отношению к сложению. Поскольку для левого идеала $[\mathfrak{g},\mathfrak{a}]=-[\mathfrak{a},\mathfrak{g}]\subset\mathfrak{a}$, то каждый левый идеал совпадает с правым и наоборот. Поэтому в алгебре Ли все идеалы являются двусторонними.

Из тождества Якоби следует, что коммутатор $[\mathfrak{a},\mathfrak{b}]$ двух идеалов $\mathfrak{a},\mathfrak{b}\subset\mathfrak{g}$ также является идеалом. Поскольку сама алгебра Ли $\mathfrak{g}$ является идеалом в себе самой, то формулы

$$\mathfrak{g}^{(1)}=\mathfrak{g},\quad \mathfrak{g}^{(2)}=[\mathfrak{g},\mathfrak{g}],\quad\ldots,\quad \mathfrak{g}^{(k)}=[\mathfrak{g}^{(k-1)},\mathfrak{g}^{(k-1)}],\quad\ldots$$

определяют в алгебре Ли $\mathfrak{g}$ невозрастающую цепочку идеалов

$$\mathfrak{g} = \mathfrak{g}^{(1)} \supset \mathfrak{g}^{(2)} \supset \ldots \supset \mathfrak{g}^{(k)} \supset \ldots$$

Определение. Алгебра Ли $\mathfrak{g}$ называется разрешимой, если существует такое число $k \geq 1$, что $\mathfrak{g}^{(k)} = 0$. □

Для любых двух идеалов $\mathfrak{a} \subset \mathfrak{g}$ и $\mathfrak{b} \subset \mathfrak{g}$ их сумма $\mathfrak{a} + \mathfrak{b}$ также является идеалом в $\mathfrak{g}$. Из первой теоремы об изоморфизмах ??? следует, что фактор алгебра $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$ изоморфна фактор алгебре $\mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$ и поэтому разрешима, если идеал $\mathfrak{a}$ разрешим. Следовательно, если идеал $\mathfrak{b}$ также разрешим, то разрешим и идеал $\mathfrak{a} + \mathfrak{b}$. Таким образом, сумма $\mathfrak{a}+\mathfrak{b}$ двух разрешимых идеалов $\mathfrak{a}$ и $\mathfrak{b}$ является разрешимым идеалом. Поэтому в любой конечномерной алгебре Ли $\mathfrak{g}$ существует наибольший разрешимый идеал $\mathfrak{r}$, который называется *радикалом*, содержащий все разрешимые идеалы: им является сумма всех разрешимых идеалов.

Определение. Алгебра Ли $\mathfrak{g}$ называется полупростой, если ее радикал тривиален $\mathfrak{r} = 0$. □

Теорема 9.11.1. *Пусть $\mathfrak{r}$ – радикал алгебры Ли $\mathfrak{g}$, тогда фактор алгебра Ли $\mathfrak{g}/\mathfrak{r}$ полупроста.*

Теорема 9.11.2 (Картан). *Для того, чтобы алгебра Ли $\mathfrak{g}$ была полупростой, необходимо и достаточно, чтобы она не содержала абелевых идеалов, отличных от 0.*

Доказательство. □

Определение. *Центром* алгебры Ли $\mathfrak{g}$ называется ее аннулятор, т.е. наибольшее подпространство $\mathfrak{z} \subset \mathfrak{g}$, для которого $[\mathfrak{z}, \mathfrak{g}] = 0$. □

Очевидно, что центр алгебры Ли $\mathfrak{g}$ является абелевым идеалом, и алгебра Ли $\mathfrak{g}$ абелева тогда и только тогда, когда ее центр совпадает со всей алгеброй $\mathfrak{z} = \mathfrak{g}$. Из теоремы 9.11.2 следует, что алгебра Ли является полупростой тогда и только тогда, когда ее центр равен нулю.

Сформулируем без доказательства несколько утверждений для групп и алгебр Ли.

Теорема 9.11.3. *Для того, чтобы подгруппа Ли $\mathbb{H}$ группы Ли $\mathbb{G}$ была инвариантной, необходимо и достаточно, чтобы ее алгебра Ли была идеалом в алгебре Ли группы $\mathbb{G}$.*

Введем понятие простой алгебры и группы Ли.

Определение. Алгебра Ли $\mathfrak{g}$ называется *простой*, если выполнены следующие два условия: алгебра $\mathfrak{g}$ неабелева, и единственные ее идеалы есть 0 и $\mathfrak{g}$.

Группа Ли $\mathbb{G}$ называется *простой*, если она неабелева и не имеет инвариантных подгрупп, отличных от единицы и всей группы.

Всякие простые алгебры и группы Ли являются также полупростыми. Обратное утверждение неверно, но справедливы следующие теоремы.

Теорема 9.11.4. *Полупростая алгебра Ли* $\mathfrak{g}$ *представима одним и только одним способом в виде прямой суммы конечного числа простых идеалов*

$$\mathfrak{g} = \bigoplus_{i=1}^{k} \mathfrak{s}_i$$

для некоторого k*. При этом каждый идеал алгебры Ли* $\mathfrak{g}$ *является прямой суммой некоторых идеалов* $\mathfrak{s}_i$*.*

Теорема 9.11.5. *Полупростая группа Ли* $\mathbb{G}$ *представима одним и только одним способом в виде прямого произведения конечного числа простых групп Ли*

$$\mathbb{G} = \bigotimes_{i=1}^{k} \mathbb{G}_i.$$

При этом каждая инвариантная подгруппа группы Ли $\mathbb{G}$ *является прямым произведением некоторых подгрупп* $\mathbb{G}_i$*.*

Введем важное понятие полупрямого произведения двух групп. Для этого нам понадобится несколько предварительных определений и утверждений. Напомним, что гомоморфное отображение группы (в том числе группы Ли) на себя называется автоморфизмом.

Пример 9.11.1. Рассмотрим группу $\mathbb{G}$ и произвольный элемент $b \in \mathbb{G}$. Тогда преобразование подобия

$$a \mapsto bab^{-1},$$

как легко проверить, задает автоморфизм. Такие автоморфизмы называются *внутренними.* Все другие виды автоморфизмов называются *внешними.* □

Теорема 9.11.6. *Множество автоморфизмов группы* $\mathbb{G}$ *само образуют группу, которая называется группой автоморфизмов группы* $\mathbb{G}$ *и обозначается* $\mathsf{aut}\,\mathbb{G}$*.*

Доказательство. Проводится простой проверкой групповых аксиом, причем единицей группы $\mathsf{aut}\,\mathbb{G}$ служит тождественное отображение $\mathbb{G} \to \mathbb{G}$. □

Поскольку внутренние автоморфизмы сами образуют группу, то они образуют подгруппу группы $\mathsf{aut}\,\mathbb{G}$. Если группа $\mathbb{G}$ допускает гомоморфное отображение на свою подгруппу $\mathbb{H} \subset \mathbb{G}$, то говорят, что она допускает *эндоморфизм.*

Теорема 9.11.7. *Ядром гомоморфного отображения* $\mathbb{G}_1 \to \mathbb{G}_2$ *всегда является инвариантная подгруппа* $\mathbb{H}_1$ *группы* $\mathbb{G}_1$*. При этом имеет место изоморфизм*

$$\frac{\mathbb{G}_1}{\mathbb{H}_1} \simeq \mathbb{G}_2.$$

Пусть $\mathbb{W}$ подгруппа центра группы $\mathbb{G}$. Как и центр группы, она всегда абелева. Пусть

$$\mathbb{H} = \frac{\mathbb{G}}{\mathbb{W}}.$$

В этом случае $\mathbb{G}$ называется *накрывающей* группой группы $\mathbb{H}$. В частном случае, группа $\mathbb{G}$ является накрывающей группой своей фактор группы по своему центру. Доказательство следующих трех теорем приведено в [**?**].

Теорема 9.11.8. *Каждая связная группа Ли $\mathbb{G}$ обладает универсальным накрывающим пространством $\widetilde{\mathbb{G}}$, которое само является группой Ли, причем накрывающее отображение – гомоморфизм групп Ли.*

Теорема 9.11.9. *Пусть $\mathbb{G}$ и $\mathbb{H}$ – связные группы Ли, и $f : \mathbb{G} \to \mathbb{H}$ – гомоморфизм. Отображение f является накрытием тогда и только тогда, когда дифференциал отображения $f_* : \mathbb{T}_e(\mathbb{G}) \to \mathbb{T}_e(\mathbb{H})$ является изоморфизмом касательных пространств к единицам групп.*

Теорема 9.11.10. *Пусть $\mathbb{G}$ и $\mathbb{H}$ – связные ? группы Ли с алгебрами Ли $\mathfrak{g}$ и $\mathfrak{h}$, и пусть группа Ли $\mathbb{G}$ является односвязной. Пусть $\varphi : \mathfrak{g} \to \mathfrak{h}$ – гомоморфизм алгебр Ли. Тогда существует единственный гомоморфизм групп Ли $f : \mathbb{G} \to \mathbb{H}$, такой, что $f_* = \varphi$.*

Следствие. Если алгебры Ли связных ? и односвязных групп Ли $\mathbb{G}$ и $\mathbb{H}$ изоморфны, то сами группы $\mathbb{G}$ и $\mathbb{H}$ изоморфны. □

В последней теореме важно, чтобы группы Ли были связны, т.к. в противном случае теорема не верна.

Теорема 9.11.11. *Для любой связной группы Ли $\mathbb{G}$ ее фундаментальная группа $\pi(\mathbb{G})$ лежит в центре универсальной накрывающей группы $\widetilde{\mathbb{G}}$.*

Теорема 9.11.12. *Существует взаимно однозначное соответствие между классами изоморфизма алгебр Ли и классами изоморфизма связных ? и односвязных групп Ли.*

Прямое произведение групп (см. раздел 2.2), было определено для произвольных групп. Полупрямое произведение двух групп $\mathbb{G}$ и $\mathbb{H}$ определено только в том случае, когда $\mathbb{G}$ является группой автоморфизмов $\mathbb{H}$.

Определение. Рассмотрим две группы $\mathbb{H}$ и $\mathbb{G} \subset \mathsf{aut}\,\mathbb{H}$. Обозначим их элементы через $a, b, \ldots$ и $\alpha, \beta, \ldots$ соответственно. Тогда *полупрямым произведением* групп $\mathbb{G}$ и $\mathbb{H}$ называются все возможные упорядоченные пары (α, a) со следующим законом композиции

$$(\alpha, a)(\beta, b) = \Big(\alpha\beta, \beta(a)b\Big),$$

где $\beta(a)$ – образ элемента a при автоморфизме β. □

Пример 9.11.2. В физике важную роль играет группа Пуанкаре, рассмотренная в разделе 2.9.1, равная полупрямому произведению группы Лоренца на группу сдвигов. При этом каждому элементу группы Лоренца соответствует вращение (автоморфизм) пространства Минковского. □

Следующая теорема говорит о том как устроены некоторые некомпактные группы Ли с топологической точки зрения.

Теорема 9.11.13. *Пусть $\mathbb{G}$ – связная некомпактная полупростая группа Ли $\mathbb{G}$ и $\mathbb{K} \subset \mathbb{G}$ – ее некоторая максимальная компактная подгруппа. Тогда существует такое подмногообразие $\mathbb{U}$ в $\mathbb{G}$, диффеоморфное евклидову пространству, что отображение*

$$\mathbb{K} \times \mathbb{U} \ni (k, u) \mapsto ku \in \mathbb{G}$$

является диффеоморфизмом прямого произведения $\mathbb{K} \times \mathbb{U}$ на $\mathbb{G}$.

Доказательство. См., например, [**?**]. □

Данная теорема означает, что как многообразие любая связная некомпактная полупростая группа Ли представляет собой прямое произведение

$$\mathbb{G} \approx \mathbb{K} \times \mathbb{R}^{\dim \mathbb{G} - \dim \mathbb{K}} \tag{9.71}$$

некоторого компактного многообразия и евклидова пространства.

Пример 9.11.3. Рассмотрим произвольную $n \times n$ матрицу $M \in \mathbb{SL}(n,\mathbb{R})$ с единичным определителем. Согласно теореме о полярном разложении матриц 28.1.6, она единственным образом представляется в виде произведения $M = SA$, где $S \in \mathbb{SO}(n)$ – некоторая специальная ортогональная матрица и A – симметричная положительно определенная матрица с единичным определителем, $\det A = 1$. Группа $\mathbb{SL}(n,\mathbb{R})$ связна, некомпактна и полупроста. Группа $\mathbb{SO}(n)$ связна, компактна и является подгруппой в $\mathbb{SL}(n,\mathbb{R})$. Ее размерность равна $n(n-1)/2$. Множество симметричных положительно определенных матриц с единичным определителем диффеоморфно евклидову пространству $\mathbb{R}^{\frac{n(n+1)}{2}-1}$. Таким образом, как многообразие группа $\mathbb{SL}(n,\mathbb{R})$ диффеоморфна прямому произведению

$$\mathbb{SL}(n,\mathbb{R}) \approx \mathbb{SO}(n) \times \mathbb{R}^{\frac{n(n+1)}{2}-1}.$$ □

Если параллельный перенос отождествить с дифференциалом отображения, генерируемого векторным полем, то на связность необходимо наложить условие

$$R_{\alpha\beta\gamma}{}^{\delta} - 2\nabla_{\alpha}T_{\beta\gamma}{}^{\delta} + 2\nabla_{\beta}T_{\alpha\gamma}{}^{\delta} - 4T_{\alpha\beta}{}^{\epsilon}T_{\epsilon\gamma}{}^{\delta} - 4T_{\gamma\alpha}{}^{\epsilon}T_{\epsilon\beta}{}^{\delta} - 4T_{\beta\gamma}{}^{\epsilon}T_{\epsilon\alpha}{}^{\delta} = 0. \tag{9.72}$$

9.12 Группа Ли $\mathbb{GL}(n,\mathbb{C})$

В настоящем разделе мы приведем некоторые сведения из теории линейных групп Ли $\mathbb{GL}(n,\mathbb{C})$ преобразований комплексного n-мерного пространства $\mathbb{C}^n$ и их алгебр Ли $\mathfrak{gl}(n,\mathbb{C})$, которые важны с точки зрения дифференциальной геометрии и физических приложений. Напомним, что при общих преобразованиях координат в касательных пространствах к точкам многообразия действует матрица Якоби (2.57), которая является элементом группы $\mathbb{GL}(n,\mathbb{R}) \subset \mathbb{GL}(n,\mathbb{C})$. Кроме того удобным выбором независимых переменных в аффинной геометрии являются переменные Картана: репер и линейная или $\mathbb{GL}(n,\mathbb{R})$-связность (см. главу **??**).

Сначала мы рассмотрим алгебры Ли, как более простые объекты, а затем перейдем к изучению соответствующих групп Ли.

9.12.1 Алгебра Ли $\mathfrak{gl}(n,\mathbb{C})$

Обозначим множество всех квадратных $n \times n$ матриц с комплексными элементами через $\mathfrak{gl}(n,\mathbb{C})$. Это обозначение обосновано в дальнейшем тем, что множество комплексных квадратных матриц можно рассматривать, как алгебру Ли группы Ли $\mathbb{GL}(n,\mathbb{C})$. Множество матриц $\mathfrak{gl}(n,\mathbb{C})$ с обычным умножением матриц на числа и сложением можно рассматривать как векторное пространство размерности n^2 над полем комплексных чисел (комплексная размерность) или как вещественное векторное пространство размерности $2n^2$ (вещественная размерность). Чтобы превратить это множество в алгебру необходимо ввести дополнительную бинарную операцию –

умножение. Если в качестве умножения рассматривать обычное умножение матриц, то мы получим ассоциативную алгебру над полем комплексных чисел. Однако это не единственная возможность. В качестве алгебраической операции мы будем рассматривать коммутатор матриц, т.е. любым двум матрицам $X, Y \in \mathfrak{gl}(n,\mathbb{C})$ мы ставим в соответствие их коммутатор

$$\mathfrak{gl}(n,\mathbb{C}) \times \mathfrak{gl}(n,\mathbb{C}) \ni X, Y \mapsto [X,Y] = XY - YX \ \in \mathfrak{gl}(n,\mathbb{C}).$$

Эта операция антисимметрична, $[X,Y] = -[Y,X]$, и для нее выполняется тождество Якоби

$$\Big[[X,Y],Z\Big] + \Big[[Y,Z],X\Big] + \Big[[Z,X],Y\Big] = 0.$$

Тем самым множество всех матриц, включая вырожденные, становится алгеброй Ли $\mathfrak{gl}(n,\mathbb{C})$. Эта алгебра неассоциативна, что следует из тождеств Якоби.

В дальнейшем мы всегда будем рассматривать множество комплексных $n \times n$ матриц, как алгебру Ли $\mathfrak{gl}(n,\mathbb{C})$.

Множество матриц можно рассматривать, как множество операторов, действующих в комплексном векторном пространстве $\mathbb{V}$. Обозначим элементы векторного пространства через $\boldsymbol{x} = x^a \boldsymbol{e}_a \in \mathbb{V}$, где $\boldsymbol{e}_a$, $a = 1, \ldots, n$ – некоторый фиксированный базис, и $x^a \in \mathbb{C}$ – комплексные компоненты вектора. Комплексная размерность этого векторного пространства равна n, а вещественная – $2n$. При фиксированном базисе векторное пространство $\mathbb{V}$ естественным образом отождествляется с комплексным евклидовым пространством $\mathbb{C}^n$ и вещественным евклидовым пространством удвоенной размерности $\mathbb{R}^{2n}$. Мы предполагаем, что топология $\mathbb{V}$ индуцируется взаимно однозначным отображением $\mathbb{V} \leftrightarrow \mathbb{R}^{2n}$.

Преобразование элементов векторного пространства $\mathbb{C}^n$ мы записываем в виде

$$x^a \mapsto x^b X_b{}^a, \qquad X = \{X_b{}^a\} \in \mathfrak{gl}(n,\mathbb{C}),$$

т.е. вектор-строка $\{x^a\}$ умножается справа на матрицу преобразований $X = \{X_b{}^a\}$. Такая запись вызвана принятыми ранее правилами: компоненты вектора мы нумеруем верхним индексом и придерживаемся правила записи индексов суммирования “с десяти до четырех”.

Базис алгебры Ли $\mathfrak{gl}(n,\mathbb{C})$ можно выбрать из квадратных матриц $\boldsymbol{e}_a{}^b$, имеющих один ненулевой элемент, который равен единице и находится на a-той строке и в b-том столбце. Очевидно, что любую матрицу можно записать в виде

$$X = X_a{}^b \boldsymbol{e}_b{}^a \ \in \ \mathfrak{gl}(n,\mathbb{C}), \qquad X_a{}^b \in \mathbb{C}.$$

Выбранный базис удовлетворяет соотношениям коммутации

$$[\boldsymbol{e}_a{}^b, \boldsymbol{e}_c{}^d] = \delta_c^b \boldsymbol{e}_a{}^d - \delta_a^d \boldsymbol{e}_c{}^b, \tag{9.73}$$

что проверяется прямой проверкой.

Рассмотренное представление естественным образом отождествляет алгебру Ли $\mathfrak{gl}(n,\mathbb{C})$ с комплексным евклидовым пространством $\mathbb{C}^{n^2}$ и вещественным евклидовым пространством $\mathbb{R}^{2n^2}$. Мы предполагаем, что топология алгебры Ли $\mathfrak{gl}(n,\mathbb{C})$ индуцирована вложением $\mathfrak{gl}(n,\mathbb{C}) \hookrightarrow \mathbb{R}^{2n^2}$.

Базис $\boldsymbol{e}_a{}^b$ алгебры Ли $\mathfrak{gl}(n,\mathbb{C})$ можно также представить в виде дифференциальных операторов

$$\boldsymbol{e}_a{}^b = -x^b \partial_a,$$

действующих на компоненты векторов из $\mathbb{V}$. Это представление определяет действие генераторов алгебры Ли $\mathfrak{gl}(n,\mathbb{C})$ на любую дифференцируемую функцию от декартовых координат точки $\mathbb{C}^n$.

Определение. Матрица A называется *невырожденной* или *регулярной*, если для нее существует обратная матрица A^{-1}, т.е. выполнено равенство $AA^{-1} = A^{-1}A = \mathbb{1}$. □

Для того, чтобы матрица A была регулярной необходимо и достаточно, чтобы ее определитель был отличен от нуля, $\det A \neq 0$. Если эндоморфизм A векторного пространства $\mathbb{C}^n$ отображает $\mathbb{C}^n$ *на* себя (сюръективен), а не на какое-нибудь подпространство низшей размерности, то соответствующая матрица A регулярна, и существует обратный эндоморфизм A^{-1}. В этом случае эндоморфизм A является автоморфизмом.

Замечание. Пусть $A = \{A_a{}^b\}$ – квадратная $n \times n$ матрица, тогда для определителя справедливо разложение, в частности, по первой строке

$$\det A = \sum_\sigma A_1{}^{\sigma(1} \dots A_n{}^{n)} \operatorname{sgn}\sigma,$$

где сумма берется по всем перестановкам $\sigma(1,\dots,n)$. □

Перепишем соотношения коммутации (9.73) в виде

$$[\boldsymbol{e}_{\text{A}}, \boldsymbol{e}_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} \boldsymbol{e}_{\text{C}},$$

где пару индексов мы для краткости обозначили одной буквой $\boldsymbol{e}_{\text{A}} = \boldsymbol{e}_a{}^b$, $\text{A} = 1,\dots,n^2$. Тогда

$$f_{\text{AB}}{}^{\text{C}} = f_a{}^b{}_c{}^d{}_e{}^f = \delta^{fbd}_{ace} - \delta^{dfb}_{ace}.$$

Простые вычисления приводят к следующей форме Киллинга–Картана

$$\eta_{\text{AB}} = f_{\text{AC}}{}^{\text{D}} f_{\text{BD}}{}^{\text{C}} = \eta_a{}^b{}_c{}^d = 2\left(n\delta^{db}_{ac} - \delta^{bd}_{ac}\right).$$

Форма Киллинга–Картана задает инвариантное скалярное произведение в алгебре Ли $\mathfrak{gl}(n,\mathbb{C})$:

$$(X,Y) = 2n \operatorname{tr}(XY) - 2 \operatorname{tr} X \operatorname{tr} Y. \tag{9.74}$$

Инвариантность в данном случае означает независимость результата скалярного произведения от преобразования подобия:

$$(X',Y') = (X,Y),$$

где

$$X' = SXS^{-1}, \quad Y' = SYS^{-1}, \qquad \forall S \in \mathbb{GL}(n,\mathbb{C}).$$

Скалярное произведение (9.74) вырождено. Действительно, поскольку $\operatorname{tr}\mathbb{1} = n$, то скалярное произведение всех матриц, кратных единичной матрице $E = a\mathbb{1}$, $a \in \mathbb{C}$, равно нулю со всеми матрицами из алгебры $\mathfrak{gl}(n,\mathbb{C})$,

$$(E,X) = a(\mathbb{1},X) = 2an \operatorname{tr} X - 2an \operatorname{tr} X = 0, \qquad \forall X \in \mathfrak{gl}(n,\mathbb{C}).$$

Отсюда следует, что форма Киллинга–Картана вырождена $\det \eta_{\text{AB}} = 0$, и поэтому алгебра Ли $\mathfrak{gl}(n,\mathbb{C})$ не является полупростой. Это значит, что на групповом многообразии группы Ли $\mathbb{GL}(n,\mathbb{C})$ не существует двусторонне инвариантной метрики и, следовательно, возникают серьезные проблемы с построением инвариантов.

Поскольку единичная матрица коммутирует со всеми матрицами, то множество матриц, кратных единице, $\mathbb{E} = \{E\}$ является центром и образует идеал в алгебре Ли $\mathfrak{gl}(n,\mathbb{C})$. Его вещественная размерность равна двум, $\dim \mathbb{E} = 2$.

Алгебра Ли $\mathfrak{gl}(n,\mathbb{C})$ содержит много подалгебр. Рассмотрим некоторые подалгебры, которые наиболее часто встречаются в приложениях. Нетрудно видеть, что множество всех квадратных вещественных $n\times n$ матриц также является алгеброй Ли, которую мы обозначим через $\mathfrak{gl}(n,\mathbb{R})$. Эта алгебра является подалгеброй в алгебре комплексных матриц $\mathfrak{gl}(n,\mathbb{R})\subset\mathfrak{gl}(n,\mathbb{C})$. Базис $\boldsymbol{e}_a{}^b$ в $\mathfrak{gl}(n,\mathbb{R})$ можно выбрать таким же, как и в случае $\mathfrak{gl}(n,\mathbb{C})$, только коэффициенты разложения теперь будут не комплексные, а действительные числа. Таким образом алгебра Ли $\mathfrak{gl}(n,\mathbb{R})$ естественным образом отождествляется с евклидовым пространством $\mathbb{R}^{n^2}$. Формы Киллинга–Картана для $\mathfrak{gl}(n,\mathbb{C})$ и $\mathfrak{gl}(n,\mathbb{R})$ совпадают и, следовательно, алгебра Ли $\mathfrak{gl}(n,\mathbb{R})$ не является полупростой.

Максимальные полупростые подалгебры – это алгебры комплексных и вещественных матриц с нулевым следом:

$$\mathfrak{sl}(n,\mathbb{C})=\{A\in\mathfrak{gl}(n,\mathbb{C}):\ \ \mathsf{tr}\,A=0\},\tag{9.75}$$

$$\mathfrak{sl}(n,\mathbb{R})=\{A\in\mathfrak{gl}(n,\mathbb{R}):\ \ \mathsf{tr}\,A=0\}.\tag{9.76}$$

Они имеют следующие вещественные размерности:

$$\mathsf{dim}\,\mathfrak{sl}(n,\mathbb{C})=2n^2-2,\qquad \mathsf{dim}\,\mathfrak{sl}(n,\mathbb{R})=n^2-1.$$

Каждую матрицу можно однозначно представить в виде

$$X=\widetilde{X}+\frac{1}{n}\,\mathsf{tr}\,X\mathbb{1},\quad \mathsf{tr}\,\widetilde{X}=0,$$

выделив из нее след. Будем считать две матрицы X и Y эквивалентными, если равны их бесследовые части, $\widetilde{X}=\widetilde{Y}$. Такие матрицы связаны преобразованием

$$X=Y+\frac{1}{n}(\,\mathsf{tr}\,X-\,\mathsf{tr}\,Y)\mathbb{1}.$$

Поэтому

$$\mathfrak{sl}(n,\mathbb{C})\simeq\frac{\mathfrak{gl}(n,\mathbb{C})}{\mathbb{E}},$$

где $\mathbb{E}$ – идеал в алгебре Ли $\mathfrak{gl}(n,\mathbb{C})$, состоящий из матриц, кратных единичной.

Чтобы доказать полупростоту этих подалгебр, вычислим форму Киллинга–Картана. Выберем базис в алгебрах Ли $\mathfrak{sl}(n,\mathbb{C})$ и $\mathfrak{sl}(n,\mathbb{R})$

$$\tilde{\boldsymbol{e}}_a{}^b=\boldsymbol{e}_a{}^b-\delta^b_a\boldsymbol{e}_n{}^n,\quad a,b=1,\ldots,n,\qquad\text{(суммирования по } n \text{ нет)},$$

который удовлетворяет условию $\mathsf{tr}\,\tilde{\boldsymbol{e}}_a{}^b=0$. Это соотношение определяет только n^2-1 базисных векторов, поскольку $\boldsymbol{e}_n{}^n=0$. Коммутационные соотношения принимают вид

$$[\tilde{\boldsymbol{e}}_a{}^b,\tilde{\boldsymbol{e}}_c{}^d]=\delta^b_c\tilde{\boldsymbol{e}}_a{}^d-\delta^d_a\tilde{\boldsymbol{e}}_c{}^b-\delta^{bn}_{ac}\tilde{\boldsymbol{e}}_n{}^d+\delta^{bd}_{an}\tilde{\boldsymbol{e}}_c{}^n-\delta^{db}_{cn}\tilde{\boldsymbol{e}}_a{}^n+\delta^{dn}_{ca}\tilde{\boldsymbol{e}}_n{}^b.$$

Теперь нетрудно вычислить форму Киллинга–Картана

$$\eta_{\text{AB}}=\eta_a{}^b{}_c{}^d=2n\left(\delta^{db}_{ac}+\delta^{bd}_{ac}-\delta^{bnd}_{acn}-\delta^{dnb}_{can}\right).$$

Отсюда следует, что, если хотя бы одна из пар индексов имеет вид $(a,b)=(n,n)$ или $(c,d)=(n,n)$, то

$$\eta_a{}^b{}_n{}^n=\eta_n{}^n{}_c{}^d=0.$$

Форма Киллинга–Картана задает инвариантное скалярное произведение в алгебрах Ли:

$$(X,Y) = 2n\,\mathsf{tr}\,(XY), \qquad X,Y \in \mathfrak{sl}(n,\mathbb{C}) \;\text{ или }\; X,Y \in \mathfrak{sl}(n,\mathbb{R}).$$

Эта форма Киллинга–Картана невырождена, и, значит, алгебры матриц с нулевым следом полупросты. Можно доказать более сильное утверждение.

Теорема 9.12.1. *Алгебра Ли $\mathfrak{sl}(n,\mathbb{R})$ проста.*

Доказательство. □

Определение. Пусть $A \in \mathfrak{gl}(n,\mathbb{C})$ и $\lambda \in \mathbb{C}$. Алгебраическое уравнение

$$\mathsf{det}\,(A - \lambda \mathbb{1}) = 0 \tag{9.77}$$

n-того порядка относительно λ называется *характеристическим* или *вековым* уравнением. Решения этого уравнения называются *собственными числами* матрицы A. □

Согласно основной теореме алгебры каждая матрица A имеет в точности n собственных чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ с учетом их кратности. В общем случае собственные числа комплексны даже для вещественной матрицы $A \in \mathfrak{gl}(n,\mathbb{R})$.

Матрицы A и A' называются *подобными* или *сопряженными* , если существует невырожденная матрица B такая, что $A' = BAB^{-1}$. Из уравнения (9.77) следует, что собственные числа подобных матриц совпадают.

Опишем экспоненциальное отображение матриц.

Определение. Пусть $A \in \mathfrak{gl}(n,\mathbb{C})$ – произвольная матрица с ограниченными элементами. Тогда ряд

$$\mathsf{e}^A = \mathsf{exp}\,A := \mathbb{1} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

всегда сходится, и притом равномерно, когда A остается в ограниченной области пространства $\mathfrak{gl}(n,\mathbb{C}) \approx \mathbb{R}^{2n^2}$. Этот ряд называется *экспоненциалом* матрицы A. □

Функция $\mathsf{exp}\,A$ определена и непрерывна на $\mathfrak{gl}(n,\mathbb{C})$ и отображает $\mathfrak{gl}(n,\mathbb{C})$ в себя.

Сформулируем некоторые свойства экспоненциала матрицы [**?**].

Предложение 9.12.1. *Пусть B – невырожденная $n \times n$ матрица. Тогда*

$$\mathsf{e}^{BAB^{-1}} = B\,\mathsf{e}^A B^{-1}.$$

Предложение 9.12.2. *Для произвольной матрицы $A \in \mathfrak{gl}(n,\mathbb{C})$ существует такая невырожденная матрица B, что матрица BAB^{-1} является верхнетреугольной. Тогда матрица $\mathsf{e}^{BAB^{-1}}$ также является верхнетреугольной. При этом на диагонали матрицы BAB^{-1} стоят собственные числа $\lambda_1, \ldots, \lambda_n$ матрицы A, а на диагонали матрицы $\mathsf{e}^{BAB^{-1}}$ стоят собственные числа $\mathsf{e}^{\lambda_1}, \ldots, \mathsf{e}^{\lambda_n}$ матрицы e^A.*

Следствие. Для любой матрицы $A \in \mathfrak{gl}(n,\mathbb{C})$ справедливо равенство

$$\mathsf{det}\,(\,\mathsf{exp}\,A) = \mathsf{exp}\,(\,\mathsf{tr}\,A) \quad \Leftrightarrow \quad \mathsf{det}\,B = \mathsf{exp}\,(\,\mathsf{tr}\,\ln B), \tag{9.78}$$

где $B = \mathsf{e}^A$. □

Отсюда следует, в частности, что $\det(\exp A) \neq 0$. То есть экспоненциал произвольной матрицы является невырожденной (регулярной) матрицей и, следовательно, принадлежит группе общих преобразований $\exp A \in \mathbb{GL}(n,\mathbb{C})$ (см. следующий раздел).

Предложение 9.12.3. *Если матрицы A, B коммутируют, то*

$$\mathrm{e}^{A+B} = \mathrm{e}^{A}\,\mathrm{e}^{B}.$$

Следствие. Экспоненциальное отображение $t \mapsto \mathrm{e}^{tA}$, где $t \in \mathbb{R}$ и $A \in \mathfrak{gl}(n,\mathbb{C})$ – фиксированная матрица, есть гладкое гомоморфное отображение аддитивной группы вещественных чисел в группу $\mathbb{GL}(n,\mathbb{C})$. □

Касательным вектором к отображению $t \mapsto \mathrm{e}^{tA}$ в точке $t = 0$ является матрица A (достаточно почленно продифференцировать степенной ряд). Экспоненциальное отображение $\mathbb{R} \to \mathbb{GL}(n,\mathbb{C})$ является единственной однопараметрической подгруппой в $\mathbb{GL}(n,\mathbb{C})$ с касательным вектором A в нуле.

Приведем несколько очевидных формул для экспоненциального отображения матриц

$$\exp A^T = (\exp A)^T,$$
$$\exp A^\dagger = (\exp A)^\dagger,$$
$$\exp(-A) = (\exp A)^{-1},$$

где индексы T и $\dagger$ обозначают транспонирование и эрмитово сопряжение матриц.

Экспоненциальное отображение матриц аналогично экспоненциальному отображению, которое генерируется векторными полями (см. раздел 9.9). Оно задает отображение алгебры Ли $\mathfrak{gl}(n,\mathbb{C})$ в группу Ли $\mathbb{GL}(n,\mathbb{C})$.

Предложение 9.12.4. *В алгебре $\mathfrak{gl}(n,\mathbb{C})$ существует окрестность $\mathbb{U}$ нулевой матрицы $0 \in \mathbb{U} \subset \mathfrak{gl}(n,\mathbb{C})$, которая удовлетворяет следующим условиям:*

1) при экспоненциальном отображении $A \mapsto \exp A$ окрестность $\mathbb{U}$ непрерывно отображается на некоторую окрестность единичной матрицы $\mathbb{1} \in \mathbb{GL}(n,\mathbb{C})$;

2) $|\operatorname{tr} A| < 2\pi$;

3) при $A \in \mathbb{U}$ справедливы включения $-A, A^T, A^\dagger \in \mathbb{U}$.

При построении экспоненциального отображение поле комплексных чисел можно заменить на поле вещественных чисел. При этом конструкция не изменится, если собственные числа матрицы с вещественными элементами также вещественны. (В общем случае это, конечно, не так.)

Экспоненциальное отображение задает отображение множества эндоморфизмов векторного пространства в множество его автоморфизмов

$$\exp : \quad \mathrm{end}(\mathbb{V}) \to \mathrm{aut}(\mathbb{V}),$$

где $\mathbb{V}$ – произвольное векторное пространство над полем вещественных или комплексных чисел.

Можно доказать [**?**], что экспоненциальное отображение для группы $\mathbb{GL}(n,\mathbb{C})$ является сюръективным. В то же время для группы общих линейных преобразований над полем вещественных чисел $\mathbb{GL}(n,\mathbb{R})$ это не так.

Пример 9.12.1. Матрицу

$$\begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$$

нельзя представить в виде e^A ни для какой $A \in \mathfrak{gl}(2,\mathbb{R})$. □

9.12.2 Группа Ли $\mathbb{GL}(n,\mathbb{C})$

Определение. Группой *общих линейных преобразований* n-мерного векторного пространства над полем комплексных чисел $\mathbb{GL}(n,\mathbb{C})$ называется множество всех невырожденных квадратных $n\times n$ матриц с комплексными элементами и обычным правилом умножения. Множество всех невырожденных $n\times n$ матриц с вещественными элементами образует группу общих линейных преобразований n-мерного вещественного пространства $\mathbb{GL}(n,\mathbb{R})$. □

Единственное условие, которое накладывается на матрицы – это отличие от нуля их определителей, которое необходимо и достаточно для существования обратных матриц.

Множество всех, в том числе вырожденных, комплексных $n\times n$ матриц, как многообразие, представляет собой евклидово пространство $\mathbb{R}^{2n^2}$, т.к. параметризуется n^2 комплексными числами, на которые не наложено никаких ограничений. Поскольку определитель матрицы является непрерывной функцией от матрицы, то множество матриц с нулевым определителем представляет собой замкнутое подмножество в $\mathbb{R}^{2n^2}$, поскольку 0 является замкнутым подмножеством в $\mathbb{R}$. Поэтому множество невырожденных матриц $\mathbb{GL}(n,\mathbb{C})$ как дополнение замкнутого подмножества образует в евклидовом пространстве $\mathbb{R}^{2n^2}$ открытое подмногообразие и, следовательно, имеет ту же размерность $2n^2$, что и евклидово пространство.

Нетрудно проверить, что групповая операция является гладкой, и, следовательно, группа $\mathbb{GL}(n,\mathbb{C})$ является группой Ли. Эта группа связна некомпактна, но локально компактна.

Аналогично, группа всех невырожденных вещественных $n\times n$ матриц, как многообразие, представляет собой открытое подмногообразие в евклидовом пространстве $\mathbb{R}^{n^2}$ с гладкой групповой операцией. В отличие от группы комплексных матриц $\mathbb{GL}(n,\mathbb{C})$, группа вещественных матриц не является связной. Она состоит из двух связных компонент: матриц с положительным и отрицательным определителем. Множество матриц с положительным определителем $\mathbb{GL}_+(n,\mathbb{R})$ само является группой Ли и представляет собой связную компоненту единицы в $\mathbb{GL}(n,\mathbb{R})$. Вторая компонента представляет собой ее смежный класс и состоит из матриц с отрицательным определителем.

Это различие легко понять. В комплексном случае матрицу A с вещественным положительным определителем, $\det A>0$, можно непрерывно трансформировать в матрицу B с вещественным отрицательным определителем, $\det B<0$, вдоль пути, целиком лежащим в $\mathbb{GL}(n,\mathbb{C})$ и состоящим из матриц с комплексным определителем. Здесь прослеживается связь с вещественной прямой и комплексной плоскостью. Если из вещественной прямой удалить точку 0, то она распадается на два несвязных многообразия. В то же время, удаление начала координат из комплексной плоскости оставляет ее связной.

Очевидно что группа вещественных матриц $\mathbb{GL}(n,\mathbb{R})$ является подгруппой в $\mathbb{GL}(n,\mathbb{C})$. Поскольку размерность $\dim\mathbb{GL}(n,\mathbb{R})<\dim\mathbb{GL}(n,\mathbb{C})$, то согласно теоремам 3.10.1 и 3.10.2 эта подгруппа замкнута.

Произвольную невырожденную матрицу $A=\{A_a{}^b\}$, $a,b,\ldots=1,2,\ldots,n$ можно представить в виде

$$A=a^{1/n}B,\qquad a=\det A,$$

где

$$\det B=1.$$

Определение. Совокупность всех квадратных $n\times n$ матриц с комплексными элементами, определитель которых равен единице, является группой, которая называется группой *специальных линейных преобразований* n-мерного векторного пространства над полем комплексных чисел и обозначается $\mathbb{SL}(n,\mathbb{C})$. Если элементы матриц вещественны, то соответствующая группа обозначается $\mathbb{SL}(n,\mathbb{R})$. Она является подгруппой в $\mathbb{SL}(n,\mathbb{C})$. □

Очевидно, что $\mathbb{SL}(n,\mathbb{C}) \subset \mathbb{GL}(n,\mathbb{C})$. Это максимальная полупростая подгруппа группы $\mathbb{GL}(n,\mathbb{C})$ размерности $2n^2-1$. Можно доказать, что она является простой. Алгебра Ли $\mathfrak{sl}(n,\mathbb{C})$ изоморфна алгебре комплексных матриц с нулевым следом.

Группа $\mathbb{SL}(n,\mathbb{R})$ является подгруппой группы общих линейных преобразований $\mathbb{GL}(n,\mathbb{R})$ и является связной некомпактной, но локально компактной группой Ли размерности n^2-1. Односвязность ??? Эта группа является максимальной полупростой подгруппой группы $\mathbb{GL}(n,\mathbb{R})$ (и простой). Алгебра Ли $\mathfrak{sl}(n,\mathbb{R})$ изоморфна алгебре вещественных матриц с нулевым следом.

Вблизи единичной матрицы $\mathbb{1} \in \mathbb{SL}(n,\mathbb{K})$, где $\mathbb{K}=\mathbb{R}$ или $\mathbb{K}=\mathbb{C}$, элемент группы $S \in \mathbb{SL}(n,\mathbb{K})$ взаимно однозначно представляется экспоненциальным отображением элемента алгебры $A \in \mathfrak{sl}(n,\mathbb{K})$:

$$S = \mathsf{e}^A, \qquad \mathsf{tr}\, A = 0.$$

У группы общих преобразований $\mathbb{GL}(n,\mathbb{C})$ есть много других подгрупп.

Пример 9.12.2. Группа вещественных ортогональных $n\times n$ матриц $\mathbb{O}(n)$ состоит из матриц $S \in \mathbb{O}(n) \subset \mathbb{GL}(n,\mathbb{R})$, удовлетворяющих условию ортогональности

$$SS^T = S^T S = 1,$$

где S^T обозначает транспонированную матрицу. След от этого равенства приводит к ограничению на матричные элементы $S_{ij}S^{ij} = n$. Это значит, что групповое многообразие является замкнутым подмножеством сферы $\mathbb{S}^{n^2-1} \subset \mathbb{R}^{n^2}$. Отсюда, по теореме Гейне–Бореля, следует компактность группы вращений. Группа $\mathbb{O}(n)$ является компактной группой Ли размерности $n(n-1)/2$ и состоит из двух компонент. Матрицы с единичным определителем образуют связную компоненту единицы $\mathbb{SO}(n)$ (собственные вращения). Вторую компоненту образует ее смежный класс, состоящий из несобственных ортогональных матриц с определителем, равным -1. Группа $\mathbb{SO}(n)$ не является односвязной.

Группа двумерных вращений $\mathbb{SO}(2)$ как многообразие представляет собой окружность $\mathbb{S}^1$. Ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$. Полная группа вращений $\mathbb{O}(2)$ состоит из двух несвязных компонент (см. раздел 2.6). Как многообразие она представляет собой прямое произведение окружностей $\mathbb{O}(2) = \mathbb{S}^1 \times \mathbb{S}^1$ с фундаментальной группой $\mathbb{Z}\times\mathbb{Z}$, где каждый из сомножителей действует на своей окружности.

Группа трехмерных вращений $\mathbb{SO}(3)$ является группой симметрии сферы $\mathbb{S}^2$, вложенной в трехмерное евклидово пространство $\mathbb{R}^3$. Каждое вращение можно параметризовать вектором $\vec{n}$, направленным вдоль оси вращения, длина которого равна углу поворота против часовой стрелки $|\vec{n}| \le \pi$. Множество векторов поворота заполняет шар $\mathbb{B}^3_\pi \subset \mathbb{R}^3$. При этом противоположные точки граничной сферы $\mathbb{S}^2$ необходимо отождествить, т.к. поворот на угол π совпадает с поворотом на угол $-\pi$. Таким образом как многообразие группа $\mathbb{SO}(3)$ представляет собой трехмерный

шар с отождествленными противоположными точками граничной сферы. Фундаментальная группа группы $\mathbb{SO}(3)$ является группой второго порядка $\mathbb{Z}_2$. Полная группа трехмерных вращений $\mathbb{O}(3)$ как многообразие состоит из двух одинаковых несвязных компонент с фундаментальной группой $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Группы вращений $\mathbb{SO}(n)$ при $n \geq 4$ являются группами симметрии сферы $\mathbb{S}^{n-1}$, вложенной в евклидово пространство $\mathbb{R}^n$. Их многообразия устроены гораздо сложнее. Можно доказать, что фундаментальные группы групп вращений $\mathbb{SO}(n)$ при всех $n \geq 3$ равны $\mathbb{Z}_2$. Полная группа вращений $\mathbb{O}(n)$ как многообразие состоит из двух несвязных компонент с фундаментальной группой $\mathbb{Z}_2 \times \mathbb{Z}_2$. ???

Алгебры Ли групп $\mathbb{O}(n)$ и $\mathbb{SO}(n)$ совпадают и изоморфны алгебре антисимметричных $n \times n$ матриц $\mathfrak{so}(n)$. Вблизи единицы $\mathbb{1} \in \mathbb{SO}(n)$ матрица вращений $S \in \mathbb{SO}(n)$ взаимно однозначно представляется в виде экспоненциала отображения некоторого элемента алгебры $A \in \mathfrak{so}(n)$

$$S = \mathsf{e}^A, \qquad A = -A^T. \qquad \square$$

Пример 9.12.3. Пусть в декартовых координатах евклидова пространства $\mathbb{R}^n$ задана метрика, представляющая собой диагональную матрицу, у которой на диагонали стоят p положительных и q отрицательных единиц

$$\eta_{ab} = \mathsf{diag}\,(\underbrace{+\ldots+}_{p}\underbrace{-\ldots-}_{q}), \qquad p+q=n. \tag{9.79}$$

Для определенности мы считаем, что все положительные единицы идут вначале. Сигнатура этой метрики равна $p-q$. Матрицы преобразований евклидова пространства $S_a{}^b$, сохраняющие метрику (9.79),

$$\eta_{ab} = S_a{}^c S_b{}^d \eta_{cd}, \tag{9.80}$$

образуют группу, которая называется группой вращений $\mathbb{O}(p,q)$. В частности, при $q=0$ группа $\mathbb{O}(n,0)=\mathbb{O}(n)$, а при $p=1$ имеем группу Лоренца $\mathbb{O}(1,n-1)$. Из определения (9.80) следует, что $\det S = \pm 1$. То есть группа вращений $\mathbb{O}(p,q)$ состоит по крайней мере из двух компонент: матриц с положительным и отрицательным определителем. Можно показать, что при $q \neq 0, n$ группа вращений $\mathbb{O}(p,q)$, как и группа Лоренца, состоит из четырех несвязных компонент и содержит три нетривиальные подгруппы. ??? При $q \neq 0, n$ группа вращений $\mathbb{O}(p,q)$ является некомпактной.

Группа вращений $\mathbb{O}(p,q)$ является группой симметрии пространств постоянной кривизны: гиперболоидов $\mathbb{H}$ размерности $\dim \mathbb{H} = n-1$, вложенных в псевдоевклидово пространство $\mathbb{R}^{p,q}$ уравнением

$$(x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots - (x^n)^2 = \pm r^2, \qquad r = \mathsf{const}.$$

Это утверждение является следствием определения (9.80). В зависимости от знака правой части мы получаем различные гиперболоиды, которые в общем случае состоят из нескольких несвязных компонент. $\square$

Пример 9.12.4. Рассмотрим комплексное векторное пространство $\mathbb{C}^n$ с естественным базисом $\boldsymbol{e}_a$, $a = 1,\ldots,n$. Пусть $X = X^a \boldsymbol{e}_a$ и $Y = Y^a \boldsymbol{e}_a$ – два произвольных вектора из $\mathbb{C}^n$. Определим *эрмитово* скалярное произведение в комплексном векторном пространстве $\mathbb{C}^n$ следующим соотношением

$$(X,Y) = X^{\dagger a} Y^b \delta_{ab}, \tag{9.81}$$

где символ † обозначает комплексное сопряжение. Отсюда следует, что $(X,X)\geq 0$ и $(X,X)=0$ тогда и только тогда, когда $X=0$. Вещественное число

$$\|X\|=\sqrt{(X,X)}$$

называется *длиной* вектора $X\in\mathbb{C}^n$.

Группа унитарных матриц $\mathbb{U}(n)$ состоит из комплексных $n\times n$ матриц $U\in\mathbb{U}(n)$, которые сохраняют эрмитово скалярное произведение (9.81). В матричных обозначениях это условие записывается в виде *условия унитарности*

$$UU^\dagger=U^\dagger U=1,$$

где символ † обозначает эрмитово сопряжение (транспонирование и комплексное сопряжение всех элементов). Из условия унитарности следует, что определитель унитарной матрицы равен по модулю единице. Группа $\mathbb{U}(n)$ является связной группой Ли размерности $\dim\mathbb{U}(n)=n^2$. Так же как и группа вещественных ортогональных матриц эта группа компактна. Группа $\mathbb{U}(n)$ не является односвязной, и ее фундаментальная группа изоморфна группе целых чисел $\mathbb{Z}$ [?]. Как многообразие группа $\mathbb{U}(n)$ диффеоморфна прямому произведению окружности на группу специальных унитарных матриц (см. пример 9.12.5) $\mathbb{U}(n)\approx\mathbb{S}^1\times\mathbb{SU}(n)$.

При $n=1$ имеем абелеву группу $\mathbb{U}(1)$, которая состоит из комплексных чисел, равных по модулю единице, с обычным правилом умножения комплексных чисел. Как многообразие эта группа представляет собой окружность $\mathbb{U}(1)\approx\mathbb{S}^1$. Легко построить изоморфизм групп, $\mathbb{U}(1)\approx\mathbb{SO}(2)$:

$$\mathbb{U}(1)\ni\mathsf{e}^{i\alpha}\leftrightarrow\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}\in\mathbb{SO}(2),\qquad \alpha\in[0,2\pi).$$

Алгебра Ли группы $\mathfrak{u}(n)$ изоморфна алгебре антиэрмитовых $n\times n$ матриц. Вблизи единицы $\mathbb{1}\in\mathbb{SU}(n)$ произвольная унитарная матрица $U\in\mathbb{SU}(n)$ взаимно однозначно представляется в виде экспоненциала от некоторой антиэрмитовой матрицы $A\in\mathfrak{u}(n)$

$$U=\mathsf{e}^A,\qquad A=-A^\dagger.\qquad\square$$

Множество эрмитовых матриц играет важную роль. Остановимся на этом более подробно. Матрица $A\in\mathfrak{gl}(n,\mathbb{C})$ называется *эрмитовой*, если выполнено условие

$$A=A^\dagger.\tag{9.82}$$

Множество эрмитовых матриц $A,B,\ldots$ в отличии от множества антиэрмитовых матриц не образует подалгебры в $\mathfrak{gl}(n,\mathbb{C})$, т.к.

$$[A,B]^\dagger=[B,A]=-[A,B].$$

Заметим также, что отображение $A\to A^*$ не является автоморфизмом группы $\mathbb{GL}(n,\mathbb{C})$ и что эрмитовы матрицы не образуют в $\mathbb{GL}(n,\mathbb{C})$ подгруппы.

Сформулируем некоторые свойства эрмитовых матриц [?].

Предложение 9.12.5. *Если A – эрмитова матрица, то UAU^{-1} – снова эрмитова для любой унитарной матрицы $U\in\mathbb{U}(n)$. Далее, существует такая унитарная матрица U_0 (не единственная), что $U_0AU_0^{-1}$ является диагональной матрицей. Если матрица A вещественна, то матрицу U_0 можно выбрать ортогональной.*

Предложение 9.12.6. *Все собственные числа эрмитовой матрицы вещественны. Обратно ???*

Предложение 9.12.7. *Если A – эрмитова матрица, то в пространстве $\mathbb{C}^n$ существует ортонормальный базис, составленный из ее собственных векторов.*

Эрмитова матрица называется *положительно определенной*, если все ее собственные числа положительны.

Предложение 9.12.8. *Соответствие $A \to \exp A$ взаимно однозначно и гладко ??? отображает совокупность всех эрмитовых матриц на совокупность всех положительно определенных матриц. ???*

Пример 9.12.5. Группа специальных унитарных (Special Unitary) матриц $\mathbb{SU}(n) \subset \mathbb{U}(n)$, состоит из унитарных $n \times n$ матриц $U \in \mathbb{SU}(n)$ с единичным определителем

$$UU^\dagger = U^\dagger U = 1, \qquad \det U = 1.$$

Эта группа является связной компактной группой Ли размерности $\dim \mathbb{SU}(n) = n^2 - 1$. Она образует замкнутое подмножество в $\mathbb{U}(n)$. Группа $\mathbb{SU}(n)$ является также односвязной. Доказательство этого утверждения сложно, и интересующийся читатель может найти его в [**?**]. (с.239)

Группа $\mathbb{SU}(1)$ тривиальна и состоит из единственного элемента – единицы.

Алгебра Ли $\mathfrak{su}(n)$ изоморфна алгебре антиэрмитовых $n \times n$ матриц с нулевым следом. Вблизи единицы $\mathbb{1} \in \mathbb{SU}(n)$ каждая унитарная матрица с единичным определителем $U \in \mathbb{SU}(n)$ взаимно однозначно представляется экспоненциальным отображением элемента алгебры $A \in \mathfrak{su}(n)$

$$U = \exp A, \quad A = -A^\dagger, \quad \operatorname{tr} A = 0. \qquad \square$$

Предложение 9.12.9. *Любая невырожденная матрица $A \in \mathbb{GL}(n,\mathbb{C})$ может быть записана, и притом лишь единственным способом, в виде произведения $A = UP$ или $A = PU$ унитарной матрицы $U \in \mathbb{U}(n)$ и положительно определенной матрицы P. Сомножители U и P являются непрерывными функциями A.*

Поскольку множество всех положительно определенных матриц гомеоморфно $\mathbb{R}^{n^2}$, то группа $\mathbb{GL}(n,\mathbb{C})$ гомеоморфна прямому произведению $\mathbb{GL}(n,\mathbb{C}) \approx \mathbb{U}(n) \times \mathbb{R}^{n^2}$.

Пример 9.12.6. Всякая компактная связная абелева группа Ли $\mathbb{G}$ размерности n изоморфна n-мерному тору $\mathbb{T}^n$. $\square$

Пример 9.12.7. Независимое определение ???

Группа $\mathbb{SPIN}(n)$ является группой Ли размерности $n(n-1)/2$. Эта группа является связной и односвязной, и отображение $\mathbb{SPIN}(n) \to \mathbb{SO}(n)$ при $n \geq 3$ представляет собой двулистное универсальное накрытие. При малых n имеют место следующие изоморфизмы:

$$\begin{aligned}
&\mathbb{SPIN}(2) \approx \mathbb{U}(1) \approx \mathbb{SO}(2) \approx \mathbb{S}^1,\\
&\mathbb{SPIN}(3) \approx \mathbb{SU}(2) \approx \mathbb{SP}(1) \approx \mathbb{S}^3,\\
&\mathbb{SPIN}(4) \approx \mathbb{SPIN}(3) \times \mathbb{SPIN}(3) \approx \mathbb{S}^3 \times \mathbb{S}^3,\\
&\mathbb{SPIN}(5) \approx \mathbb{SP}(2),\\
&\mathbb{SPIN}(6) \approx \mathbb{SU}(4),
\end{aligned} \tag{9.83}$$

$\square$

Пример 9.12.8. Построим универсальную накрывающую группу $\widetilde{\mathbb{SL}(2,\mathbb{R})}$ для группы $\mathbb{SL}(2,\mathbb{R})$. Это построение будет конструктивным, т.е. мы определим групповую операцию на связном и односвязном многообразии $\widetilde{\mathbb{SL}(2,\mathbb{R})}$, а затем построим гомоморфизм групп $\widetilde{\mathbb{SL}(2,\mathbb{R})} \to \mathbb{SL}(2,\mathbb{R})$ (см., например, [**?**]).

Рассмотрим трехмерное многообразие $\widetilde{\mathbb{SL}(2,\mathbb{R})} = \mathbb{R}\times\mathbb{D}$, которое представляет собой прямое произведение вещественной прямой $\mathbb{R}$ на единичный диск $\mathbb{D}$ плоскости комплексного переменного, состоящего из таких комплексных чисел z, что $|z|<1$. Поскольку каждый из сомножителей является связным и односвязным многообразием, то и их произведение также связно и односвязно.

Функция

$$w = \frac{1+z}{1+\bar z}$$

осуществляет непрерывное отображение диска $\mathbb{D}$ на окружность $|w|=1$ с выколотой точкой $w=-1$. Поэтому для любого числа $z\in\mathbb{D}$ существует единственное число t, $-\pi<t<\pi$, для которого

$$\mathsf{e}^{it} = \frac{1+z}{1+\bar z}.$$

Обозначим это число символом

$$t = \frac{1}{i}\,\mathsf{ln}\frac{1+z}{1+\bar z}.$$

Определим на многообразии $\widetilde{\mathbb{SL}(2,\mathbb{R})}$ умножение следующей формулой

$$(x,u)(y,v) = \left(x+y+t, \frac{u+\mathsf{e}^{2iy}v}{\mathsf{e}^{2iy}+u\bar v}\right), \qquad x,y\in\mathbb{R},\quad u,v\in\mathbb{D}, \tag{9.84}$$

где

$$t = \frac{1}{2i}\,\mathsf{ln}\frac{1+\mathsf{e}^{-2iy}u\bar v}{1+\mathsf{e}^{2iy}\bar u v}.$$

Так как

$$|\mathsf{e}^{2iy}+u\bar v|^2 - |u+\mathsf{e}^{2iy}v|^2 = \left(1-|u|^2\right)\left(1-|v|^2\right) > 0,$$

то

$$\left|\frac{u+\mathsf{e}^{2iy}v}{\mathsf{e}^{2iy}+u\bar v}\right| < 1.$$

Поэтому формула (9.84) корректно определяет в $\widetilde{\mathbb{SL}(2,\mathbb{R})} = \mathbb{R}\times\mathbb{D}$ умножение.

Это умножение обладает единицей $(0,0)$, и для любого элемента (x,u) существует обратный элемент

$$(x,u)^{-1} = (-x, -\mathsf{e}^{2ix}u).$$

Кроме того, прямые вычисления показывают, что умножение (9.84) ассоциативно. Это умножение гладко, и тем самым доказано, что многообразие $\widetilde{\mathbb{SL}(2,\mathbb{R})} = \mathbb{R}\times\mathbb{D}$ является группой Ли относительно умножения (9.84).

Центр этой группы состоит из таких элементов (x,u), для которых, в частности, равенство

$$\frac{u+\mathsf{e}^{2iy}v}{\mathsf{e}^{2iy}+u\bar v} = \frac{v+\mathsf{e}^{2ix}u}{\mathsf{e}^{2ix}+v\bar u} \tag{9.85}$$

выполнено для всех $y\in\mathbb{R}$ и $v\in\mathbb{D}$. В частном случае при $v=0$ должно выполняться равенство $u = \mathsf{e}^{2iy}u$ для всех y. Это возможно только при $u=0$. Тогда из

(9.85) следует равенство $v = \mathsf{e}^{2ix}v$, которое должно выполняться для всех v, что возможно только при $x = \pi n$, $n \in \mathbb{Z}$. Нетрудно проверить, что все элементы вида $(\pi n, 0)$ принадлежат центру группы $\widetilde{\mathbb{SL}(2,\mathbb{R})}$. Следовательно, центр группы $\widetilde{\mathbb{SL}(2,\mathbb{R})}$ бесконечен.

Теперь построим гомоморфизм $\widetilde{\mathbb{SL}(2,\mathbb{R})} \to \mathbb{SL}(2,\mathbb{R})$. Легко проверить, что матрица

$$A = \frac{1}{\sqrt{1-|u|^2}} \begin{pmatrix} \cos x + |u| \cos(x+\varphi) & |u| \sin(x+\varphi) - \sin x \\ |u| \sin(x+\varphi) + \sin x & \cos x - |u| \cos(x+\varphi) \end{pmatrix}, \qquad (9.86)$$

где $\varphi = \arg u$, имеет единичный определитель. Прямые вычисления позволяют проверить, что отображение $\widetilde{\mathbb{SL}(2,\mathbb{R})} \to \mathbb{SL}(2,\mathbb{R})$ определено корректно и является гомоморфизмом групп. Его ядро состоит из элементов вида $(2\pi n, 0)$ и, значит, дискретно. Поскольку размерности групп совпадают, то гомоморфизм $\widetilde{\mathbb{SL}(2,\mathbb{R})} \to \mathbb{SL}(2,\mathbb{R})$ является групповым накрытием. Таким образом мы построили универсальную накрывающую для группы $\mathbb{SL}(2,\mathbb{R})$:

$$\mathbb{SL}(2,\mathbb{R}) = \frac{\widetilde{\mathbb{SL}(2,\mathbb{R})}}{\mathbb{Z}},$$

где группа преобразований

$$\mathbb{Z}: \quad x \to x + 2\pi n$$

действует на прямой $\mathbb{R}$. С топологической точки зрения группа $\mathbb{SL}(2,\mathbb{R})$ является произведением окружности на диск

$$\mathbb{SL}(2,\mathbb{R}) = \mathbb{S}^1 \times \mathbb{D}.$$

Группа специальных матриц $\mathbb{SL}(2,\mathbb{R})$ дважды накрывает собственную ортохронную группу Лоренца $\mathbb{SO}_\uparrow(1,2)$ (2.173). При этом одному лоренцеву вращению соответствует пара матриц (9.86), отличающихся знаком. Изменению знака матрицы A соответствует сдвиг координаты x на π. Поэтому

$$\mathbb{SO}_\uparrow(1,2) = \frac{\mathbb{SL}(2,\mathbb{R})}{\mathbb{Z}_2} = \frac{\widetilde{\mathbb{SL}(2,\mathbb{R})}}{\mathbb{Z}},$$

где группа преобразований

$$\mathbb{Z}: \quad x \to x + \pi n.$$

Группа $\mathbb{Z}_2$ действует на окружности $\mathbb{S}^1$ путем отождествления противоположных точек и превращает окружность в проективное пространство $\mathbb{RP}^1$. Поскольку одномерное проективное пространство с топологической точки зрения является окружностью, то

$$\mathbb{SO}_\uparrow(1,2) = \mathbb{S}^1 \times \mathbb{D}.$$

Поскольку диск диффеоморфен всей плоскости $\mathbb{D} \approx \mathbb{R}^2$, то

$$\mathbb{SO}_\uparrow(1,2) \approx \mathbb{S}^1 \times \mathbb{R}^2.$$

Предложенная конструкция универсальной накрывающей имеет следующее объяснение. Согласно теореме о полярном разложении матриц любая унимодулярная матрица $A \in \mathbb{SL}(2,\mathbb{R})$ единственным образом раскладывается в произведение $A =$

UP некоторой матрицы вращения U и некоторой положительно определенной унимодулярной матрицы P. Матрица U задается углом поворота x, а матрица P, имеющая вид

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

где $a > 0$ и $ac - b^2 = 1$, задается двумя числами $a > 0$ и $b \in \mathbb{R}$, т.е. комплексным числом $z = a + ib$, принадлежащим правой полуплоскости $a > 0$. Переходя с помощью дробно-линейного преобразования от правой полуплоскости к единичному кругу, мы получаем, следовательно, что каждая матрица $A \in \mathbb{SL}(2, \mathbb{R})$ однозначно характеризуется парой (x, u), где $x \in \mathbb{R}$ и $u \in \mathbb{D}$. При этом умножению матриц соответствует умножение пар, которое задается формулой (9.84). Чтобы получить теперь универсальную накрывающую $\widetilde{\mathbb{SL}(2, \mathbb{R})}$ достаточно считать координату x не углом, а точкой вещественной прямой $\mathbb{R}$. $\square$

Глава 10

Группы преобразований

Теория групп имеет важные приложения в математической физике, и почти всегда группы появляются как группы преобразований чего либо. Это может быть группа преобразований пространства-времени, группа симметрий изотопического пространства (например, $\mathbb{SU}(3)$ симметрия в физике элементарных частиц), группа симметрий уравнений движения или что нибудь более экзотическое. В настоящем разделе мы рассмотрим группы преобразований $\mathbb{G}$ многообразия $\mathbb{M}$. В общем случае группа $\mathbb{G}$ может быть конечной, счетной или группой Ли. В основном мы будем рассматривать группы Ли преобразований, хотя многие определения и утверждения справедливы и для произвольных групп.

10.1 Действие групп преобразований

Пусть на многообразии $\mathbb{M}$ задано *правое* действие группы $\mathbb{G}$, т.е. задано отображение

$$\mathbb{M} \times \mathbb{G} \ni (x, a) \mapsto xa \in \mathbb{M}, \tag{10.1}$$

которое предполагается достаточно гладким.

Определение. Пара $(\mathbb{M}, \mathbb{G})$ называется *группой правых преобразований* многообразия $\mathbb{M}$, если действие группы $\mathbb{G}$ на $\mathbb{M}$ справа удовлетворяет следующим условиям:

1) при любом $a \in \mathbb{G}$ отображение $\mathbb{M} \ni x \mapsto xa \in \mathbb{M}$ является диффеоморфизмом,

2) $(xa)b = x(ab)$ для всех $x \in \mathbb{M}$ и $a, b \in \mathbb{G}$.

В этом случае многообразие $\mathbb{M}$ называется $\mathbb{G}$-*многообразием.* □

Замечание. Если $\mathbb{G}$ – группа Ли, то отображение (10.1) предполагается дифференцируемым и, следовательно, непрерывным для всех $a \in \mathbb{G}$ и для всех $x \in \mathbb{M}$. Для каждой фиксированной точки $x \in \mathbb{M}$ мы имеем дифференцируемое отображение $\mathbb{G} \to \mathbb{M}$. □

Замечание. Иногда в качестве $\mathbb{M}$ мы будем рассматривать многообразия с краем. Под дифференцируемой функцией f на $\mathbb{M}$ в таком случае мы понимаем наличие всех частных производных у f во всех внутренних точках и существование предела этих производных на крае $\partial\mathbb{M}$. □

Каждому элементу группы $a \in \mathbb{G}$ соответствует дифференцируемое отображение $\mathbb{M} \to \mathbb{M}$, заданное правилом $x \mapsto xa$. Поэтому группа преобразований определяет отображение $\mathbb{G} \to \mathsf{diff}\,(\mathbb{M})$ группы $\mathbb{G}$ в группу диффеоморфизмов многообразия $\mathsf{diff}\,(\mathbb{M})$.

Предложение 10.1.1. *Пара* $(\mathbb{M}, \mathbb{G})$ *является группой преобразований многообразия* $\mathbb{M}$ *тогда и только тогда, когда отображение* $\mathbb{G} \to \mathsf{diff}\,(\mathbb{M})$ *является гомоморфизмом групп и отображение (10.1) дифференцируемо.*

Доказательство. Условие 1) в определении группы преобразований означает, что отображение $x \mapsto xa$ принадлежит группе $\mathsf{diff}\,(\mathbb{M})$. Групповая операция сохраняется при отображении в силу свойства 2). Осталось доказать, что единица $e \in \mathbb{G}$ отображается в тождественное преобразование многообразия $\mathbb{M}$. Действительно, отображение $x \mapsto xa$ является диффеоморфизмом и, следовательно, биекцией для всех $a \in \mathbb{G}$. В частности, для каждого $x \in \mathbb{M}$ найдется такая точка $y \in \mathbb{M}$, что $x = ye$. Умножив это уравнение справа на e, получаем равенство $xe = x$ для всех $x \in \mathbb{M}$. То есть единичный элемент группы не сдвигает ни одну точку многообразия.

Обратно. Каждому элементу a соответствует некоторый диффеоморфизм, т.е. условие 1) выполнено. Поскольку отображение $\mathbb{G} \to \mathsf{diff}\,(\mathbb{M})$ – гомоморфизм, то выполнено условие 2). □

Аналогично определяется *левое* действие группы $\mathbb{G}$ на $\mathbb{M}$, т.е. отображение $\mathbb{G} \times \mathbb{M} \ni (a, x) \mapsto ax \in \mathbb{M}$.

В дальнейшем, если не оговорено противное, под *группой преобразований* многообразия будем понимать группу правых преобразований и будем говорить, что $\mathbb{G}$ действует на $\mathbb{M}$, подразумевая действие группы справа.

Пример 10.1.1. Если задано представление группы $\mathbb{G}$, то задана группа линейных преобразований $(\mathbb{V}, \mathbb{G})$ векторного пространства $\mathbb{V}$, которое снабжено естественной структурой многообразия. □

Пример 10.1.2. Группа Ли является группой преобразований самой себя $\mathbb{G}\times\mathbb{G} \to \mathbb{G}$, которую можно рассматривать, как действующую слева или справа. □

Определение. Группа $\mathbb{G}$ действует на $\mathbb{M}$ *тривиально*, если равенство $xa = x$ выполнено для всех $x \in \mathbb{M}$ и всех $a \in \mathbb{G}$. В этом случае группа преобразований превращается в тривиальное главное расслоение $\mathbb{M} \times \mathbb{G} \xrightarrow{\pi} \mathbb{M}$ с естественной проекцией $\pi:\ (x, a) \mapsto x$.

Говорят, что $\mathbb{G}$ действует *эффективно* на $\mathbb{M}$, если из равенства $xa = x$ для всех $x \in \mathbb{M}$ следует, что $a = e$. Если для любой точки $x \in \mathbb{M}$ уравнение $xa = x$ имеет единственное решение $a = e$, то говорят, что группа преобразований действует *свободно*. □

Другими словами, действие группы $\mathbb{G}$ на $\mathbb{M}$ является эффективным, если любой элемент группы, отличный от единичного, перемещает хотя бы одну точку. Действие группы свободно, если любой элемент группы, отличный от единичного, перемещает все точки многообразия. Группа, действующая свободно, является эффективной. Обратное утверждение неверно, что показывает следующий

Пример 10.1.3. Группа вращений $\mathbb{SO}(3)$ трехмерного евклидова пространства $\mathbb{R}^3$ действует эффективно, но не свободно, т.к. начало координат остается неподвижным. □

Если действие группы преобразований свободно, то это значит, что для любой пары точек $x_1, x_2 \in \mathbb{M}$ либо не существует преобразования, связывающего эти точки, $x_2 = x_1 a$, либо элемент $a \in \mathbb{G}$ *единственен.*

Определение. *Ядром* $\mathbb{K}$ *группы преобразований* $(\mathbb{M}, \mathbb{G})$ называется множество элементов группы, которое действует тривиально на все точки многообразия,

$$\mathbb{K} := \{a \in \mathbb{G} : \ xa = x, \ \forall x \in \mathbb{M}\}. \qquad \square$$

Действие группы эффективно, если $\mathbb{K} = \{e\}$.

Предложение 10.1.2. *Ядро группы преобразований является нормальной подгруппой в* $\mathbb{G}$. *Кроме того, эта подгруппа замкнута в* $\mathbb{G}$.

Доказательство. Отображение $\mathbb{G} \to \mathsf{diff}\,(\mathbb{M})$ непрерывно и ядро $\mathbb{K}$ отображается в единицу группы диффеоморфизмов $\mathsf{diff}\,(\mathbb{M})$. Поскольку единица замкнута в $\mathsf{diff}\,(\mathbb{M})$, то ее прообраз $\mathbb{K}$ замкнут в $\mathbb{G}$ по критерию 1) теоремы 2.4.3. $\square$

Ядро группы преобразований представляет собой то множество элементов группы, которое вообще не действует на многообразии $\mathbb{M}$. Поэтому действие группы $\mathbb{G}$ естественным образом сводится к действию фактор группы $\mathbb{G}/\mathbb{K}$, которая действует эффективно.

Пример 10.1.4. Рассмотрим присоединенное действие группы Ли на себя $\mathbb{G} \times \mathbb{G} \to \mathbb{G}$, заданное преобразованием подобия $a \mapsto bab^{-1}$, для всех $a, b \in \mathbb{G}$. Ядром этого действия является центр группы $\mathbb{K} \subset \mathbb{G}$. Присоединенное действие группы Ли не является свободным, так как единичный элемент группы $e \in \mathbb{G}$ остается неподвижным, $aea^{-1} = e$, $\forall a \in \mathbb{G}$. Присоединенное действие фактор группы $\mathbb{G}/\mathbb{K}$ является эффективным. $\square$

Определение. Пусть точка $x \in \mathbb{M}$ фиксирована. Множество точек

$$x\mathbb{G} := \{xa \in \mathbb{M} : \ \forall a \in \mathbb{G}\}$$

называется *орбитой* точки x относительно действия группы $\mathbb{G}$. $\square$

Используя групповые свойства, нетрудно проверить, что орбиты двух точек $x_1, x_2 \in \mathbb{M}$ либо совпадают, либо не имеют общих точек. Орбитой произвольного подмножества $\mathbb{U} \subset \mathbb{M}$ будем называть объединение орбит всех точек из $\mathbb{U}$. Обозначим его $\mathbb{U}\mathbb{G}$.

Если группа $\mathbb{G}$ конечна, то орбита любой точки также конечна. Если группа преобразований счетна, то ее орбиты либо счетны, либо конечны.

Пример 10.1.5. Рассмотрим поворот евклидовой плоскости $\mathbb{R}^2$ на конечный угол θ. Начало координат неподвижно относительно вращений. Все остальные точки имеют нетривиальные орбиты, если угол поворота не кратен 2π. Пусть $\theta = 2\pi k$, где $k \in \mathbb{R}$. Если число k рационально, $k = m/n$, где m и n – взаимно простые натуральные числа, то орбиты состоят из n точек. Если k – иррациональное число, то орбита произвольной точки $x \neq 0$ счетна и являются всюду плотным подмножеством на окружности, проходящей через точку x. $\square$

Если задана группа Ли преобразований $(\mathbb{M}, \mathbb{G})$, то орбита $x\mathbb{G}$ произвольной точки $x \in \mathbb{M}$ является подмногообразием в $\mathbb{M}$. Это следует из дифференцируемости отображения (10.1) для всех $x \in \mathbb{M}$ и групповых свойств $\mathbb{G}$. Размерность орбиты не превосходит размерности группы, $\mathsf{dim}\,(x\mathbb{G}) \leq \mathsf{dim}\,\mathbb{G}$.

Пример 10.1.6. Рассмотри евклидово пространство $\mathbb{R}^n$ с декартовой системой координат x^α, $\alpha = 1, \ldots, n$. Пусть группа вещественных чисел $\mathbb{R}$ по сложению действует на $\mathbb{R}^n$, как трансляции вдоль первой оси,

$$\mathbb{R}^n \times \mathbb{R} \ni \quad \Big(\{x^1, x^2, \ldots, x^n\}, a\Big) \mapsto \{x^1 + a, x^2, \ldots, x^n\} \quad \in \mathbb{R}^n.$$

Тогда орбитой произвольной точки $x \in \mathbb{R}^n$ является прямая, проходящая через x и параллельная оси x^1. □

Предложение 10.1.3. *Орбита $x\mathbb{G}$ произвольной точки $x \in \mathbb{M}$ является замкнутым подмногообразием в $\mathbb{M}$. Кроме того, если $\mathbb{G}$ – связная группа Ли, то орбита является связным подмногообразием.*

Доказательство. Подмножество $\mathbb{M}\backslash x$ открыто в $\mathbb{M}$. Поэтому объединение открытых подмножеств $(\mathbb{M} \setminus x)\mathbb{G}$ открыто в $\mathbb{M}$. Следовательно, орбита произвольной точки замкнута, как дополнение открытого подмножества, $x\mathbb{G} = \mathbb{M}\setminus\Big((\mathbb{M}\setminus x)\mathbb{G}\Big)$. Поскольку отображение (10.1) является непрерывным для всех x, то связность сохраняется. □

Определение. Обозначим через $\mathbb{M}/\mathbb{G}$ множество орбит $\breve{x} := x\mathbb{G}$ относительно действия группы $\mathbb{G}$ на $\mathbb{M}$. То есть $\breve{x}_1 = \breve{x}_2$ тогда и только тогда, когда точки x_1 и x_2 лежат на одной орбите: $x_2 = x_1 a$ для некоторого $a \in \mathbb{G}$. Пусть $\pi : \mathbb{M} \to \mathbb{M}/\mathbb{G}$ – проекция, сопоставляющая каждой точке $x \in \mathbb{M}$ ее орбиту $\breve{x}$. Тогда пространство орбит наделяется фактортопологией, т.е. множество $\mathbb{V} \in \mathbb{M}/\mathbb{G}$ является открытым тогда и только тогда, когда множество $\pi^{-1}(\mathbb{V})$ открыто в $\mathbb{M}$. Если $\mathbb{U} \subset \mathbb{M}$ – окрестность точки $x \in \mathbb{M}$, то множество орбит $\mathbb{V} = \mathbb{U}\mathbb{G}$ – окрестность точки $\breve{x} \in \mathbb{M}/\mathbb{G}$. Множество орбит $\mathbb{M}/\mathbb{G}$ называется *пространством орбит* многообразия $\mathbb{M}$ относительно действия группы $\mathbb{G}$ или *факторпространством*. □

Замечание. Хотя $\mathbb{M}$ и $\mathbb{G}$ – многообразия, пространство орбит с фактортопологией может оказаться весьма сложным. Оно, как покажут дальнейшие примеры, может не иметь структуры многообразия. В общем случае пространство орбит может не быть даже хаусдорфовым топологическим пространством. □

Определение. Действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ называется *транзитивным*, если для любых двух точек $x_1, x_2 \in \mathbb{M}$ найдется такой элемент $a \in \mathbb{G}$, что $x_1 a = x_2$. □

Если действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ транзитивно, то орбита произвольной точки $x \in \mathbb{M}$ совпадает со всем $\mathbb{M}$. Если на многообразии существует одна точка $x_0 \in \mathbb{M}$, которую можно преобразовать во все другие точки $\mathbb{M}$, то этого достаточно для того, чтобы действие группы $\mathbb{G}$ на $\mathbb{M}$ было транзитивным. Действительно, если $x_1 = x_0 a$ и $x_2 = x_0 b$, то $x_2 = x_1 a^{-1} b$.

Пример 10.1.7. Действие группы Ли $\mathbb{G}$ на себе является транзитивным и свободным. □

Пример 10.1.8. Пусть $\mathbb{G}$ – группа и $\mathbb{H}$ – ее нетривиальная подгруппа. Рассмотрим пространство правых смежных классов $\mathbb{G}/\mathbb{H}$, состоящее из элементов вида $\mathbb{H}a$, где $a \in \mathbb{G}$. Определим действие группы $\mathbb{G}$ в пространстве смежных классов $\mathbb{G}/\mathbb{H}$ соотношением

$$\mathbb{G}/\mathbb{H} \times \mathbb{G} \ni \quad \mathbb{H}a \times b \mapsto \mathbb{H}ab \quad \in \mathbb{G}/\mathbb{H}.$$

Это действие транзитивно для любых $\mathbb{G}$ и $\mathbb{H}$. Действие $\mathbb{G}$ на $\mathbb{G}/\mathbb{H}$ эффективно тогда и только тогда, когда $\mathbb{H}$ не содержит нормальных подгрупп. Действительно, если $\mathbb{H}$ – нормальная подгруппа, то для всех $h \in \mathbb{H}$ и $a \in \mathbb{G}$ выполнено равенство $(\mathbb{H}a)h = \mathbb{H}a$. Действие группы никогда не является свободным, т.к. $(\mathbb{H}e)h = \mathbb{H}e$ для всех $h \in \mathbb{H}$. □

В теории групп преобразований важную роль играет

Теорема 10.1.1. *Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H}$ – ее замкнутая подгруппа. Тогда на фактор пространстве правых смежных классов $\mathbb{G}/\mathbb{H}$ существует единственная структура вещественно аналитического многообразия такая, что действие $\mathbb{G}$ на $\mathbb{G}/\mathbb{H}$ вещественно аналитично. В частности, проекция $\mathbb{G} \to \mathbb{G}/\mathbb{H}$ вещественно аналитична.*

Доказательство. См., например, в [**?**]. □

Определение. Точка $x \in \mathbb{M}$ называется *стационарной* или *неподвижной* относительно действия группы преобразований, если $xa = x$ для всех $a \in \mathbb{G}$. □

Если группа действует на многообразии свободно или транзитивно, то стационарные точки отсутствуют.

Пример 10.1.9. Представление группы $(\mathbb{V}, \mathbb{G})$ не может быть транзитивной или свободной группой преобразований, т.к. нулевой вектор из $\mathbb{V}$ является стационарной точкой. □

Определение. *Группой изотропии* $\mathbb{G}_x$ точки многообразия $x \in \mathbb{M}$ называется множество элементов группы $\mathbb{G}$, оставляющих точку x неподвижной,

$$\mathbb{G}_x := \{a \in \mathbb{G} : \ xa = x\}\,. \qquad \square$$

Нетрудно проверить, что это множество элементов образует группу. Группы изотропии $\mathbb{G}_{x_1}$ и $\mathbb{G}_{x_2}$ двух точек, лежащих на одной орбите, $x_2 a = x_1$, являются сопряженными подгруппами в $\mathbb{G}$:

$$\mathbb{G}_{x_2} = a^{-1}\mathbb{G}_{x_1} a.$$

Ядро $\mathbb{K}$ группы преобразований $(\mathbb{M}, \mathbb{G})$ связано с группами изотропии простым соотношением

$$\mathbb{K} = \bigcap_{x \in \mathbb{M}} \mathbb{G}_x,$$

где объединение берется по всем точкам $x \in \mathbb{M}$.

Из теоремы 10.1.1 следует, что пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ является группой транзитивных преобразований. Представляет большой интерес обратное утверждение: если пара $(\mathbb{M}, \mathbb{G})$ – транзитивная группа преобразований, то она имеет вид $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ для некоторой подгруппы $\mathbb{H} \subset \mathbb{G}$ и при некоторых дополнительных предположениях, которые мы уточним ниже. Для ясности сначала будет доказана теорема безотносительно топологии и дифференцируемой структуры, заданной на $\mathbb{M}$.

Теорема 10.1.2. *Пусть $\mathbb{G}$ – группа, которая действует транзитивно на некотором множестве $\mathbb{M}$. Тогда для каждой точки $x \in \mathbb{M}$ существует биекция правых смежных классов $\mathbb{G}/\mathbb{G}_x$ на множество $\mathbb{M}$, заданное отображением*

$$\mathbb{G}/\mathbb{G}_x \ni \mathbb{G}_x a \mapsto xa \in \mathbb{M}, \tag{10.2}$$

где $\mathbb{G}_x$ – группа изотропии точки x.

Доказательство. Отображение (10.2) определено для каждого представителя из правого смежного класса. Покажем сначала, что это отображение не зависит от выбора представителя. Пусть a и b – два представителя какого либо одного смежного класса. Тогда существует элемент $h \in \mathbb{G}_x$ такой, что $a = hb$ и поэтому $xa = xhb = xb$. То есть отображение (10.2) не зависит от выбора представителя смежного класса и, следовательно, определено корректно.

Выберем по представителю из двух смежных классов $\mathbb{G}_x a$ и $\mathbb{G}_x b$ и подействуем на точку x. В результате получим точки $x_1 = xa$ и $x_2 = xb$. Эти точки совпадают тогда и только тогда, когда группа изотропии содержит элемент $ba^{-1} \in \mathbb{G}_x$. Это означает, что отображение (10.2) является взаимно однозначным.

Поскольку группа преобразований $(\mathbb{M}, \mathbb{G})$ является транзитивной, то отображение (10.2) является взаимно однозначным отображением на, т.е. биекцией. □

Следствие. Если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна и свободна, т.е. $\mathbb{G}_x = e$ для всех точек $x \in \mathbb{M}$, то отображение

$$\mathbb{G} \ni a \mapsto xa \in \mathbb{M}$$

является биекцией. □

Теперь разберемся с дифференцируемой структурой.

Предложение 10.1.4. *Пусть $(\mathbb{M}, \mathbb{G})$ – группа транзитивных преобразований. Выберем произвольную точку $x \in \mathbb{M}$. Тогда ее группа изотропии $\mathbb{G}_x$ является замкнутой подгруппой в $\mathbb{G}$.*

Доказательство. Поскольку группа преобразований транзитивна, то отображение $\mathbb{G} \ni a \mapsto xa \in \mathbb{M}$, определенное для каждого $x \in \mathbb{M}$, является сюръективным и непрерывным. Точка x является замкнутым подмножеством в $\mathbb{M}$, и, в силу теоремы критерия 1) из теоремы 2.4.3, группа изотропии $\mathbb{G}_x$ замкнута в $\mathbb{G}$. □

Следствие. Группа изотропии $\mathbb{G}_x$ для всех $x \in \mathbb{M}$ является замкнутой подгруппой в $\mathbb{G}$ для произвольной группы Ли преобразований $(\mathbb{M}, \mathbb{G})$.

Доказательство. Каждая точка $x \in \mathbb{M}$ принадлежит орбите $x\mathbb{G}$, которая является подмногообразием в $\mathbb{M}$. На каждой орбите группа $\mathbb{G}$ действует транзитивно, и поэтому пара $(x\mathbb{G}, \mathbb{G})$ – транзитивная группа преобразований. Следовательно, группа изотропии $\mathbb{G}_x$ является замкнутой подгруппой в $\mathbb{G}$. □

Согласно теореме 10.1.1 на пространстве смежных классов $\mathbb{G}/\mathbb{G}_x$ существует единственная дифференцируемая структура. Поскольку отображение (10.2) является биекцией, то эту дифференцируемую структуру можно перенести на $\mathbb{M}$, превратив его тем самым в многообразие. Тогда отображение (10.2) становится диффеоморфизмом.

Таким образом, если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна, то многообразие $\mathbb{M}$ представляет собой фактор пространство $\mathbb{G}/\mathbb{G}_x$. При этом

$$\mathsf{dim}\,\mathbb{M} = \mathsf{dim}\,\mathbb{G} - \mathsf{dim}\,\mathbb{G}_x.$$

Отсюда следует, что размерность многообразия $\mathbb{M}$, на котором действует транзитивная группа преобразований, не может превышать размерность группы.

Определение. Произвольное множество $\mathbb{M}$, на котором группа преобразований $(\mathbb{M}, \mathbb{G})$ действует транзитивно называется *однородным пространством*. Если, вдобавок, действие группы эффективно, то это множество называется *главным однородным пространством* над $\mathbb{G}$. □

Пример 10.1.10. Пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ (пример 10.1.8) является однородным пространством. Если подгруппа $\mathbb{H} \subset \mathbb{G}$ не содержит нормальных подгрупп, то пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$ – главное однородное пространство. □

10.2 Инфинитезимальные преобразования

Пусть задана группа преобразований $(\mathbb{M}, \mathbb{G})$, где $\mathbb{G}$ – группа Ли. Тогда имеет смысл говорить об инфинитезимальных преобразованиях, которые взаимно однозначно определяются векторными полями на $\mathbb{M}$ и $\mathbb{G}$. Ниже мы рассмотрим соотношения между алгебрами Ли векторных полей на $\mathbb{G}$ и $\mathbb{M}$.

Построим отображение $\mu: \mathfrak{g} \to \mathcal{X}(\mathbb{M})$ алгебры Ли $\mathfrak{g}$ группы Ли $\mathbb{G}$ (множества левоинвариантных векторных полей в $\mathcal{X}(\mathbb{G})$) в бесконечномерную алгебру Ли векторных полей $\mathcal{X}(\mathbb{M})$ на многообразии $\mathbb{M}$. Это можно сделать тремя эквивалентными способами.

1) Каждому элементу алгебры Ли $X \in \mathfrak{g}$ соответствует однопараметрическая подгруппа (см. раздел 9.9)

$$\mathsf{exp}_X: \quad \mathbb{R} \ni t \mapsto a(t) \in \mathbb{G},$$

где $a(t) := \mathsf{exp}_X(t)$ (экспоненциальное отображение) и $a(0) = e$ – единица группы. Тогда через каждую точку $x \in \mathbb{M}$ проходит единственная кривая $xa(t)$. Касательные векторы к этим кривым образуют векторное поле $X^*(x)$ на многообразии $\mathbb{M}$. Таким образом мы построили отображение μ алгебры Ли $\mathfrak{g}$ в алгебру векторных полей $\mathcal{X}(\mathbb{M})$:

$$\mu: \quad \mathfrak{g} \ni X \mapsto \mu(X) := X^* \in \mathcal{X}(\mathbb{M}).$$

2) Для каждой точки $x \in \mathbb{M}$ определим отображение $\mu_x: G \ni a \mapsto xa \in \mathbb{M}$. Тогда отображение μ определено соотношением

$$X^*(x) = \mu(X)|_x := (\mu_x)_* X_e, \tag{10.3}$$

где $(\mu_x)_*$ – дифференциал отображения μ_x, действующий на касательный вектор $X_e := X(e)$ к единице группы, который соответствует элементу алгебры Ли $X \in \mathfrak{g}$.

3) Каждое векторное поле взаимно однозначно определяется дифференцированием в алгебре функций. Поэтому рассмотрим гладкую функцию $f \in \mathcal{C}^\infty(\mathbb{M})$ и определим векторное поле соотношением

$$X^*(f) := \frac{d}{dt} f\Big(xa(t)\Big)\Big|_{t=0},$$

где $a(t) := \mathsf{exp}_X(t)$.

Поскольку произвольное левоинвариантное векторное поле X на группе Ли $\mathbb{G}$ полно, то полно также векторное поле X^* на $\mathbb{M}$. Это следует из определения, т.к. отображение μ определено для всех $t \in (-\infty, \infty)$.

Предложение 10.2.1. *Отображение μ есть гомоморфизм алгебр Ли. Если $\mathbb{G}$ действует эффективно на $\mathbb{M}$, то μ есть изоморфизм алгебры Ли $\mathfrak{g}$ и ее образа $\mu(\mathfrak{g}) \subset \mathcal{X}(\mathbb{M})$. Если $\mathbb{G}$ действует свободно на $\mathbb{M}$, то для каждого ненулевого элемента алгебры Ли $X \in \mathfrak{g}$, его образ $\mu(X) = X^*$ не обращается в нуль на $\mathbb{M}$.*

Доказательство. Поскольку дифференциал отображения – это линейная операция, то из определения (10.3) следует, что μ есть линейное отображение из $\mathfrak{g}$ в $\mathcal{X}(\mathbb{M})$. Доказательство того факта, что отображение μ сохраняет коммутатор векторных полей

$$[X^*, Y^*] = \mu[X, Y], \qquad \text{где } \; X^*, Y^* \in \mathcal{X}(\mathbb{M}),\; X, Y \in \mathfrak{g},$$

технично и приведено, например, в [**?**].

Докажем второе утверждение предложения. Пусть $\mu(X) = 0$ всюду на $\mathbb{M}$. Это значит, что однопараметрическая группа действует тривиально на $\mathbb{M}$, т.е. $a(t) = e$ для всех t. Верно и обратное утверждение. Пусть теперь $X \in \ker \mu$, т.е. $\mu(X) = 0$. Это значит, что $a(t) = e$ для всех t. Отсюда следует, что ядро $\ker \mu$ тривиально, $\mu(X) = 0$. Поэтому отображение μ – изоморфизм.

Пусть $\mu(X) = 0$ в некоторой точке $x \in \mathbb{M}$. Тогда точка x неподвижна, $xa(t) = x$ для всех t. Обратно. Если точка x неподвижна, то $a(t) = e$ для всех t и, следовательно, $X = 0$. □

Это предложение является основой для рассмотрения главных расслоений со структурной группой Ли в следующем разделе.

Мы видим, что транзитивные группы преобразований устроены довольно просто. В этом случае многообразие $\mathbb{M}$ – это одна орбита, и фактор пространство $\mathbb{M}/\mathbb{G}$ состоит из одной точки. Если действие группы $\mathbb{G}$ на многообразии $\mathbb{M}$ не является транзитивным, то орбит может быть много. В общем случае пространство орбит устроено сложно, что показывает следующий

Пример 10.2.1. Стационарной точкой в евклидовом пространстве $\mathbb{R}^3$ относительно действия группы трехмерных вращений $\mathbb{SO}(3)$ является начало координат. Группа изотропии начала координат совпадает со всей группой, а группой изотропии любой другой точки является подгруппа двумерных вращений $\mathbb{SO}(2) \subset \mathbb{SO}(3)$ вокруг оси, проходящей через данную точку и начало координат. Пространство орбит представляет собой множество сфер $\mathbb{S}^2_r$, которое можно параметризовать их радиусом $r \in \mathbb{R}_+$, объединенное с началом координат $\{0\} \in \mathbb{R}^3$,

$$\mathbb{R}^3 \Big/ \mathbb{SO}(3) \approx \mathbb{R}_+ \cup \{0\}.$$

То есть факторпространство представляет собой положительную полуось $\overline{\mathbb{R}}_+$, включая начало координат, и не является многообразием, т.к. не является открытым множеством (оно является многообразием с краем). □

Если группа преобразований $(\mathbb{M}, \mathbb{G})$ не является транзитивной, то многообразие $\mathbb{M}$ представляет собой объединение орбит, на каждой из которых группа $\mathbb{G}$ действует транзитивно. Выберем точку x_i на каждой орбите и обозначим через $\mathbb{G}_i$ ее группу изотропии. Тогда существует диффеоморфизм каждой орбиты на фактор пространство $\mathbb{G}/\mathbb{G}_i$. Другими словами, каждая орбита является однородным пространством и на нем существует естественная дифференцируемая структура (теорема 10.1.1). Согласно предложению 10.1.3 каждая орбита является замкнутым подмногообразием в $\mathbb{M}$.

Рассмотрим частный случай. Пусть L_α, $\alpha = 1, \ldots, \dim \mathbb{G}$, – базис алгебры Ли. Если векторные поля $\mu(L_\alpha)$ линейно независимы в каждой точке, то они определяют инволютивное распределение на многообразии $\mathbb{M}$. Тогда, согласно теореме Фробениуса, каждая орбита с естественной дифференцируемой структурой является подмногообразием в $\mathbb{M}$. В общем случае, несмотря на изоморфизм отображения μ, векторные

поля $\mu(L_{\text{A}})$ могут быть линейно зависимы. Это показывает пример 10.2.1, в котором орбиты двумерны (за исключением 0-мерного начала координат), а алгебра Ли $\mathfrak{so}(3)$ – трехмерна. Вопрос о том можно ли ввести на пространстве орбит (фактор пространстве $\mathbb{M}/\mathbb{G}$) дифференцируемую структуру сложен и выходит за рамки данной книги.

В заключение настоящего раздела докажем утверждение, которое будет использовано при построении связностей в главном расслоении.

Предложение 10.2.2. *Пусть a есть диффеоморфизм многообразия $\mathbb{M}$,*

$$a: \quad \mathbb{M} \ni x \mapsto a(x) \in \mathbb{M}.$$

Если векторное поле $X \in \mathcal{X}(\mathbb{M})$ порождает однопараметрическую группу преобразований $s_t: \mathbb{M} \to \mathbb{M}$, то векторное поле a_*X, где a_* – дифференциал отображения a, порождает однопараметрическую группу преобразований $a \circ s_t \circ a^{-1}$.

Доказательство. Ясно, что $a \circ s_t \circ a^{-1}$ есть однопараметрическая группа преобразований. Покажем, что эта группа преобразований индуцирует векторное поле a_*X. Пусть $x \in \mathbb{M}$ – произвольная точка и $y = a^{-1}(x)$. Так как отображение s_t индуцирует векторное поле X, то вектор $X_y \in \mathbb{T}_y(\mathbb{M})$ касается кривой $\gamma_1(t) = s_t(y)$ в точке $y = \gamma_1(0)$. Из этого следует, что вектор

$$(a_*X)_x = a_*X_y \in \mathbb{T}_x(\mathbb{M})$$

касается кривой $\gamma_2(t) = a \circ s_t(y) = a \circ s_t \circ a^{-1}(x)$. □

Замечание. В данном предложении преобразование a не обязано принадлежать какой либо группе. Это может быть произвольный фиксированный диффеоморфизм многообразия $\mathbb{M}$. □

Следствие. Векторное поле $X \in \mathcal{X}(\mathbb{M})$ инвариантно относительно диффеоморфизма $a: \mathbb{M} \to \mathbb{M}$, т.е. $a_*X = X$, тогда и только тогда, когда диффеоморфизм a перестановочен с однопараметрической группой преобразований s_t, порожденной векторным полем X.

10.3 Инвариантные структуры

Рассмотрим группу преобразований $(\mathbb{M}, \mathbb{G})$. Предположим, что на многообразии задана некоторая геометрическая структура, например, тензорное поле. Возникает вопрос при каких условиях эта структура является инвариантной относительно данной группы преобразований ? Ответ на этот вопрос очень важен для приложений, поскольку часто мы ищем решения некоторой системы уравнений, обладающее определенной симметрией. Такие структуры называются $\mathbb{G}$-*инвариантными.*

Начнем с простейшего случая. Предположим, что на многообразии $\mathbb{M}$ задана функция и действие группы преобразований $(\mathbb{M}, \mathbb{G})$. Определим действие элемента группы преобразований $a \in \mathbb{G}$ на функцию формулой

$$f(x) \mapsto (\hat{a}f)(x) := f(xa^{-1})$$

или $(hataf)(xa) = f(x)$. То есть значение новой функции $\hat{f}$ в точке xa равно значению исходной функции в точке x.

Определение. Назовем функцию $\mathbb{G}$-*инвариантной*, если она удовлетворяет условию

$$f(xa) = f(x). \quad \square \tag{10.4}$$

$\square$

В силу этого равенства $\mathbb{G}$-инвариантные функции постоянны на орбитах группах.

Пример 10.3.1. Сферически симметричные функции в трехмерном евклидовом пространстве $\mathbb{R}^3$ постоянны на сферах $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ с центром в начале координат. Эти функции в сферической системе координат зависят только от радиуса $f = f(r)$. $\square$

Пример 10.3.2. Если группа преобразований $(\mathbb{M}, \mathbb{G})$ транзитивна, то $\mathbb{G}$-инвариантные функции на $\mathbb{M}$ постоянны. $\square$

Рассмотрим более сложную ситуацию.

Определение. Пусть многообразии $\mathbb{M}$, на котором действует группа преобразований $(\mathbb{M}, \mathbb{G})$, задано поле $\varphi = \{\varphi^i(x)\}$, имеющее N компонент, $i = 1, \ldots, N$, которое преобразуются по некоторому представлению группы $\mathbb{G}$. Это значит, что группа преобразований действует не только на точки многообразия, но одновременно и на компоненты φ^i по правилу

$$\begin{aligned} x &\mapsto x' = xa, \\ \varphi(x) &\mapsto \varphi'^i(x') = \varphi^j(x) S_j{}^i(a), \end{aligned} \tag{10.5}$$

где $S_j{}^i(a)$ – матрица представления, соответствующая элементу группы $a \in \mathbb{G}$. Поле $\varphi^i(x)$ называется $\mathbb{G}$-инвариантным, если выполнено условие

$$\varphi^i(xa) = \varphi^j(x) S_j{}^i(a). \quad \square \tag{10.6}$$

В отличие от правила преобразования поля (10.5) в левой части этого равенства опущен штрих у поля, и поэтому условие $\mathbb{G}$-инвариантности представляет собой уравнение на поле $\varphi^i(x)$. В случае, если представление группы преобразований тривиально, $S_j{}^i(a) = \delta^i_j$ для всех $a \in \mathbb{G}$, мы имеем набор $\mathbb{G}$-инвариантных скалярных функций (10.4).

Условие $\mathbb{G}$-инвариантности можно переписать в эквивалентном виде

$$\varphi^i(x) = \varphi^j(xa^{-1}) S_j{}^i(a).$$

Пример 10.3.3. Рассмотрим сферически инвариантные векторные поля в трехмерном евклидовом пространстве $\mathbb{R}^3$. Нетрудно проверить, что векторное поле с компонентами

$$\varphi^i(x) = x^i f(r), \qquad i = 1, 2, 3,$$

заданными в декартовой системе координат x^i, где $f(r)$ – произвольная функция от радиуса, является сферически симметричным. Можно также доказать обратное утверждение, что любое сферически симметричное векторное поле имеет такой вид для некоторой функции $f(r)$. Таким образом, сферически симметричные векторные поля на $\mathbb{R}^3$ параметризуются одной функцией радиуса $f(r)$. В разделе 8.1 приведены также сферически симметричные тензоры второго ранга. $\square$

Замечание. При исследовании моделей математической физики предположение о симметрии полей позволяет существенно уменьшить число независимых переменных и число аргументов. Это позволяет в ряде случаев найти явно точное решение уравнений Эйлера–Лагранжа. □

Замечание. При рассмотрении групп преобразований мы предполагаем, что преобразования многообразия $\mathbb{M}$ являются глобальными, т.е. элемент группы $a \in \mathbb{G}$ не зависит от точки многообразия $x \in \mathbb{M}$. При таких преобразованиях калибровочные поля преобразуются, как тензоры. Поэтому условие $\mathbb{G}$-инвариантности (10.6) можно распространить и на калибровочные поля. □

Рассмотрим общую ситуацию. Пусть задана группа преобразований $(\mathbb{M}, \mathbb{G})$. Каждому элементу группы $a \in \mathbb{G}$ соответствует диффеоморфизм $\mathbb{M} \xrightarrow{a} \mathbb{M}$, который мы обозначим той же буквой, что и элемент группы. У этого отображения существует дифференциал a_* и возврат отображения a^*. Возврат отображения обратим, так как a – диффеоморфизм. Пусть в точке $x \in \mathbb{M}$ задан тензор $T(x) \in \mathbb{T}^r_{s,x}(\mathbb{M})$ типа (r, s). Тогда его можно перенести из точки x в точку xa с помощью дифференциала отображения и возврата отображения, которые действуют соответственно на контравариантные и ковариантные индексы, $T(xa) = (a_*)^r (a^{*-1})^s T(x)$.

В координатах это отображение записывается следующим образом. Пусть $x = \{x^\alpha\}$ и $xa = \{y^\alpha\}$, $\alpha = 1, \ldots, \mathsf{dim}\,\mathbb{M}$, – координаты точек. Тогда

$$T_{\beta_1 \ldots \beta_s}{}^{\alpha_1 \ldots \alpha_r}(xa) = \frac{\partial x^{\gamma_1}}{\partial y^{\beta_1}} \ldots \frac{\partial x^{\gamma_s}}{\partial y^{\beta_s}} T_{\gamma_1 \ldots \gamma_s}{}^{\delta_1 \ldots \delta_r}(x) \frac{\partial y^{\alpha_1}}{\partial x^{\delta_1}} \ldots \frac{\partial y^{\alpha_r}}{\partial x^{\delta_r}},$$

где матрицы Якоби вычисляются в точке x.

Определение. Произвольное тензорное поле $T \in \mathcal{T}^r_s(\mathbb{M})$, заданное на многообразии $\mathbb{M}$, называется $\mathbb{G}$-*инвариантным*, если выполнено условие

$$T(xa) = (a_*)^r (a^{*-1})^s T(x) \tag{10.7}$$

для всех $x \in \mathbb{M}$ и $a \in \mathbb{G}$. □

Замечание. Это определение применимо как для счетных групп, так и для групп Ли. □

Поскольку действие группы преобразований транзитивно на каждой орбите, то тензорное поле на всей орбите однозначно задается его значением в некоторой точке x на данной орбите. Это не значит, что тензор в точке x можно задать произвольно, а затем разнести по орбите. Если группа изотропии $\mathbb{G}_x$ нетривиальна, то в соответствии с определением (10.7) тензор в точке x должен быть выбран симметричным относительно преобразований из $\mathbb{G}_x$.

В случае, когда $\mathbb{G}$ – группа Ли, условие (10.7) можно записать в дифференциальной форме. Генераторы группы Ли L_A, $\mathrm{A} = 1, \ldots, \mathsf{dim}\,\mathbb{G}$, – это базис левоинвариантных векторных полей на $\mathbb{G}$. Им соответствуют векторные поля L^*_A на многообразии $\mathbb{M}$. Из определения производной Ли следует, что если тензорное поле является $\mathbb{G}$-инвариантным, то

$$\mathsf{L}_{L^*_\mathrm{A}} T = 0$$

для всех A и $x \in \mathbb{M}$. Это – система дифференциальных уравнений в частных производных первого порядка на компоненты $\mathbb{G}$-инвариантного тензорного поля.

Замечание. В математической физике, как правило, ставится одна из двух задач. 1) По заданному тензорному полю определить его группу симметрии и 2) найти тензорное поле, обладающее определенной симметрией. □

В моделях гравитации важную роль играют однопараметрические группы преобразований, которые оставляют инвариантной метрику, заданную на многообразии $\mathbb{M}$. Этим группам преобразований соответствуют векторные поля, которые называются векторными полями Киллинга. Им будет посвящен отдельный раздел 16.

10.4 Отображения групп преобразований

Пусть задано две группы преобразований $(\mathbb{M},\mathbb{G})$ и $(\mathbb{N},\mathbb{G})$ двух многообразий $\mathbb{M}$ и $\mathbb{N}$ с одной и той же группой $\mathbb{G}$. В частном случае многообразия могут совпадать, $\mathbb{M}=\mathbb{N}$, но действия группы $\mathbb{G}$ на $\mathbb{M}$ можно задавать различными способами.

Пример 10.4.1. Можно рассматривать группу вращений плоскости $\left(\mathbb{R}^2,\mathbb{SO}(2)\right)$ относительно произвольной точки $x\in\mathbb{R}^2$. □

Обозначим действие элемента группы $a\in\mathbb{G}$ на $\mathbb{M}$ и $\mathbb{N}$ соответственно через $a_{\mathbb{M}}$ и $a_{\mathbb{N}}$:

$$a_{\mathbb{M}}:\ \mathbb{M}\times\mathbb{G}\ni x,a\mapsto xa\in\mathbb{M},$$
$$a_{\mathbb{N}}:\ \mathbb{N}\times\mathbb{G}\ni y,a\mapsto ya\in\mathbb{N}.$$

Определение. Отображение $f:\ \mathbb{M}\to\mathbb{N}$ называется *эквивариантным*, если диаграмма

$$\begin{array}{ccc}\mathbb{M}\times\mathbb{G} & \xrightarrow{f\times\mathsf{id}} & \mathbb{N}\times\mathbb{G}\\ \Big\downarrow{\scriptstyle a_{\mathbb{M}}} & & \Big\downarrow{\scriptstyle a_{\mathbb{N}}}\\ \mathbb{M} & \xrightarrow{f} & \mathbb{N}\end{array},$$

где id – тождественное отображение $\mathbb{G}\to\mathbb{G}$, коммутативна. То есть

$$f\circ a_{\mathbb{M}}=a_{\mathbb{N}}\circ f\quad\text{или}\quad f(xa)=f(x)a$$

для всех $a\in\mathbb{G}$ и $x\in\mathbb{M}$. □

Понятие эквивариантного отображения можно обобщить. Рассмотрим две группы преобразований $(\mathbb{M},\mathbb{G})$ и $(\mathbb{N},\mathbb{H})$. Пусть задан гомоморфизм групп $\rho:\ \mathbb{G}\to\mathbb{H}$ и дифференцируемое отображение многообразий $f:\ \mathbb{M}\to\mathbb{N}$.

Определение. Отображение $f\times\rho:\ \mathbb{M}\times\mathbb{G}\to\mathbb{N}\times\mathbb{H}$ называется *эквивариантным*, если диаграмма

$$\begin{array}{ccc}\mathbb{M}\times\mathbb{G} & \xrightarrow{f\times\rho} & \mathbb{N}\times\mathbb{H}\\ \Big\downarrow{\scriptstyle a_{\mathbb{M}}} & & \Big\downarrow{\scriptstyle \rho(a)_{\mathbb{N}}}\\ \mathbb{M} & \xrightarrow{f} & \mathbb{N}\end{array}$$

коммутативна. То есть

$$f \circ a_{\mathbb{M}} = \rho(a)_{\mathbb{N}} \circ (f \times \rho) \quad \text{или} \quad f(xa) = f(x)\rho(a)$$

для всех $a \in \mathbb{G}$ и $x \in \mathbb{M}$. □

Пример 10.4.2. Важным физическим приложением указанной конструкции является понятие спинора. Рассмотрим две группы преобразований: группу вращений $\mathbb{SO}(3)$ трехмерного евклидова пространства $\mathbb{R}^3$ и группу $\mathbb{SU}(2)$ линейных преобразований двумерного комплексного пространства $\mathbb{C}^2$. Согласно теореме 2.8.1, группа $\mathbb{SU}(2)$ два раза накрывает группу вращений $\mathbb{SO}(3)$. Обозначим соответствующий гомоморфизм $\rho :\ \mathbb{SU}(2) \to \mathbb{SO}(3)$. Тогда для спиноров имеем коммутативную диаграмму

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{SU}(2) & \xrightarrow{f \times \rho} & \mathbb{R}^3 \times \mathbb{SO}(3) \\ \Big\downarrow{\scriptstyle a_{\mathbb{C}^2}} & & \Big\downarrow{\scriptstyle \rho(a)_{\mathbb{R}^3}}, \\ \mathbb{C}^2 & \xrightarrow{\qquad f \qquad} & \mathbb{R}^3 \end{array}$$

где $a_{\mathbb{C}^2} = U$ – унитарная 2×2 матрица с единичным определителем линейных преобразований двумерного комплексного пространства $\mathbb{C}^2$, и $\rho(a)_{\mathbb{R}^3}$ – соответствующая 3×3 матрица вращений трехмерного евклидова пространства $\mathbb{R}^3$. Отображение f задается формулой $x^i = (\bar{\psi}\sigma^i\psi)$, где σ^i, $i = 1, 2, 3$, – матрицы Паули, $\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} \in \mathbb{C}^2$ – столбец из двух комплексных чисел (компоненты спинора) и $\bar{\psi} = (\psi_1^*, \psi_2^*)$ – строка из двух комплексно сопряженных чисел (компоненты сопряженного спинора). В физике принято называть матрицы U – спинорным представлением группы вращений $\mathbb{SO}(3)$. Строго говоря, соответствие

$$\mathbb{SO}(3) \ni \rho(a) \mapsto a \in \mathbb{SU}(2)$$

не является отображением и, следовательно, представлением, так как каждому вращению $\rho(a) \in \mathbb{SO}(3)$ соответствуют две унитарные матрицы $\pm a \in \mathbb{SU}(2)$, которые отличаются знаком. Поэтому спинорное представление называют двузначным представлением. □

Глава 11

Гомотопии и фундаментальная группа

Все многообразия локально устроены одинаково так же как и евклидовы пространства. Однако их глобальная структура может существенно отличаться. Описание глобальной структуры многообразий является сложной задачей, и для ее решения применяются различные методы. Например, на каждом многообразии можно задать некоторую алгебраическую структуру, которая является топологическим инвариантом, т.е. сохраняется при диффеоморфизмах. В настоящем разделе мы опишем один из простейших топологических инвариантов – фундаментальную группу. Этот инвариант довольно груб и не различает все топологически различные многообразия. Однако он полезен в приложениях, т.к. если фундаментальная группа двух многообразий различна, то можно с уверенностью сказать, что данные многообразия недиффеоморфны.

11.1 Пути

Важным методом алгебраического анализа глобальной структуры многообразий является нахождение фундаментальной группы многообразия, которая строится на основе понятия *пути* или *кривой*. Определение кривой в евклидовом пространстве (2.9) дословно переносится на произвольное многообразие $\mathbb{M}$ размерности n.

Определение. *Путем* γ называется кусочно дифференцируемое отображение единичного отрезка $[0,1] \in \mathbb{R}$ в $\mathbb{M}$,

$$\gamma : \quad [0,1] \ni t \mapsto x(t) \in \mathbb{M}, \tag{11.1}$$

где t – вещественный параметр вдоль кривой. Говорят, что путь соединяет *начало* пути $x(0)$ и его *конец* $x(1)$. Если начало и конец пути совпадают, $x(0) = x(1)$, то путь называют *замкнутым*. Замкнутый путь называют также *петлей*. Мы будем также писать $\gamma = x(t)$. □

В координатах отображение (11.1) задается набором n кусочно дифференцируемых функций $x(t) = \{x^\alpha(t)\}$, $\alpha = 1, \ldots, n$, одной переменной $t \in [0,1]$.

Замечание. В общем случае кривая $x(t)$ может иметь точки самопересечения, и тогда отображение γ не является взаимно однозначным отображением отрезка $[0,1]$ на его образ $\gamma([0,1])$. □

Замечание. В алгебраической топологии, где в роли $\mathbb{M}$ выступают топологические пространства, от отображения (11.1) требуется только непрерывность. Для алгебраического анализа многообразий удобно рассматривать кусочно дифференцируемые пути, т.е. функции $x^\alpha(t)$ считать кусочно дифференцируемыми. Это удобно при умножении путей, которое будет введено позже. □

Замкнутую кривую в $\mathbb{M}$ можно рассматривать, как непрерывный образ окружности, т.к. при совпадении начала и конца можно отождествить концы единичного интервала $t \in [0,1]$.

Определение. *Простой замкнутой* или *жордановой* кривой называется гомеоморфный образ окружности, т.е. замкнутая кривая без точек самопересечения. В общем случае простая замкнутая кривая может иметь изломы, т.е. от нее требуется только непрерывность. Жорданова кривая на евклидовой плоскости, состоящая из конечного числа прямолинейных отрезков, называется *жордановым* многоугольником. □

Теорема 11.1.1 (Жордан). *Пусть $\mathbb{S}$ – простая замкнутая кривая на евклидовой плоскости $\mathbb{R}^2$. Тогда $\mathbb{R}^2 \setminus \mathbb{S}$ несвязно и состоит из двух компонент, общей границей которых является $\mathbb{S}$. В точности одна из этих компонент ограничена.*

Доказательство. Несмотря на интуитивную очевидность утверждения теоремы, ее доказательство сложно. Оно приведено, например, в [**?**]. □

В определении пути (11.1) область определения параметра вдоль пути фиксирована единичным отрезком. Этого можно не делать, считая, что параметр принимает значения на отрезке произвольной длины $[a,b] \subset \mathbb{R}$. В этом случае все пути можно разделить на классы эквивалентности. А именно, два пути $x_1(t)$ и $x_2(s)$ называются *эквивалентными*, если они связаны гладкой монотонной заменой параметра $s = s(t)$:

$$x_1(t) = x_2\Big(s(t)\Big), \qquad \frac{ds}{dt} > 0.$$

где $s(a) = 0$ и $s(b) = 1$. В частном случае, возможно преобразование параметра t внутри единичного отрезка, $a = 0$, $b = 1$. В дальнейшем под путем будем понимать класс эквивалентных путей, определенных с точностью до гладкой репараметризации

Замечание. Иногда понятие пути и кривой различают, понимая под путем отображение (11.1), а под кривой – образ пути $\gamma([0,1])$ в многообразии $\mathbb{M}$. Как правило, мы этого не делаем и употребляем термины путь и кривая как синонимы. При изучении фундаментальных групп отображение (11.1) будем называть путем, как это принято в алгебраической топологии. □

Замечание. В математической физике под многообразием $\mathbb{M}$ часто понимается пространство-время, а под вещественным параметром t – время. Тогда путь $x(t) = \{x^\alpha(t)\}$ совпадает с траекторией точечной частицы. □

Пример 11.1.1. Простейшим примером пути является *тождественный* путь e_x в некоторой точке $x \in \mathbb{M}$, определенный равенством

$$e_x : \quad [0,1] \ni t \mapsto x \in \mathbb{M} \tag{11.2}$$

для всех t, где $x = \{x^\alpha = \mathsf{const}\} \in \mathbb{M}$ – произвольная фиксированная точка многообразия. □

При рассмотрении евклидова пространства $\mathbb{R}^n$ как топологического пространства в разделе 2.3.2 было введено понятие связности. Введем новое понятие связности, которое удобно, в частности, при рассмотрении фундаментальных групп.

Определение. Топологическое пространство $\mathbb{M}$ называется *линейно связным*, если для любых двух точек $x_1, x_2 \in \mathbb{M}$ найдется соединяющий их путь, целиком лежащий в $\mathbb{M}$. Топологическое пространство $\mathbb{M}$ называется *локально линейно связным*, если для любого $x \in \mathbb{M}$ всякая открытая окрестность точки x содержит линейно связную окрестность этой точки. □

Если какую либо фиксированную точку $x_0 \in \mathbb{M}$ можно соединить путем с любой другой точкой $x_1 \in \mathbb{M}$, то этого достаточно для линейной связности многообразия, потому что тогда две произвольные точки $x_1, x_2 \in \mathbb{M}$ можно соединить путем, проходящим через x_0.

Между просто связностью и линейной связностью топологических пространств имеется тесная связь, которая дается следующей теоремой.

Теорема 11.1.2. *Всякое линейно связное топологическое пространство связно.*

Доказательство. Зафиксируем точку $x_0 \in \mathbb{M}$. Тогда ее можно соединить некоторой кривой γ_x, целиком лежащей в $\mathbb{M}$, с произвольной точкой $x \in \mathbb{M}$, потому что $\mathbb{M}$ – линейно связно. Каждая кривая γ_x является связным подмножеством в $\mathbb{M}$. Тогда все пространство $\mathbb{M}$ является объединением, $\mathbb{M} = \bigcup_x \gamma_x$, и связно, как объединение связных подмножеств, имеющих непустое пересечение (теорема 2.3.14). □

Обратное утверждение теоремы 11.1.2 неверно. Приведем пример связного, но не линейно связного топологического пространства.

Пример 11.1.2 (Блоха и гребенка). Пусть подмножество комплексной плоскости $\mathbb{W} \subset \mathbb{C}$ является объединением двух множеств $\mathbb{W} = \mathbb{U} \bigcup \mathbb{V}$ (см. рис. 11.1), где

$$\begin{array}{rcll} \mathbb{U} & = & \{i\} & \text{– блоха} \\ \mathbb{V} & = & [0,1] \bigcup \{1/n + iy : \quad n \in \mathbb{N},\ 0 \le y \le 1\} & \text{– гребенка.} \end{array}$$

Будем считать, что топология $\mathbb{W}$ индуцирована вложением в комплексную плоскость

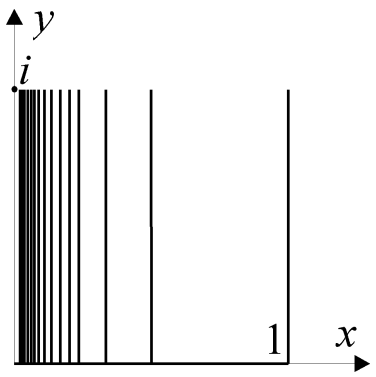


Рис. 11.1: Пример связного, но не линейно связного топологического пространства "блоха и гребенка".

$\mathbb{C} \approx \mathbb{R}^2$. Прежде всего заметим, что множество $\mathbb{V}$ линейно связно как объединение линейно связных множеств (теорема 2.3.14) и потому связно. Пусть $\mathbb{U}_\epsilon$ – ϵ-окрестность

точки i, т.е. пересечение множества $\{z: \ |z-i|<\epsilon\}$ с $\mathbb{W}$. Подмножество $\mathbb{U}$, состоящее их одной точки, принадлежит замыканию $\overline{\mathbb{V}}$, т.к. для любого $\epsilon>0$, пересечение $\mathbb{U}_\epsilon \bigcap \mathbb{V}$ не пусто, поскольку для достаточно больших n точки $1/n+i \in \mathbb{V}$ лежат в ϵ-окрестности точки i (лемма 2.3.1). Отсюда вытекает, что связная компонента $\mathbb{V}$ содержит $\mathbb{U}$ и поэтому совпадает со всем $\mathbb{W}$.

Покажем теперь, что топологическое пространство $\mathbb{W}$ не является линейно связным, т.к. точку $i \in \mathbb{W}$ нельзя соединить путем с любой другой точкой из $\mathbb{W}$. Предположим, что существует путь с началом в точке $\gamma(0)=i$ и концом в некоторой точке $\gamma(1) \in \mathbb{V}$. Не ограничивая общности можно считать, что $\gamma(t) \neq i$ при $t>0$. Из непрерывности отображения $\gamma(t)$ следует, что для любого положительного ϵ существует такое положительное число δ, что при $t<\delta$ выполняется неравенство $|x(t)-i|<\epsilon$, т.е. путь лежит в ϵ-окрестности $\mathbb{U}_\epsilon$ точки i. Выберем конечное число $\epsilon<1$ и некоторое значение параметра $0<t_0<\delta$. По предположению, $x(t_0) \in \mathbb{V}$. Окрестность $\mathbb{U}_\epsilon$ многосвязна, т.к. состоит из самой точки i и объединения непересекающихся интервалов. Отрезок $[0,t_0]$ связен и, как показано в примере 2.4.10, его невозможно непрерывно отобразить на многосвязную область таким образом, чтобы начало и конец отрезка лежали в разных компонентах.

Построенное топологическое пространство не является многообразием, хотя бы потому, что точка i не имеет окрестности, гомеоморфной прямой $\mathbb{R}$. □

Замечание. В рассмотренном примере топологическое пространство $\mathbb{W}$ не является многообразием по двум причинам. Во-первых, окрестность точки i нельзя отобразить на интервал и, во вторых, множество $\mathbb{V}$ имеет точки самопересечения. □

Следующий пример показывает, что не всякое линейно связное топологическое пространство является локально линейно связным.

Пример 11.1.3 (Польская окружность). Рассмотрим подмножество на евклидовой плоскости $\mathbb{U} \subset \mathbb{R}^2$, которое имеет вид $\mathbb{U}=\mathbb{U}_1 \cup \mathbb{U}_2 \cup \mathbb{U}_3$, где

$$\begin{aligned}
\mathbb{U}_1 &= \{(x,y) \in \mathbb{R}^2: \ \ x^2+y^2=1, \ \ y \geq 0\},\\
\mathbb{U}_2 &= \{(x,y) \in \mathbb{R}^2: \ \ -1 \leq x \leq 0, \ \ y=0\},\\
\mathbb{U}_3 &= \{(x,y) \in \mathbb{R}^2: \ \ 0<x\leq 1, \ \ y=\frac{1}{2}\sin\frac{\pi}{x}\},
\end{aligned}$$

с индуцированной топологией, см.рис.11.2. Это пространство линейно связно, но не локально линейно связно, т.к. начало координат $(0,0)$ не имеет окрестности, содержащей линейно связную окрестность. Это топологическое пространство не является многообразием, поскольку начало координат не имеет окрестности, гомеоморфной прямой $\mathbb{R}$. □

Следующие свойства справедливы для произвольных топологических пространств.

Предложение 11.1.1. *Образ линейно связного многообразия при непрерывном отображении линейно связен.*

Следствие. Если $\mathbb{M}_1$ и $\mathbb{M}_2$ два гомеоморфных топологических пространства, то $\mathbb{M}_1$ линейно связно тогда и только тогда, когда $\mathbb{M}_2$ линейно связно. □

Доказательство. См., например, в [**?**]. □

Предложение 11.1.2. *Многообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ линейно связны тогда и только тогда, когда их прямое произведение $\mathbb{M}_1 \times \mathbb{M}_2$ линейно связно.*

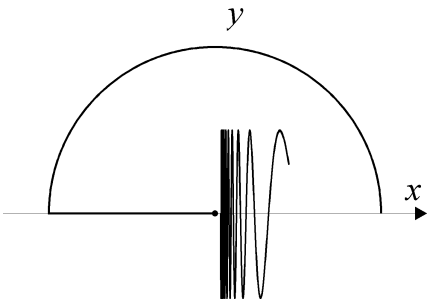


Рис. 11.2: Пример линейно связного, но не локально линейного связного топологического пространства.

Доказательство. См., например, в [**?**]. □

Для многообразий понятие связности и линейной связности эквивалентны.

Теорема 11.1.3. *Многообразие* $\mathbb{M}$ *связно тогда и только тогда, когда* $\mathbb{M}$ *– линейно связно.*

Доказательство. Достаточность является утверждением теоремы 11.1.2.

Необходимость докажем от противного. Пусть $\mathbb{M}$ связное, но не линейно связное многообразие. Прежде всего заметим, что каждое многообразие можно покрыть счетным числом карт. Каждая карта является линейно связным подмножеством, потому что гомеоморфна $\mathbb{R}^n$. Если две карты пересекаются, то их объединение также линейно связно. Это значит, что многообразие можно представить в виде объединения не более, чем счетного числа непересекающихся подмножеств, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, где каждое $\mathbb{U}_i$ линейно связно и $\mathbb{U}_i \cap \mathbb{U}_j = \emptyset$ при $i \neq j$. Поскольку $\mathbb{M}$ связно, то среди подмножеств $\mathbb{U}_i$ существует по крайней мере одно подмножество, пусть это будет $\mathbb{U}_i$, которое содержит предельную точку некоторого другого подмножества $\mathbb{U}_j$. Пусть точка $x_0 \in \mathbb{U}_i$ является предельной для подмножества $\mathbb{U}_j$. Это значит, что в $\mathbb{U}_j$ существует последовательность точек $\{x_k\} \in \mathbb{U}_j$, которая сходится к x_0. Поэтому, начиная с некоторого N, все точки последовательности x_k, $k >$ N, лежат в окрестности $\mathbb{U}_0 \ni x_0$. Следовательно, точку x_0 можно соединить кривой с x_k и подмножества $\mathbb{U}_i$ и $\mathbb{U}_j$ – линейно связны, что противоречит предположению. □

Пример 11.1.4. Евклидово пространство $\mathbb{R}^n$ линейно связно, т.к. любые две точки $x_0, x_1 \in \mathbb{R}^n$ можно соединить путем $x(t) = x_0 + (x_1 - x_0)t$, который представляет собой прямолинейный отрезок. □

Нетрудно также доказать

Предложение 11.1.3. *Каждое многообразие* $\mathbb{M}$ *локально линейно связно.*

В общем случае многосвязное многообразие представляет собой набор компонент связности, каждая из которых является линейно связной.

Для полноты картины приведем достаточное условие линейной связности произвольного топологического пространства.

Предложение 11.1.4. *Если топологическое пространство* $\mathbb{M}$ *связно и локально линейно связно, то оно линейно связно.*

Доказательство. См., например, [**?**]. □

11.2 Гомотопия непрерывных отображений

На множестве непрерывных отображений двух топологических пространств $f:\ \mathbb{M}_1 \to \mathbb{M}_2$ и, в частности, многообразий можно ввести отношение эквивалентности – гомотопию. Понятие гомотопии является топологическим, потому что при его определении используется только непрерывность отображений.

Определение. Два непрерывных отображения

$$f_0, f_1:\quad \mathbb{M}_1 \ni x \mapsto f_0(x), f_1(x) \in \mathbb{M}_2$$

называются *гомотопными*, если существует непрерывное отображение

$$F:\quad \mathbb{M}_1 \times [0,1] \ni x, t \mapsto F(x,t) \in \mathbb{M}_2$$

такое, что $F(x,0) = f_0(x)$ и $F(x,1) = f_1(x)$. При этом само отображение F называется *гомотопией* между f_0 и f_1:

$$F:\quad f_0 \sim f_1.$$

Гомотопия $F(x,t)$ называется также *деформацией* отображения f_1 в f_2. □

Другими словами, гомотопия позволяет непрерывно деформировать одно отображение в другое. Нетрудно проверить, что гомотопия является отношением эквивалентности на множестве всех отображений $\mathbb{M}_1 \to \mathbb{M}_2$. Отображения, гомотопные отображению f, образуют гомотопический класс отображений, который будем обозначать квадратными скобками $[f]$. Множество всех гомотопических классов отображений $\mathbb{M}_1 \to \mathbb{M}_2$ будем обозначать через $[\mathbb{M}_1; \mathbb{M}_2]$.

Гомотопию $F(x,t)$ можно представить себе следующим образом. Зафиксируем точку $x \in \mathbb{M}_1$ и будем менять параметр $t \in [0,1]$. Тогда мы получим траекторию точки x в пространстве $\mathbb{M}_2$. Тогда гомотопия $F(x,t)$ представляет собой семейство всех траекторий, зависящих от параметра $x \in \mathbb{M}_1$.

Пример 11.2.1. Любой путь $\gamma:\ [0,1] \ni s \mapsto x(s) \in \mathbb{M}$ гомотопен постоянному пути e_{x_0} с началом и концом в точке $x_0 = x(0)$ посредством гомотопии

$$F:\quad [0,1] \times [0,1] \ni s, t \mapsto F(s,t) \in \mathbb{M},\quad \text{где}\quad F(s,t) = x\Big((1-t)s\Big).$$

В этой гомотопии конец пути меняется: он скользит вдоль кривой $\gamma(s)$, приближаясь к началу пути при $t \to 1$. При данном выше определении гомотопии, когда отображения f_0 и f_1 не обязаны совпадать ни в одной точке, все пути на связном многообразии гомотопны между собой. □

Понятие гомотопных отображений можно использовать для определения отношения эквивалентности самих топологических пространств.

Определение. Говорят, что пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ имеют один и тот же *гомотопический тип*, если существуют непрерывные отображения $f:\ \mathbb{M}_1 \to \mathbb{M}_2$ и $g:\ \mathbb{M}_2 \to \mathbb{M}_1$ такие, что композиции этих отображений гомотопны тождественным отображениям:

$$g \circ f \sim \mathsf{id}_1:\quad \mathbb{M}_1 \to \mathbb{M}_1, \qquad f \circ g \sim \mathsf{id}_2:\quad \mathbb{M}_2 \to \mathbb{M}_2,$$

где $\mathsf{id}_{1,2}$ – тождественные отображения пространств $\mathbb{M}_1$ и $\mathbb{M}_2$. Отображения f и g называются в этом случае *гомотопическими эквивалентностями*. Говорят также, что пространства $\mathbb{M}_1$ и $\mathbb{M}_2$ в этом случае *гомотопически эквивалентны*. □

Это определение отличается от определения гомеоморфизма тем, что вместо знака равенства стоит знак гомотопической эквивалентности. Поэтому два гомеоморфных пространства имеют одинаковый гомотопический тип. Обратное утверждение неверно.

Пример 11.2.2. n-мерный шар $\mathbb{B}^n \in \mathbb{R}^n$ гомотопически эквивалентен точке $x \in \mathbb{B}^n$, но не гомеоморфен ей. □

Определение. В общем случае топологическое пространство, гомотопически эквивалентное точке, называется *стягиваемым*. □

Другими словами, пространство стягиваемо, если его можно непрерывно деформировать по себе в точку.

Пример 11.2.3. Нестягиваемыми пространствами являются n-мерная сфера $\mathbb{S}^n$ и тор $\mathbb{T}^n$. □

Пример 11.2.4. Евклидово пространство $\mathbb{R}^n$, а также любая выпуклая область $\mathbb{U} \subset \mathbb{R}^n$ гомотопически эквивалентна одной точке $x_0 \in \mathbb{U}$. Напомним, что область $\mathbb{U} \subset \mathbb{R}^n$ называется *выпуклой*, если наряду с произвольными точками $x_1, x_2 \in \mathbb{U}$ она содержит также и отрезок $x_1 + (x_2 - x_1)t$, их соединяющий. □

Пример 11.2.5. Евклидово пространство с выколотой точкой, $\mathbb{R}^n \setminus x_0$, гомотопически эквивалентно сфере $\mathbb{S}^{n-1}$, содержащей точку x_0 внутри. В частности, плоскость с выколотой точкой гомотопически эквивалентна окружности. □

Пример 11.2.6. Плоскость с двумя выколотыми точками $x_1, x_2 \in \mathbb{R}^2$ гомотопически эквивалентна *букету окружностей*, изображенному на рис. 11.3. Букет окружностей представляет собой объединение двух окружностей, имеющих одну общую точку, и не является многообразием. □

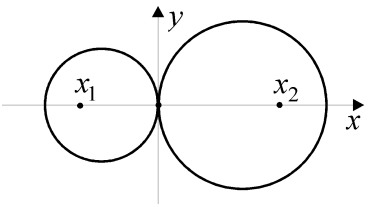


Рис. 11.3: Букет двух окружностей.

Пример 11.2.7. Гомотопически эквивалентными пространствами является тройка: цилиндр, лист Мебиуса и окружность. □

Из этих примеров видно, что гомотопически эквивалентные многообразия могут иметь разную размерность и ориентируемость. Многообразия могут быть также гомотопически эквивалентны топологическим пространствам, которые не являются многообразиями, как показывает пример плоскости с двумя выколотыми точками. Тем не менее некоторые топологические свойства являются общими для гомотопически эквивалентных пространств.

Предложение 11.2.1. *Если пространство $\mathbb{M}_1$ связно и имеет тот же гомотопический тип, что и $\mathbb{M}_2$, то пространство $\mathbb{M}_2$ также связно.*

Доказательство. □

Понятие гомотопии может быть использовано для доказательства нетривиальных утверждений.

Теорема 11.2.1 (**О невозможности причесать ежа**). *На четномерной сфере $\mathbb{S}^n$ нельзя задать непрерывное векторное поле, которое не обращается в нуль ни в одной точке.*

Доказательство. От противного. Рассмотрим сферу единичного радиуса

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \quad x^\alpha x_\alpha = 1, \quad \alpha = 1, \ldots, n+1\},$$

вложенную в евклидово пространство $\mathbb{R}^{n+1}$. Допустим, что на сфере задано непрерывное векторное поле X, не имеющее нулей. Тогда векторное поле $T = X/\sqrt{X^2}$ имеет единичную длину и поэтому задает непрерывное отображение сферы на себя $\mathbb{S}^n \to \mathbb{S}^n$. Касательный вектор к сфере $T = \{T^\alpha(x)\}$ ортогонален радиус-вектору $\{x^\alpha\}$, т.е. $x^\alpha T_\alpha = 0$. При этом мы рассматриваем координаты точки на сфере и касательный вектор как векторы евклидова пространства $\mathbb{R}^{n+1}$. Тогда отображение

$$F(x,t): \qquad x^\alpha, t \mapsto x^\alpha \cos(\pi t) + T^\alpha \sin(\pi t)$$

задает гомотопию

$$F: \qquad \mathbb{S}^n \times [0,1] \to \mathbb{S}^n.$$

Эта гомотопия переводит произвольную точку на сфере $\{x^\alpha\}$ в ее антиподальную точку с координатами $\{-x^\alpha\}$. Выберем ориентацию на сфере, согласованную с ориентацией евклидова пространства (**??**). Тогда объем сферы равен

$$V = \int_{\mathbb{S}^n} \upsilon \neq 0, \tag{11.3}$$

где форма объема на сфере υ получена сужением канонической формы объема евклидова пространства υ_0 на сферу $\mathbb{S}^n$:

$$\upsilon = \mathsf{i}_N \upsilon_0, \qquad \upsilon_0 = dx^1 \wedge dx^2 \wedge \ldots \wedge dx^{n+1}.$$

Здесь i_N – внутреннее умножение формы объема υ_0 на внешнюю нормаль к сфере N. Интеграл (11.3) задает непрерывное отображение $\Lambda_{n+1}(\mathbb{R}^{n+1}) \to \mathbb{R}$. При отображении $x^\alpha \to -x^\alpha$ объем сферы преобразуется по правилу

$$V \to (-1)^{n+1} V.$$

Поскольку изменение знака объема при гомотопии невозможно, т.к. она является непрерывным отображением, то для четных n приходим к противоречию. □

Замечание. На нечетномерной сфере непрерывные векторные поля, нигде не обращающиеся в нуль, существуют. А именно, можно доказать [**?**], что максимальное число линейно независимых нигде не обращающихся в нуль векторных полей на нечетномерной сфере $\mathbb{S}^{n-1}$ равно $2^c + 8d - 1$, где c и d – неотрицательные целые числа, которые определяются следующим образом. Поскольку n четно, то его единственным образом можно представить в виде $n = (2a-1)2^b$, $a, b \in \mathbb{N}$. Тогда $c = b \bmod 4$ и $d = (b-c)/4$. □

Рассмотренный выше пример 11.2.1 гомотопных путей показывает, что определение гомотопии является грубым, т.к. гомотопические классы путей содержат слишком много элементов. Более тонкое деление на классы дает понятие отображений, гомотопных относительно некоторого подмножества $\mathbb{U} \subset \mathbb{M}_1$.

Определение. Два непрерывных отображения топологических пространств

$$f_0, f_1 : \quad \mathbb{M}_1 \to \mathbb{M}_2$$

называются *гомотопными относительно подмножества* $\mathbb{U} \subset \mathbb{M}_1$, если существует гомотопия

$$F : \quad \mathbb{M}_1 \times [0,1] \to \mathbb{M}_2,$$

такая, что $F(x,t)$ не зависит от t при $x \in \mathbb{U}$. Мы пишем $f_0 \sim f_1(\,\mathsf{rel}\mathbb{U})$. □

В частности, отображения f_0 и f_1 должны совпадать на подмножестве $\mathbb{U}$. Другими словами, при непрерывной деформации отображений допускается их изменение только на множестве $\mathbb{M}\setminus\mathbb{U}$. Соответствующая гомотопия называется *относительной гомотопией* и также определяет отношение эквивалентности на множестве непрерывных отображений топологических пространств, которое является более тонким.

Пример 11.2.8. Можно рассмотреть класс путей, гомотопных относительно начала $x(0)$ и конца $x(1)$ некоторого пути. Этот класс содержит все пути, которые имеют фиксированное начало и конец и которые можно непрерывно деформировать друг в друга по многообразию $\mathbb{M}$. Более того, как показывает пример тора, не все пути с одинаковым началом и концом можно непрерывно деформировать друг в друга, не выходя из многообразия $\mathbb{M}$. Если начало и конец пути отличаются, то пути негомотопны тождественному пути. Это говорит о том, что понятие относительной гомотопии более тонко, чем просто гомотопия. □

Понятие относительной гомотопии для путей будет использовано при определении фундаментальной группы многообразия.

11.3 Фундаментальная группа

Имеется два простых, но важных способа получения новых путей из старых. А именно, можно определить произведение путей, когда сначала проходится путь γ_1, а затем – путь γ_2, если такое возможно. Кроме того, можно определить *обратный* путь γ^{-1}, который проходится в обратном направлении.

Предложение 11.3.1. *1) Если начало второго пути $\gamma_2 = x_2(t)$ совпадает с концом первого пути $\gamma_1 = x_1(t)$, то можно определить произведение путей $\gamma_2 \circ \gamma_1$ следующей формулой:*

$$(\gamma_2 \circ \gamma_1) = x(t) := \begin{cases} x_1(2t) & при \quad 0 < t < \frac{1}{2}, \\ x_2(2t-1) & при \quad \frac{1}{2} < t < 1, \end{cases} \tag{11.4}$$

которое является путем в $\mathbb{M}$.

2) Для любого пути γ можно определить путь γ^{-1}, получаемый как путь γ, проходимый в обратном направлении:

$$\gamma^{-1} = x(1-t). \tag{11.5}$$

Эти операции похожи, но не являются групповыми операциями на множестве всех путей. Действительно, умножение путей определено только в том случае, если конец первого пути совпадает с началом второго. Кроме того, произведения $\gamma^{-1} \circ \gamma$ и $\gamma \circ \gamma^{-1}$ дают не тождественные, а замкнутые пути в началом и концом в точках $x(0)$ и $x(1)$ соответственно.

Разобьем все множество путей γ на классы эквивалентных путей $[\gamma]$, которые гомотопны относительно начала и конца пути. То есть каждый элемент $[\gamma]$ содержит все пути, которые можно непрерывно деформировать друг в друга при фиксированном начале и конце путей: $\gamma_1 \in [\gamma]$ тогда и только тогда, когда $\gamma_1 \sim \gamma(\,\mathsf{rel}\{0,1\})$. В классах эквивалентных путей также можно ввести умножение путей и обратный элемент, выбрав из каждого класса по представителю и воспользовавшись формулами (11.4) и (11.5). Нетрудно показать, что эти операции на множестве относительно гомотопных путей корректно определены, т.к. не зависят от выбора представителя в каждом классе. В этих классах произведения $[\gamma^{-1}]\circ[\gamma] = [e_{\gamma(0)}]$ и $[\gamma]\circ[\gamma^{-1}] = [e_{\gamma(1)}]$ дают классы путей, гомотопных тождественным путям. Однако умножение определено не для всех классов путей.

Ситуацию можно исправить и превратить некоторое подмножество всех путей на $\mathbb{M}$ в группу, если зафиксировать точку многообразия и ограничиться замкнутыми путями, имеющими начало и конец в данной точке.

Зафиксируем произвольную точку многообразия $x_0 \in \mathbb{M}$ и рассмотрим все замкнутые пути, имеющие начало и конец в данной точке. Множество этих путей обозначим через $\Omega(\mathbb{M}, x_0)$. При этом допускается, чтобы пути имели точки самопересечения. Поскольку начала и концы всех путей совпадают, то на этом множестве определена операция умножения (11.4). Для каждого пути определен также обратный путь (11.5). Множество $\Omega(\mathbb{M}, x_0)$ с введенной операцией умножения и обратного элемента группы не образует, потому что произведения двух различных путей на их обратные не совпадают, и, следовательно, невозможно ввести понятие единственного единичного элемента. Эту трудность можно обойти, если разбить множество всех путей $\Omega(\mathbb{M}, x_0)$ на классы путей, гомотопных относительно начала и конца, которые обозначим через $\pi(\mathbb{M}, x_0)$. Справедлива следующая

Теорема 11.3.1. *Гомотопические классы путей $\pi(\mathbb{M}, x_0)$ образуют группу относительно операции умножения путей, причем обратным элементом является гомотопический класс обратного пути, а единицей – гомотопический класс единичного пути e_{x_0}.*

Доказательство. Сводится к проверке корректности операций и групповых свойств. Детали можно найти, например, в [**?**]. □

Таким образом, для определения группы было сделано два шага. Во-первых, чтобы определить операцию умножения на всех элементах, множество всех путей было ограничено только замкнутыми путями, имеющими начало и конец в фиксированной точке многообразия. Во-вторых, для корректного определения единичного элемента мы перешли от умножения отдельных путей к умножению классов относительно гомотопных путей.

Определение. Группа $\pi(\mathbb{M}, x_0)$ называется *фундаментальной группой* топологического пространства и, следовательно, многообразия $\mathbb{M}$ в точке x_0. □

В этом определении используется фиксированная точка x_0. Для связных многообразий роль фиксированной точки не является существенной.

Теорема 11.3.2. *Пусть* $\mathbb{M}$ *– линейно связное топологическое пространство (многообразие). Тогда для любых точек* $x_0, x_1 \in \mathbb{M}$ *фундаментальные группы* $\pi(\mathbb{M}, x_0)$ *и* $\pi(\mathbb{M}, x_1)$ *изоморфны.*

Доказательство. Пусть δ – путь из x_0 в x_1 и γ – произвольный замкнутый путь в точке x_0. Тогда $\delta \circ \gamma \circ \delta^{-1}$ – замкнутый путь в точке x_1. Поэтому между множествами замкнутых путей $\Omega(\mathbb{M}, x_0)$ и $\Omega(\mathbb{M}, x_1)$ устанавливается взаимно однозначное соответствие. Отсюда следует, что между фундаментальными группами существует биекция $f_\delta : \ \pi(\mathbb{M}, x_0) \to \pi(\mathbb{M}, x_1)$. Эта биекция – гомоморфизм групп, т.к.

$$f_\delta([\gamma_2] \circ [\gamma_1]) = [\delta] \circ [\gamma_2] \circ [\gamma_1] \circ \delta^{-1} = [\delta] \circ [\gamma_2] \circ \delta^{-1} \circ \delta \circ [\gamma_1] \circ \delta^{-1} = f_\delta([\gamma_2]) \circ f_\delta([\gamma_2]).$$

Обратное отображение f_δ^{-1} определяется обратным путем δ^{-1}. □

В силу этой теоремы при обозначении фундаментальной группы фиксированную точку x_0 мы часто будем опускать: $\pi(\mathbb{M}, x_0) = \pi(\mathbb{M})$. Изоморфизм между фундаментальными группами f_δ определяется гомотопическим классом пути $[\delta]$ из x_0 в x_1. Поскольку нет естественного выбора такого класса путей, то естественного изоморфизма между группами не существует.

Пример 11.3.1. В евклидовом пространстве $\mathbb{R}^n$ при любом n произвольный замкнутый путь для любой отмеченной точки гомотопен тождественному пути в этой точке. Поэтому фундаментальная группа евклидова пространства тривиальна $\pi(\mathbb{R}^n) = e$. □

Определение. Топологическое пространство (многообразие) $\mathbb{M}$ называется *односвязным*, если его фундаментальная группа тривиальна $\pi(\mathbb{M}) = e$. □

Предложение 11.3.2. *Если* $\mathbb{M}_1$ *и* $\mathbb{M}_2$ *– односвязные многообразия, то их прямое произведение* $\mathbb{M}_1 \times \mathbb{M}_2$ *односвязно.*

Доказательство. Простая проверка. □

Пример 11.3.2. Фундаментальная группа окружности с отмеченной точкой изоморфна группе целых чисел по сложению, $\pi(\mathbb{S}^1, x_0) \simeq \mathbb{Z}$. При этом целое число равно количеству полных обходов окружности, причем положительные и отрицательные числа соответствуют обходам против и по часовой стрелки соответственно. Хотя изоморфизм фундаментальной группы окружности группе целых чисел интуитивно очевиден, доказательство громоздко. Его можно найти, например, в [**?**]. □

Пример 11.3.3. Фундаментальная группа цилиндра совпадает с фундаментальной группой окружности и изоморфна группе целых чисел по сложению $\mathbb{Z}$. □

В общем случае фундаментальная группа не является абелевой.

Пример 11.3.4. Рассмотрим евклидову плоскость $\mathbb{R}^2$ с двумя выколотыми точками x_1 и x_2 и фиксированной точкой x_0 (см. рис.11.4). Пусть γ_1 и γ_2 – замкнутые пути с началом и концом в точке x_0, которые охватывают соответственно точки x_1 и x_2. На рис.11.4 показаны представители из классов гомотопных путей $[\gamma_2] \circ [\gamma_1]$ и $[\gamma_1] \circ [\gamma_2]$. Ясно, что эти пути не могут быть непрерывно деформированы друг в друга без пересечения выколотых точек и поэтому негомотопны друг другу. Следовательно, фундаментальная группа евклидовой плоскости с двумя выколотыми точками некоммутативна. □

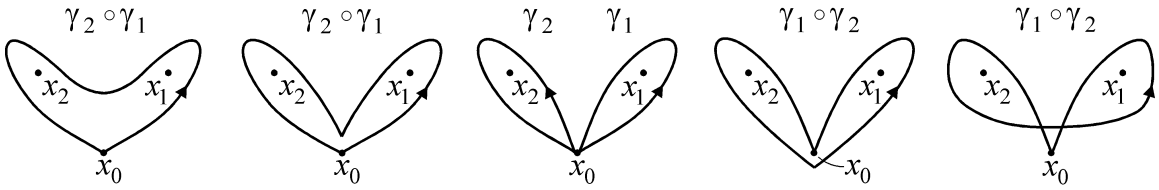


Рис. 11.4: Замкнутые пути γ_1 и γ_2 вокруг точек x_1 и x_2 на евклидовой плоскости (в центре). Непрерывная деформация произведения путей $[\gamma_2] \circ [\gamma_1]$ (влево) и $[\gamma_1] \circ [\gamma_2]$ (вправо).

В общем случае фундаментальная группа многообразия представляет собой свободную группу с конечным числом образующих. Напомним

Определение. Группа $\mathbb{G}$ называется *свободной* группой с системой E порождающих (образующих) элементов, если любое отображение множества E в любую группу $\mathbb{H}$ продолжается до гомоморфизма $\mathbb{G} \to \mathbb{H}$. Система $E \subset \mathbb{G}$ называется *системой свободных порождающих*. Ее мощность называется называется *рангом свободной группы* $\mathbb{G}$. Множество E называется также *алфавитом*. Элементы $g \in \mathbb{G}$ представляют собой *слова* в алфавите E, т.е. выражения в виде конечного произведения

$$g = e_{i_1}^{\epsilon_1} e_{i_2}^{\epsilon_2} \dots e_{i_k}^{\epsilon_k}, \qquad e_{i_j} \in E, \quad \epsilon_j = \pm 1, \quad \forall j = 1, \dots, k, \tag{11.6}$$

а также пустое слово. Слово называется *несократимым*, если

$$e_{i_j}^{\epsilon_j} \neq e_{i_{j+1}}^{-\epsilon_{j+1}}, \qquad \forall j = 1, \dots, k-1.$$

Несократимые слова являются разными элементами свободной группы $\mathbb{G}$, и каждое слово равно единственному несократимому слову. Число k называется *длиной* слова g, если оно несократимо. □

Пример 11.3.5. Рассмотрим расширенную комплексную плоскость $\overline{\mathbb{C}}$ (сферу Римана).

Определение. Область $\mathbb{D} \subset \overline{\mathbb{C}}$ называется *m-связной*, если ее граница состоит из m связных компонент. □

Любая m-связная область диффеоморфна комплексной плоскости $\mathbb{C}$ с $m-1$ дырками.

Предложение 11.3.3. *Фундаментальная группа произвольной m-связной области в расширенной комплексной плоскости $\overline{\mathbb{C}}$ является свободная группа с $m-1$ образующими.*

Доказательство. Каждой связной компоненте границы соответствует ровно одна компонента дополнения $\overline{\mathbb{C}} \backslash \mathbb{D}$. Обозначим эти компоненты $\mathbb{B}_0, \mathbb{B}_1, \dots, \mathbb{B}_{m-1}$ (см. рис.11.5). Заключим каждую компоненту $\mathbb{B}_i$ в некоторую окрестность $\mathbb{D}_i$, которая ограничена простой замкнутой кривой γ_i, целиком лежащей в $\mathbb{D}$. Области $\mathbb{D}_i$ выберем такими, чтобы кривые γ_i не пересекались. На каждой кривой выберем некоторую точку z_i, которую будем считать началом и концом кривой γ_i. Из фиксированной точки z_0 проведем в каждую точку z_i кривую β_i, целиком лежащую в области $\mathbb{D}$. Обозначим

$$e_i := \beta_i^{-1} \circ \gamma_i \circ \beta_i, \qquad i = 0, 1, \dots, m-1.$$

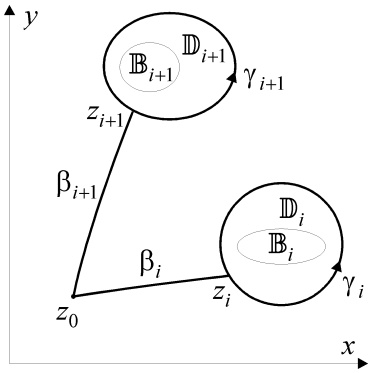


Рис. 11.5: Образующие m-связной области в расширенной комплексной плоскости.

Заметим, что справедливо тождество

$$[e_0] \circ [e_1] \circ \ldots \circ [e_{m-1}] = 1.$$

Поэтому гомотопические классы $[e_1], \ldots, [e_{m-1}]$ можно выбрать в качестве образующих фундаментальной группы $\pi(\mathbb{D}, z_0)$. Можно проверить, что любой путь γ гомотопен конечному произведению:

$$[\gamma] = [\gamma_{i_1}^{\epsilon_1}] \circ [\gamma_{i_2}^{\epsilon_2}] \circ \ldots \circ [e_{i_k}^{\epsilon_k}],$$

где все i_j, $j = 1, \ldots, k$, принадлежат множеству $\{1, \ldots, m-1\}$. При этом данное представление единственно, если оно несократимо. □ □

Нетрудно доказать, что при непрерывном отображении $f : \mathbb{M}_1 \to \mathbb{M}_2$ гомотопически эквивалентные пути $\gamma_1 \sim \gamma_2$ переходят в гомотопически эквивалентные пути $f\gamma_1 \sim f\gamma_2$ и что замкнутые пути переходят в замкнутые. При этом между фундаментальными группами возникает связь.

Теорема 11.3.3. *При непрерывном отображении $f : \mathbb{M}_1 \to \mathbb{M}_2$ фундаментальные группы испытывают гомоморфизм*

$$f_* : \quad \pi(\mathbb{M}_1, x) \to \pi\Big(\mathbb{M}_2, f(x)\Big),$$

где x – произвольная точка из $\mathbb{M}_1$, который называется индуцированным и одинаков для всех отображений, гомотопных относительно точки x.

Доказательство. См., например, [**?**]. □

Замечание. Если отображение двух пространств является гомеоморфизмом, то гомоморфизм фундаментальных групп становится изоморфизмом. □

Понятие фундаментальной группы позволяет осуществить переход от топологии пространств к алгебре и является одним из разделов алгебраической топологии. Этот переход позволяет использовать алгебраические методы при изучении топологических пространств и, следовательно, многообразий. При этом 1) каждому топологическому пространству с отмеченной точкой ставится в соответствие фундаментальная группа, 2) каждому непрерывному отображению пространств ставится в соответствие гомоморфизм групп, 3) тождественному отображению отвечает тождественный гомоморфизм, 4) гомеоморфизму отвечает изоморфизм, 5) композиции непрерывных отображений сопоставляется композиция гомоморфизмов.

Замечание. Отмеченные выше свойства 1) – 5) дают пример функтора из категории топологических пространств с отмеченной точкой и непрерывными отображениями, сохраняющими отмеченную точку, в категорию групп и их гомоморфизмов. □

Рассмотренные примеры окружности и цилиндра показывают, что изоморфизм фундаментальных групп не означает гомеоморфизма пространств. Однако, если фундаментальные группы не изоморфны, то соответствующие пространства заведомо не гомеоморфны.

Теорема 11.3.4. *Пусть* $f: \mathbb{M}_1 \to \mathbb{M}_2$ *– гомотопическая эквивалентность двух пространств, тогда индуцированное отображение* $f_*: \pi(\mathbb{M}_1, x) \to \pi\Big(\mathbb{M}_2, f(x)\Big)$ *является изоморфизмом для любой точки* $x \in \mathbb{M}_1$.

Доказательство. См., например, [**?**]. □

Следствие. Стягиваемое пространство имеет тривиальную фундаментальную группу. □

Мы видим, что любое стягиваемое пространство является односвязным. Обратное утверждение неверно.

Пример 11.3.6. Шар $\mathbb{B}^n$, $n \geq 3$, с выколотой точкой является односвязным, но не стягиваемым многообразием. □

Теорема 11.3.5. *Пусть* $\mathbb{M}_1$ *и* $\mathbb{M}_2$ *– два линейно связных топологических пространства. Тогда фундаментальная группа произведения* $\mathbb{M}_1 \times \mathbb{M}_2$ *изоморфна произведению фундаментальных групп,* $\pi(\mathbb{M}_1 \times \mathbb{M}_2) \simeq \pi(\mathbb{M}_1) \times \pi(\mathbb{M}_2)$.

Доказательство. Доказательство заключается в явном построении изоморфизма. См., например, [**?**]. □

Следствие. Произведение $\mathbb{M}_1 \times \mathbb{M}_2$ двух односвязных пространств является односвязным. □

Пример 11.3.7. Фундаментальная группа тора $\mathbb{T}^n \approx \mathbb{S}^1 \times \ldots \times \mathbb{S}^1$ изоморфна произведению групп целых чисел, $\pi(\mathbb{T}^n) \simeq \underbrace{\mathbb{Z} \times \ldots \times \mathbb{Z}}_{n}$. □

11.4 Фундаментальная группа и ориентируемость

Все связные многообразия можно разделить на два класса: ориентируемые и неориентируемые. Напомним, что многообразие $\mathbb{M}$ называется ориентируемым (см. раздел 3), если его можно покрыть системой карт, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, причем якобиан преобразования координат (2.58) во всех пересечениях $\mathbb{U}_i \cap \mathbb{U}_j$ положителен.

Рассмотрим класс многообразий, на которых можно задать репер $\{e_a\}$, $a = 1, \ldots, n$, т.е. n дифференцируемых векторных полей, которые линейно независимы в каждой точке многообразии $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$. Отсюда, в частности, следует, что векторные поля $e_a(x)$ не могут обращаться в нуль ни в какой точке многообразия $\mathbb{M}$, т.к. в противном случае они были бы линейно зависимы в этой точке. Репер $\{e_a\}$ образует базис для касательных векторных полей. Локально репер всегда существует. Однако его глобальное существование является сильным предположением, и далеко не каждое ориентируемое многообразие допускает существование репера.

Пример 11.4.1. Двумерная сфера $\mathbb{S}^2$ является ориентируемым многообразием, однако она не допускает существование репера, так как на ней не существует векторных полей, которые не обращаются в нуль (теорема 11.2.1). □

В локальной системе координат репер раскладывается по координатному базису, $e_a = e^\alpha{}_a \partial_\alpha$, $\det e^\alpha{}_a \neq 0$.

Если от отдельных реперов перейти к классам ориентирующих реперов, то можно дать новое эквивалентное определение ориентируемости многообразий. Пусть в точке $x \in \mathbb{M}$, $\dim \mathbb{M} = n$, задан *класс ориентации*, т.е. класс реперов $\{e_a\}$, $a = 1, \ldots, n$, связанных друг с другом линейным преобразованием с положительным определителем. Подробнее, два репера $\{e'_a\}$ и $\{e_a\}$ принадлежат одному классу ориентации, если они связаны преобразованием $e'_a = S_a{}^b e_b$, причем $\det S(x) > 0$. Если $\det S < 0$, то будем говорить, что в точке x реперы принадлежат классам противоположной ориентации. Таким образом, в каждой точке многообразия имеются два класса ориентации. Поскольку репер можно непрерывно смещать из точки x в близкие точки многообразия, то имеет смысл говорить о непрерывной зависимости классов ориентации от точки многообразия.

Определение. Если в каждой точке многообразия $\mathbb{M}$ существует класс ориентации реперов, который непрерывно зависит от точки многообразия, то многообразие $\mathbb{M}$ называется *ориентируемым*. □

Докажем эквивалентность нового определения и определения ориентируемости многообразий, которое было дано в разделе 3. Пусть многообразие ориентируемо и на нем существует координатное покрытие $\mathbb{M} = \bigcup_i \mathbb{U}_i$ такое, что якобиан преобразования координат положителен, $\det \partial_\alpha x^{\alpha'} > 0$, во всех областях пересечения карт. В каждой координатной окрестности выберем координатный репер $\{\partial_\alpha\}$. Множество этих реперов определяет класс ориентации, который непрерывно зависит от точки многообразия. Действительно, если точка принадлежит различным картам, то координатные реперы принадлежат одному классу ориентации. Обратно. Пусть существует класс ориентации реперов, непрерывно зависящий от точки многообразия. Выберем координатное покрытие многообразия, $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Тогда в двух пересекающихся картах $\mathbb{U}_i$ и $\mathbb{U}_j$ можно выбрать координаты x^α и $x^{\alpha'}$ таким образом, что

$$e_a = e^\alpha{}_a \partial_\alpha = e^{\alpha'}{}_a \partial_{\alpha'}, \qquad \det e^\alpha{}_a > 0, \quad \det e^{\alpha'}{}_a > 0,$$

где $\{e_a\}$ – некоторый представитель заданного класса ориентации. Действительно, если в какой то системе координат $\det e^\alpha{}_a < 0$, то, переставив две произвольные координаты, получим $\det e^\alpha{}_a > 0$. Отсюда следует положительность якобиана преобразования координат в области пересечения карт,

$$J = \det \partial_\alpha x^{\alpha'} = \det \left(e^a{}_\alpha e^{\alpha'}{}_a\right) > 0. \tag{11.7}$$

Это справедливо для любого представителя из класса ориентации.

На произвольном многообразии, в том числе и неориентируемом, ориентацию можно переносить вдоль путей. Пусть на многообразии $\mathbb{M}$ задан кусочно дифференцируемый путь $\gamma = x(t)$, $t \in [0, 1]$, и класс ориентации реперов $\{e_a\big(x(t)\big)\}$ во всех точках пути, который непрерывно зависит от точки пути. Тогда класс ориентации $\{e_a(1)\}$ в конечной точке $x(1)$ называется *переносом класса ориентации* репера $\{e_a(0)\}$ из точки $x(0)$ в точку $x(1)$ вдоль пути γ.

Операция переноса ориентации вдоль путей обладает следующими свойствами.

1) Из любой точки $x \in \mathbb{M}$ ориентацию можно однозначно перенести во все близлежащие точки многообразия вдоль путей, целиком лежащих в пределах координатной окрестности точки x.

2) Для любого кусочно дифференцируемого пути перенос ориентации существует и не зависит от выбора репера вдоль пути. Существование очевидно, т.к. достаточно задать произвольным образом компоненты n линейно независимых векторов $\{e_a(t)\}$, как дифференцируемые функции одного переменного. Независимость от выбора репера доказывается просто. Пусть $\{e_a(t)\}$ и $\{e'_a(t)\}$ – два репера вдоль одной кривой $x(t)$, имеющие одинаковую ориентацию при $t = 0$. Матрица перехода от одного репера к другому, $e'_a = S_a{}^b e_b$, невырождена, $\det S(t) \neq 0$ для всех $t \in [0, 1]$. Поэтому, если в начальный момент $\det S(0) > 0$, то из непрерывности следует, что $\det S(t) > 0$ для всех $t \in [0, 1]$, поскольку матрица S не может быть вырождена.

3) Если два пути $x_1(t)$ и $x_2(t)$ гомотопны относительно начала и конца, то переносы класса ориентации вдоль этих путей совпадают. Действительно, для гомотопных путей существует непрерывное отображение $F(t, s) : \ [0, 1] \times [0, 1] \to \mathbb{M}$ такое, что $F(t, 0) = x_1(t)$ и $F(t, 1) = x_2(t)$. Пусть $\{e_a(t)\}$ – репер вдоль $x_1(t)$. Перенесем каким либо непрерывным образом этот репер из каждой точки кривой γ_1 вдоль кривых $F(t, s)$ по второму аргументу s. Перенос ориентации в конечные точки не зависит от выбора репера $\{e_a(t)\}$ вдоль γ_1 в силу свойства 2). В результате получим некоторый репер вдоль второй кривой γ_2. Поскольку в начальной и конечной точке реперы по построению совпадают, то совпадают и переносы классов ориентации вдоль гомотопных путей.

Приведенные выше определения удобны для доказательства неориентируемости некоторых многообразий в силу следующей теоремы.

Теорема 11.4.1. *Связное многообразие ориентируемо тогда и только тогда, когда перенос класса ориентации репера вдоль любого замкнутого пути сохраняет ориентацию.*

Доказательство. Пусть многообразие $\mathbb{M}$ ориентируемо. Выберем координатное покрытие такое, чтобы якобиан преобразования координат был положителен, $\det \partial_\alpha x^{\alpha'} > 0$, во всех областях пересечения карт. Выберем произвольную точку $x_0 \in \mathbb{M}$ и зафиксируем какой либо репер, который принадлежал бы тому же классу ориентации, что и координатный базис, т.е. $\det \left(e^\alpha{}_a(x_0) \right) > 0$. Эту ориентацию можно перенести в любую точку многообразия, т.к. оно линейно связно. Результат переноса ориентации не зависит от пути, потому что ориентацию можно переносить внутри произвольной карты, а при переходе от карты к карте принадлежность репера определенному классу ориентации сохраняется. В частности, принадлежность репера данному классу ориентации сохраняется при переносе ориентации вдоль произвольного замкнутого пути.

Обратно. Выберем произвольным образом репер в какой либо точке $x_0 \in \mathbb{M}$ и разнесем эту ориентацию во все другие точки $x \in \mathbb{M}$. Результат будет однозначен, т.к. перенос класса ориентации вдоль произвольного замкнутого пути сохраняет ориентацию. Таким образом в каждой точке $x \in \mathbb{M}$ определен класс ориентации $\{e_a(x)\}$. Пусть $\mathbb{U}_i$ – координатное покрытие многообразия $\mathbb{M}$. На каждой координатной окрестности можно выбрать систему координат такую, что $\det e^\alpha{}_a > 0$ для некоторого представителя из класса ориентации. Следовательно, из (11.7) следует, что якобиан преобразования координат во всех пересекающихся областях будет положителен. □

Из этой теоремы вытекает, что если можно указать замкнутый путь, вдоль которого не существует переноса класса ориентации репера, то многообразие является неориентируемым. В частности, неориентируемое многообразие не может быть односвязным. Действительно, любой замкнутый путь на односвязном многообразии можно непрерывно стянуть в точку. Поскольку замкнутые пути в достаточно малой окрестности произвольной точки сохраняют ориентацию, то из непрерывности следует, что на односвязном многообразии любой замкнутый путь сохраняет ориентацию. То есть каждая односвязная компонента многообразия всегда ориентируема. Неодносвязные многообразия могут быть как ориентируемыми, так и неориентируемыми.

Пример 11.4.2. Цилиндр и лист Мебиуса имеют изоморфные фундаментальные группы $\pi \simeq \mathbb{Z}$, однако являются соответственно ориентируемой и неориентируемой поверхностями. □

Пусть на многообразии $\mathbb{M}$ задан ориентирующий репер $\{e_a\}$ глобально. Рассмотрим этот репер в качестве базиса касательных векторных полей $X = X^a e_a \in \mathcal{X}(\mathbb{M})$. Зададим на $\mathbb{M}$ тривиальную линейную $\mathbb{GL}(n,\mathbb{R})$-связность: $\omega_{\alpha a}{}^b = 0$ в каждой карте. Это значит, что при параллельном переносе векторов их компоненты относительно выбранного репера не меняются. Тогда рассматриваемое многообразие будет пространством абсолютного параллелизма, т.к. тензор кривизны для тривиальной связности равен нулю, $R_{\alpha\beta a}{}^b = 0$. По этой причине, если на многообразии можно задать репер глобально, то оно называется *параллелизуемым*. В то же время тензор кручения для тривиальной связности в общем случае будет отличен от нуля. При этом тензор кручения на односвязных многообразиях для заданной нулевой связности однозначно определяется репером. Верно также и обратное утверждение: в односвязном пространстве абсолютного параллелизма, в котором тензор кривизны всюду равен нулю, $R_{\alpha\beta a}{}^b = 0$, всегда можно задать репер глобально. Для этого достаточно выбрать n линейно независимых векторов в касательном пространстве произвольной точки, а затем разнести их по всему многообразию, используя то свойство, что результат не зависит от пути параллельного переноса.

Очевидно, что любое параллелизуемое многообразие является ориентируемым.

Пример 11.4.3. Любая группа Ли является параллелизуемым многообразием, если групповое умножение слева (справа) принять за определение параллельного переноса (см. раздел 9.6). Поэтому все группы Ли представляют собой ориентируемые многообразия. При этом левоинвариантные (правоинвариантные) векторные поля представляют собой глобально определенный репер. □

Обратное утверждение, что все ориентируемые многообразия являются параллелизуемыми, неверно, как показывает пример сферы 11.4.1.

Приведем критерий параллелизуемости многообразий.

Теорема 11.4.2. *Многообразие $\mathbb{M}$ является параллелизуемым тогда и только тогда, когда касательное расслоение тривиально, т.е. диффеоморфно прямому произведению $\mathbb{T}(\mathbb{M}) \approx \mathbb{M} \times \mathbb{R}^n$.*

Доказательство. Проверяется тот факт, что глобально определенный репер устанавливает диффеоморфизм расслоений. □

Ориентируемость многообразий можно также определить с помощью дифференциальных форм, рассмотренных в разделе 4. Рассмотрим связное многообразие $\mathbb{M}$,

$\dim \mathbb{M} = n$. Пространство дифференциальных форм максимальной степени $\Lambda_n(\mathbb{M})$ в каждой точке многообразия одномерно. Пусть $A, B \in \Lambda_n(\mathbb{M})$ и $A \neq 0$, $B \neq 0$ во всех точках $x \in \mathbb{M}$. Отношение $A = f(x)B$, где $f \in \mathcal{C}(\mathbb{M})$ и $f > 0$, задает отношение эквивалентности в пространстве n-форм, не обращающихся в нуль ни в одной точке многообразия. Это отношение определяет ровно два класса эквивалентности. Назовем *ориентацией* многообразия $\mathbb{M}$ соответствующий класс эквивалентности n-форм, если такой класс существует. Это определение ориентируемости эквивалентно приведенному выше в силу следующей теоремы.

Теорема 11.4.3. *Связное многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, ориентируемо тогда и только тогда, когда оно допускает непрерывную нигде не обращающуюся в нуль n-форму.*

Доказательство. Достаточность. Пусть на связном многообразии $\mathbb{M}$ ориентация задана с помощью n-формы A, которая не равна нулю ни в одной точке. Пусть $\{\mathbb{U}_i\}$ – локально конечное покрытие многообразия $\mathbb{M}$. Пусть x^α, $\alpha = 1, \ldots, n$ – система координат на некоторой окрестности $\mathbb{U}_i$. Тогда n-форма $dx^1 \wedge \ldots \wedge dx^n$ не равна нулю ни в какой точке $\mathbb{U}_i$ и поэтому отличается от A на некоторый отличный от нуля множитель $f(x)$,

$$dx^1 \wedge \ldots \wedge dx^n = fA.$$

Если $f > 0$, то мы говорим, что система координат x^α согласована с ориентацией на $\mathbb{M}$. В противном случае, при $f < 0$, достаточно переставить две координаты для изменения знака n-формы. Обозначим координаты на области $\mathbb{U}_j$, которая пересекается с $\mathbb{U}_i$, через $x^{\alpha'}$. Эти координаты всегда можно выбрать так, чтобы они были согласованы с ориентацией A. Тогда n-форма $dx^{1'} \wedge \ldots \wedge dx^{n'}$ также отличается от формы A на некоторый положительный множитель $g(x) > 0$,

$$dx^{1'} \wedge \ldots \wedge dx^{n'} = gA.$$

Поскольку

$$dx^{1'} \wedge \ldots \wedge dx^{n'} = \det\left(\frac{\partial x^{\alpha'}}{\partial x^\alpha}\right) dx^1 \wedge \ldots \wedge dx^n,$$

то якобиан соответствующего преобразования координат $x^\alpha \to x^{\alpha'}$ положителен,

$$\det\left(\frac{\partial x^{\alpha'}}{\partial x^\alpha}\right) = gf^{-1} > 0.$$

Полученное выражение для якобиана преобразования координат корректно, что легко проверяется для пересечения трех карт.

Необходимость. Пусть f_i – разбиение единицы, подчиненное заданному счетному локально конечному покрытию $\{\mathbb{U}_i\}$ (теорема 3.2.3). Выберем координаты на каждой области $\mathbb{U}_i$ таким образом, чтобы все якобианы преобразований координат в областях пересечения были положительны. На каждой области $\mathbb{U}_i$ можно построить отличную от нуля n-форму $\omega_i = dx^1 \wedge \ldots \wedge dx^n$. Тогда n-форма $\sum_i f_i \omega_i$ определена на всем многообразии и отлична от нуля во всех точках. □

Замечание. В доказательстве теоремы использовано разбиение единицы, которое всегда существует, т.к. в определение многообразия мы включили счетность базы топологии. Если это требование исключено из определения, как это часто делается, то в условии теоремы необходимо дополнительно потребовать, например, паракомпактность многообразия. □

В роли формы, которая не обращается в нуль, может выступать форма объема (4.72). Если многообразие не допускает существования непрерывной невырожденной n-формы, то оно будет неориентируемым. Если на связном многообразии задана невырожденная n-форма, то мы говорим, что на многообразии задана ориентация. Если многообразие несвязно и состоит из нескольких компонент, то на каждой компоненте ориентацию можно задавать независимо, если такое возможно.

Отметим, что каждый репер e_a, заданный глобально, определяет невырожденную n-форму. Действительно, обозначим дуальный базис кокасательного пространства через e^a. Тогда форма объема

$$\upsilon = e^1 \wedge \ldots \wedge e^n$$

определена глобально и невырождена. Обратное утверждение неверно.

Пример 11.4.4. На двумерной сфере $\mathbb{S}^2$ существует форма объема, однако репер определить глобально нельзя, так как на ней не существует векторных полей, не обращающихся в нуль. □

Теорема 11.4.4. *Для любого неориентируемого многообразия $\mathbb{M}$ существует двулистное ориентируемое накрытие $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$.*

Доказательство. См., например, [**?**]. □

Пример 11.4.5. Сфера $\mathbb{S}^2$ ориентируема и является двулистным универсальным накрывающим пространством для проективной плоскости $\mathbb{RP}^2$, которая неориентируема.

Пример 11.4.6. Лист Мебиуса можно получить из цилиндра путем отождествления центрально симметричных точек, как показано на рис. 11.6. Это накрытие двулистно, но не является универсальным. Универсальные накрывающие цилиндра и листа Мебиуса совпадают – это плоскость $\mathbb{R}^2$. □

Рис. 11.6: Накрытие листа Мебиуса цилиндром. Отождествив центрально симметричные точки цилиндра, мы получим лист Мебиуса в виде полосы с отождествленными краевыми точками.

Обсудим более подробно связь неориентируемости многообразия с фундаментальной группой. Перенос ориентации с помощью репера показывает, что каждый гомотопический класс замкнутого пути с началом и концом в точке x_0 (т.е. элемент фундаментальной группы $\pi(\mathbb{M}, x_0)$) либо сохраняет, либо меняет ориентацию репера на противоположную. Это означает, что существует гомоморфизм σ фундаментальной группы в группу $\mathbb{Z}_2$,

$$\sigma:\quad \pi(\mathbb{M}, x_0) \to \mathbb{Z}_2 = \{\pm 1\}\,.$$

Другими словами, каждому замкнутому пути ставится в соответствие $+1$ или -1, в зависимости от того, сохраняется ориентация репера при переносе или меняется на противоположную. Если многообразие ориентируемо, то гомоморфизм σ тривиален. Для неориентируемых многообразий ввиду наличия путей, обращающих ориентацию, гомоморфизм σ нетривиален. Как следствие получаем, что фундаментальная группа неориентируемых многообразий не может быть тривиальной.

Пример 11.4.7. Фундаментальная группа листа Мебиуса совпадает с группой целых чисел, $\pi(\mathbb{M}) = \mathbb{Z}$, т.к. он стягивается к центральной окружности. При гомоморфизме σ все четные и нечетные числа отображаются соответственно в $+1$ и -1. □

Пример 11.4.8. Фундаментальная группа проективного пространства $\mathbb{RP}^n$ нетривиальна: $\pi(\mathbb{RP}^n) = \mathbb{Z}_2$, и гомоморфизм σ является изоморфизмом для четных n. Для нечетных n гомоморфизм σ тривиален, т.к. проективные пространства нечетной размерности ориентируемы. Проективное пространство можно параметризовать точками шара, у которого отождествлены диаметрально противоположные точки граничной сферы (см. пример 3.1.8). Тогда диаметр шара будет замкнутым путем (образом окружности), который не стягивается в точку. В то же время, если при отображении окружности в проективное пространство диаметр проходится два раза, то такой образ окружности можно стянуть в точку, как показано на рис. 11.7. Аналогично стягивается в точку любой образ окружности, который проходит диа-

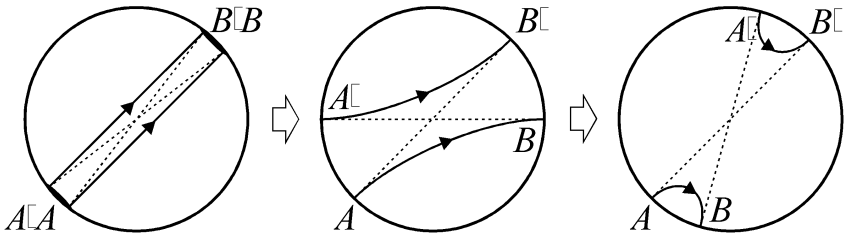


Рис. 11.7: Непрерывная деформация образа окружности в проективном пространстве $\mathbb{RP}^2$, который проходит по диаметру AB два раза, в точку. Пунктиром показано отождествление противоположных точек граничной окружности.

метр произвольное четное число раз. Это поясняет, почему фундаментальная группа проективного пространства равна $\mathbb{Z}_2$. □

Глава 12

Накрытия

Важным классом непрерывных отображений топологических пространств являются накрытия, которые представляют собой локальный гомеоморфизм.

12.1 Определения и примеры

Все, сказанное в настоящей главе о топологических пространствах, относится также к многообразиям.

Определение. Непрерывное отображение топологических пространств $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ называется *накрытием*, если выполнены следующие условия:

1) p сюръективно;
2) для любого $x \in \mathbb{M}$ найдется открытая окрестность $\mathbb{U}_x$ точки x такая, что $p^{-1}(\mathbb{U}_x) = \bigcup_{j\in\mathcal{J}} \mathbb{U}_j$ для некоторого семейства открытых подмножеств $\mathbb{U}_j \subset \widetilde{\mathbb{M}}$, удовлетворяющих условиям $\mathbb{U}_j \bigcap \mathbb{U}_k = \emptyset$ при $j \neq k$ и сужение отображения $p|_{\mathbb{U}_j}:\ \mathbb{U}_j \to \mathbb{U}_x$ – гомеоморфизм для всех $j \in \mathcal{J}$.

Мы говорим также, что окрестность $\mathbb{U}_x$ *просто накрыта* отображением p. Топологическое пространство $\mathbb{M}$ называется *базой* накрытия, а $\widetilde{\mathbb{M}}$ – *накрывающим пространством*. Каждое множество $\mathbb{U}_j$ называется *листом* накрытия области $\mathbb{U}_x$. Если множество $\mathcal{J}$ содержит n элементов, то говорят об n-листном накрытии. Если топологическое пространство $\widetilde{\mathbb{M}}$ является односвязным, т.е. фундаментальная группа тривиальна, $\pi(\widetilde{\mathbb{M}}) = e$, то накрытие называется *универсальным*. □

Предложение 12.1.1. *Если база $\mathbb{M}$ накрытия $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ линейно связна, то мощность множества $\mathcal{J}$ не зависит от точки $x \in \mathbb{M}$.*

Доказательство. См., например, [**?**]. □

Это утверждение говорит о том, что определение n-листного накрытия корректно, т.к. не зависит от точки $x \in \mathbb{M}$.

Замечание. Если $\widetilde{\mathbb{M}}$ и $\mathbb{M}$ являются не просто топологическими пространствами, а многообразиями, то вместо непрерывности отображения p мы требуем его дифференцируемость, а гомеоморфизм заменяем на диффеоморфизм областей $\mathbb{U}_j$ и $\mathbb{U}_x$ для всех $j \in \mathcal{J}$. Для многообразий $\mathsf{dim}\,\mathbb{U}_j = \mathsf{dim}\,\mathbb{U}_x$ для всех $j \in \mathcal{J}$, т.к. сужение отображения p на $\mathbb{U}_j$ – диффеоморфизм. Кроме того, каждое подмножество $\mathbb{U}_j$, по-определению, открыто в $\widetilde{\mathbb{M}}$. Это возможно только если $\mathsf{dim}\,\widetilde{\mathbb{M}} = \mathsf{dim}\,\mathbb{U}_j$. Таким образом, если $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ накрытие многообразий, то необходимо, чтобы накрывающее многообразие и база имели одинаковую размерность, $\mathsf{dim}\,\widetilde{\mathbb{M}} = \mathsf{dim}\,\mathbb{M}$. Для многообразий отображение p является погружением. □

Пример 12.1.1. Всякий гомеоморфизм является однолистным накрытием. В этом случае у каждой области $\mathbb{U}_x$ имеется только один лист. □

Пример 12.1.2. Если $\widetilde{\mathbb{M}} = \mathbb{M} \times \mathbb{X}$ – топологическое произведение топологического пространства $\mathbb{M}$ на счетное множество $\mathbb{X}$, то естественная проекция $p: \mathbb{M} \times \mathbb{X} \to \mathbb{M}$ является накрытием. Число листов равно числу элементов множества $\mathbb{X}$. □

Пример 12.1.3. Пусть задана аналитическая функция $F(z)$ в некоторой области $\mathbb{D} \subset \mathbb{C}$. Если в каждой точке $z \in \mathbb{D}$ функция $F(z)$ имеет N значений, то риманова поверхность функции $F(z)$ является N-листным накрытием области $\mathbb{D}$ (см. раздел **??**). □

Пример 12.1.4. Отображение вещественной прямой в окружность,

$$p: \quad \mathbb{R} \ni t \mapsto \mathsf{e}^{2\pi i t} \in \mathbb{S}^1,$$

является накрытием с бесконечным числом листов. Это накрытие является универсальным. □

Пример 12.1.5. Пусть окружность $\mathbb{S}^1$ задана на комплексной плоскости $\mathbb{C}$ уравнением $|z| = 1$. Тогда отображение окружности в окружность,

$$p: \quad \mathbb{S}^1 \ni z \mapsto z^n \in \mathbb{S}^1, \qquad \forall n \in \mathbb{N},$$

является накрытием с n листами. Это же отображение является n-листным накрытием комплексной плоскости с выколотым началом координат $p: \mathbb{C}\setminus\{0\} \to \mathbb{C}\setminus\{0\}$. □

Пример 12.1.6. Отображение сферы в проективное пространство $p: \mathbb{S}^n \to \mathbb{RP}^n$, $n \geq 1$, получающееся отождествлением диаметрально противоположных точек, является двулистным накрытием. Это накрытие является универсальным. □

Пример 12.1.7. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H}$ – ее дискретная подгруппа. Тогда проекция $p: \mathbb{G} \to \mathbb{G}/\mathbb{H}$, где $\mathbb{G}/\mathbb{H}$ – пространство левых или правых смежных классов, является накрытием. Число листов равно числу элементов группы $\mathbb{H}$. □

Предложение 12.1.2. *Пусть $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ расслоение, определенное в разделе 3.4, типичным слоем $\mathbb{F}$ которого является 0-мерное многообразие. Тогда проекция $\pi: \mathbb{E} \to \mathbb{M}$ является накрытием. Наоборот, любое накрытие является расслоением $\widetilde{\mathbb{M}}(\mathbb{M}, p, \mathcal{J})$ с 0-мерным типичным слоем $\mathcal{J}$.* □

Доказательство. Определение накрытия и расслоения в случае 0-мерного типичного слоя совпадают, если записать $\pi^{-1}(\mathbb{U}_x) = \chi^{-1}(\mathbb{U}_x \times \mathbb{F}) = \bigcup_{j\in\mathcal{J}} \mathbb{U}_j$, где точки 0-мерного многообразия $j \in \mathcal{J} = \mathbb{F}$ нумеруют листы накрытия области $\mathbb{U}_x$. Второе условие в определении расслоения при 0-мерном слое выполняется автоматически, т.к. отображение $\pi \circ \chi^{-1}$ дифференцируемо, а это возможно только при проекции на первый сомножитель, $\pi \circ \chi^{-1}: \mathbb{U}_x \times \mathbb{F} \ni (y, f) \mapsto y \in \mathbb{M}$. □

Замечание. Если типичным слоем расслоения $\mathbb{E}(\mathbb{M}, \pi, \mathbb{F})$ является многообразие более высокой размерности, $\dim \mathbb{F} \geq 1$, то расслоение не является накрытием, потому что $\dim \pi^{-1}(\mathbb{U}_x) = \dim \mathbb{M} + \dim \mathbb{F} > \dim \mathbb{M}$. □

Следующее утверждение облегчает изучение накрытий за счет ограничения класса рассматриваемых баз.

Теорема 12.1.1. *Если пространство* $\mathbb{M}$ *– локально связно, то непрерывное отображение* $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ *тогда и только тогда является накрытием, когда для каждой компоненты связности* $\mathbb{U} \subset \mathbb{M}$ *сужение отображения*

$$p|_{p^{-1}(\mathbb{U})}:\quad p^{-1}(\mathbb{U}) \to \mathbb{U}$$

является накрытием.

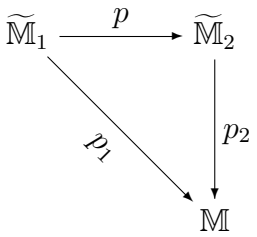
Доказательство. См., например, [**?**]. □

Поскольку все многообразия являются локально связными, то в дифференциальной геометрии мы можем ограничиться рассмотрением накрытий над связными базами $\mathbb{M}$. Поэтому в дальнейшем мы будем рассматривать только линейно связные и, следовательно, связные базы накрытий.

Из определения накрытия следует, что прообраз $p^{-1}(\mathbb{U}_x)$ является объединением непересекающихся открытых подмножеств $\mathbb{U}_j$, каждое из которых гомеоморфно $\mathbb{U}_x$. Если $\mathbb{U}_x$ связно, то каждое из подмножеств $\mathbb{U}_j$ также связно. Отсюда следует

Предложение 12.1.3. *Пусть* $\mathbb{U}$ *– открытое связное подмножество пространства* $\mathbb{M}$*, просто накрытое отображением* $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$*. Тогда* p *гомеоморфно отображает каждую компоненту связности множества* $p^{-1}(\mathbb{U})$ *на* $\mathbb{U}$*.*

Следствие. Рассмотрим два накрытия $p_1:\ \widetilde{\mathbb{M}}_1 \to \mathbb{M}$, $p_2:\ \widetilde{\mathbb{M}}_2 \to \mathbb{M}$ с одинаковыми базами и коммутативную диаграмму



где база $\mathbb{M}$ локально связна. Если p сюръективно, то оно также является накрытием. □

Теорема 12.1.2. *Пусть* $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ *– накрытие. Тогда*

1) p *– открытое отображение;*

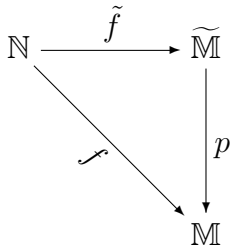
2) $\mathbb{M}$ *имеет фактор топологию относительно* p*.*

Доказательство. См., например, [**?**]. □

Если на базе накрытия $\mathbb{M}$ задан путь $\gamma:\ [0,1] \ni t \mapsto x(t) \in \mathbb{M}$, то возникает вопрос, что представляет собой прообраз $p^{-1}\Big(\gamma([0,1])\Big)$ при накрытии $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$? Для ответа введем новое понятие.

Пусть задано накрытие $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ и непрерывное отображение некоторого пространства $\mathbb{N}$ в базу накрытия, $f:\ \mathbb{N} \to \mathbb{M}$. (Для путей в роли $\mathbb{N}$ выступает единичный отрезок $[0,1]$.)

Определение. *Поднятием* непрерывного отображения f называется непрерывное отображение $\tilde{f}:\ \mathbb{N} \to \widetilde{\mathbb{M}}$ такое, что $p \circ \tilde{f} = f$, т.е. диаграмма



коммутативна. □

Следующий результат показывает, что, если поднятие отображения f существует, то оно, по-существу, единственно.

Предложение 12.1.4. *Пусть $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ – накрытие и $\tilde{f}, \tilde{\tilde{f}}:\ \mathbb{N} \to \widetilde{\mathbb{M}}$ – два поднятия отображения $f:\ \mathbb{N} \to \mathbb{M}$. Если $\mathbb{N}$ связно и поднятия совпадают $\tilde{f}(y) = \tilde{\tilde{f}}(y)$ хотя бы в одной точке $y \in \mathbb{N}$, то поднятия совпадают во всех точках, $\tilde{f} = \tilde{\tilde{f}}$.*

Доказательство. См., например, [**?**]. □

Теорема 12.1.3 (О накрывающей гомотопии для путей). *Пусть $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ – накрытие. Тогда*

1) *для пути $\gamma:\ [0,1] \to \mathbb{M}$ и точки $\tilde{x} \in \widetilde{\mathbb{M}}$ такой, что $p(\tilde{x}) = \gamma(0)$ существует единственный путь $\tilde{\gamma}:\ [0,1] \to \widetilde{\mathbb{M}}$, который является поднятием, $p \circ \tilde{\gamma} = \gamma$, пути γ и для которого $\tilde{\gamma}(0) = \tilde{x}$;*
2) *для непрерывного отображения (гомотопии двух путей в базе) $F:\ [0,1] \times [0,1] \to \mathbb{M}$ и точки $\tilde{x} \in \widetilde{\mathbb{M}}$ такой, что $p(\tilde{x}) = F(0,0)$ существует единственное непрерывное отображение (гомотопия двух путей в накрывающем пространстве) $\widetilde{F}:\ [0,1] \times [0,1] \to \widetilde{\mathbb{M}}$, которое является поднятием, $p \circ \widetilde{F} = F$, отображения F и для которого $\widetilde{F}(0,0) = \tilde{x}$.*

Доказательство. Теорема доказывается путем явного построения отображений. Детали можно найти, например, в [**?**]. □

Следствие (Теорема о монодромии). Пусть $\gamma_1 \sim \gamma_2(\ \mathsf{rel}\{0,1\})$ – два пути в базе $\mathbb{M}$, гомотопные относительно начала и конца, и начальные точки их поднятий совпадают, $\tilde{\gamma}_0(0) = \tilde{\gamma}_1(0)$. Тогда конечные точки поднятий также совпадают, $\tilde{\gamma}_0(1) = \tilde{\gamma}_1(1)$. □

Напомним, что единица фундаментальной группы $\pi(\mathbb{M}, x_0)$ – это класс замкнутых путей с началом и концом в точке x_0, которые гомотопны тождественному пути e_{x_0}. Из теоремы о монодромии следует, что поднятия всех этих путей также замкнуты в накрывающем пространстве $\widetilde{\mathbb{M}}$. Следующая теорема говорит о том, что поднятие замкнутых путей, которые не стягиваются в точку x_0 будет незамкнутым.

Теорема 12.1.4. *Пусть $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ – универсальное накрытие, т.е. накрывающее пространство $\widetilde{\mathbb{M}}$ – односвязно. Тогда существует взаимно однозначное соответствие между элементами фундаментальной группы $\pi\Big(\mathbb{M}, p(\tilde{x})\Big)$ и множеством прообразов $p^{-1}\Big(p(\tilde{x})\Big)$, где $\tilde{x}$ – произвольная точка $\widetilde{\mathbb{M}}$.*

Доказательство. Пусть $[\gamma] \in \pi\Big(\mathbb{M}, p(\tilde{x})\Big)$ – класс гомотопных замкнутых путей на базе и $\tilde{\gamma}$ – поднятие какого либо представителя из данного класса. Определим отображение

$$\varphi:\quad \pi\Big(\mathbb{M}, p(\tilde{x})\Big) \ni [\gamma] \mapsto \varphi([\gamma]) = \tilde{\gamma}(1) \in p^{-1}\Big(p(\tilde{x})\Big).$$

Это отображение определено корректно согласно теореме о монодромии.

Теперь определим обратное отображение $\varphi^{-1}:\ p^{-1}\Big(p(\tilde{x})\Big) \to \pi\Big(\mathbb{M}, p(\tilde{x})\Big)$. Выбираем точку $\tilde{x}_1 \in p^{-1}\Big(p(\tilde{x})\Big)$ и некоторый путь $\tilde{\gamma}$ из $\tilde{x}$ в $\tilde{x}_1$. Поскольку $\widetilde{\mathbb{M}}$ односвязно, то любые такие пути гомотопны. Поэтому класс замкнутых путей $[p \circ \tilde{\gamma}]$ – корректно определенный элемент фундаментальной группы $\pi\Big(\mathbb{M}, p(\tilde{x})\Big)$. Легко проверить, что $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = e$, так что φ и φ^{-1} – биекции. □

Эта теорема дает возможность в некоторых случаях вычислить фундаментальную группу. Для этого надо найти универсальное накрытие $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ и групповую структуру на прообразе $p^{-1}(x_0)$, где $x_0 \in \mathbb{M}$, так, чтобы биекция $\varphi:\ \pi(\mathbb{M}, x_0) \to p^{-1}(x_0)$ стала гомоморфизмом групп. Вообще говоря, сделать это сложно.

12.2 Фундаментальная группа пространства орбит

В настоящем разделе мы рассмотрим группы преобразований $(\widetilde{\mathbb{M}}, \mathbb{G})$ и приведем достаточные условия на группу преобразований, которые обеспечивают то, что проекция пространства $\widetilde{\mathbb{M}}$ в пространство орбит $\mathbb{M} = \widetilde{\mathbb{M}}/\mathbb{G}$ является накрытием. Обсудим связь группы преобразований $\mathbb{G}$ с фундаментальными группами базы $\mathbb{M} = \widetilde{\mathbb{M}}/\mathbb{G}$ и накрывающего пространства $\widetilde{\mathbb{M}}$. Кроме того, обсудим достаточные условия существования в пространстве орбит дифференцируемой структуры.

Замечание. Многие определения и теоремы настоящего раздела справедливы не только для многообразий, но и для произвольных топологических пространств. Поэтому мы будем употреблять термин топологическое пространство, имея в виду, что соответствующее утверждение верно и для многообразий. Там, где важна специфика дифференцируемой структуры, мы будем писать многообразие. □

Определение. Группа $\mathbb{G}$ действует на топологическом пространстве $\widetilde{\mathbb{M}}$ *собственно разрывно*, если выполнены следующие условия:

1) для любых двух точек $\tilde{x}_1$ и $\tilde{x}_2$ из $\widetilde{\mathbb{M}}$, не лежащих на одной орбите, существуют окрестности $\widetilde{\mathbb{U}}_1 \supset \tilde{x}_1$ и $\widetilde{\mathbb{U}}_2 \supset \tilde{x}_2$ такие, что $\widetilde{\mathbb{U}}_1\mathbb{G} \bigcap \widetilde{\mathbb{U}}_2 = \emptyset$,
2) группа изотропии $\mathbb{G}_{\tilde{x}}$ любой точки $\tilde{x} \in \widetilde{\mathbb{M}}$ конечна,
3) для каждой точки $\tilde{x}$ существует окрестность $\widetilde{\mathbb{U}}$, устойчивая относительно группы изотропии $\mathbb{G}_{\tilde{x}}$, $\widetilde{\mathbb{U}}\mathbb{G}_{\tilde{x}} = \widetilde{\mathbb{U}}$, и такая, что для всех $a \in \mathbb{G}$, не лежащих в $\mathbb{G}_{\tilde{x}}$, пересечение $\widetilde{\mathbb{U}}a \bigcap \widetilde{\mathbb{U}} = \emptyset$.

Предложение 12.2.1. *Если группа $\mathbb{G}$ действует на многообразии $\widetilde{\mathbb{M}}$ собственно разрывно, то она конечна или счетна.*

Доказательство. В силу условия 3) орбита произвольной точки является дискретным подмножеством многообразия. Поскольку множество таких точек на многообразии счетно, то, вместе с условием 2) это влечет счетность элементов группы $\mathbb{G}$. В частности, оно может быть конечно. □

Для доказательства следующей теоремы понадобится лемма, которая представляет и самостоятельный интерес.

Лемма 12.2.1. *Пусть $(\widetilde{\mathbb{M}}, \mathbb{G})$ – группа преобразований топологического пространства $\widetilde{\mathbb{M}}$ и фактор пространство $\mathbb{M} = \widetilde{\mathbb{M}}/\mathbb{G}$ снабжено фактор топологией. Тогда каноническая проекция $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ является открытым отображением.*

Доказательство. Пусть $\widetilde{\mathbb{U}}$ открытое подмножество $\widetilde{\mathbb{M}}$. Тогда подмножество $p(\widetilde{\mathbb{U}})$ открыто в $\widetilde{\mathbb{M}}/\mathbb{G}$, т.к. оно снабжено фактор топологией. Прообраз этого подмножества можно представить в виде

$$p^{-1}\Big(p(\widetilde{\mathbb{U}})\Big) = \bigcup_{a\in\mathbb{G}} \widetilde{\mathbb{U}}a. \tag{12.1}$$

Поскольку отображение $\widetilde{\mathbb{U}} \to \widetilde{\mathbb{U}}a$ – гомеоморфизм для всех $a \in \mathbb{G}$, то $p^{-1}\Big(p(\widetilde{\mathbb{U}})\Big)$ открыто в $\widetilde{\mathbb{M}}$ как объединение открытых подмножеств. □

Теорема 12.2.1. *Пусть группа $\mathbb{G}$ действует в топологическом пространстве $\widetilde{\mathbb{M}}$ собственно разрывно и свободно, а пространство орбит $\widetilde{\mathbb{M}}/\mathbb{G}$ снабжено фактор топологией. Тогда каноническая проекция $p:\ \widetilde{\mathbb{M}} \to \widetilde{\mathbb{M}}/\mathbb{G}$ является накрытием.*

Доказательство. По построению отображение p сюръективно и непрерывно. По лемме 12.2.1 это отображение открыто. Пусть $\widetilde{\mathbb{U}}$ – окрестность точки $\tilde{x} \in \widetilde{\mathbb{M}}$, удовлетворяющая условию 3) в определении собственно разрывной группы преобразований. Так как p – открытое отображение, то $p(\widetilde{\mathbb{U}})$ – открытая окрестность орбиты $p(\tilde{x}) = \tilde{x}\mathbb{G}$ и ее прообраз имеет вид (12.1). Поскольку действие группы свободно, то все области $\widetilde{\mathbb{U}}a$, $a \in \mathbb{G}$, открыты в $\widetilde{\mathbb{M}}$ и не пересекаются. Сужение отображения на каждую подобласть $p|_{\widetilde{\mathbb{U}}a}:\ \widetilde{\mathbb{U}}a \to p(\widetilde{\mathbb{U}})$ является непрерывным открытым биективным отображением и, следовательно, гомеоморфизмом. □

Теперь обсудим связь между группой преобразований $(\widetilde{\mathbb{M}}, \mathbb{G})$, действующей собственно разрывно и свободно, и фундаментальной группой пространства орбит $\mathbb{M} = \widetilde{\mathbb{M}}/\mathbb{G}$, на котором определена фактор топология. Пусть $\tilde{x} \in \widetilde{\mathbb{M}}$ и $x = p(\tilde{x}) \in \widetilde{\mathbb{M}}/\mathbb{G}$. Заметим, что

$$p^{-1}(x) = \{\tilde{x}a \in \widetilde{\mathbb{M}}:\quad a \in \mathbb{G}\},$$

т.е. точки орбиты находятся во взаимно однозначном соответствии с элементами группы. Если класс относительно гомотопных путей принадлежит фундаментальной группе пространства орбит, $[\gamma] \in \pi(\widetilde{\mathbb{M}}/\mathbb{G}, x)$, то существует единственное поднятие $\tilde{\gamma}$ пути γ с началом в точке $\tilde{x} \in \widetilde{\mathbb{M}}$ (теорема 12.1.3). Конец поднятого пути $\tilde{x}(1) \in p^{-1}(x)$, и, поскольку группа преобразований действует свободно, то существует единственный элемент $a_{(\gamma)} \in \mathbb{G}$ такой, что $\tilde{x}(1) = \tilde{x}a_{(\gamma)}$. Следовательно, определено отображение

$$\varphi:\quad \pi(\widetilde{\mathbb{M}}/\mathbb{G}, x) \ni [\gamma] \mapsto a_{(\gamma)} \in \mathbb{G}. \tag{12.2}$$

Теорема 12.2.2. *Отображение (12.2) является гомоморфизмом групп.*

Доказательство. Доказательство сводится к проверке равенства $\varphi([\gamma_1][\gamma_2]) = \varphi([\gamma_1])\varphi([\gamma_2])$. Детали можно найти, например, в [**?**]. □

Если фундаментальная группа накрывающего пространства нетривиальна, то ядро гомоморфизма также нетривиально.

Предложение 12.2.2. *Ядро гомоморфизма (12.2) совпадает с индуцированной подгруппой $p_*\pi(\widetilde{\mathbb{M}}, \tilde{x}) \subset \pi(\widetilde{\mathbb{M}}/\mathbb{G}, x)$.*

Доказательство. См., например, [**?**]. □

В частности, $p_*\pi(\widetilde{\mathbb{M}}, \tilde{x})$ – нормальная подгруппа в $\pi(\widetilde{\mathbb{M}}/\mathbb{G}, x)$, и поэтому определена фактор группа $\pi(\widetilde{\mathbb{M}}/\mathbb{G}, x)/p_*\pi(\widetilde{\mathbb{M}}, \tilde{x})$.

Теорема 12.2.3. *Группы $\pi(\widetilde{\mathbb{M}}/\mathbb{G}, x)/p_*\pi(\widetilde{\mathbb{M}}, \tilde{x})$ и $\mathbb{G}$ изоморфны.*

Доказательство. См., например, [**?**]. □

Следствие. Если накрывающее топологическое пространство $\widetilde{\mathbb{M}}$ односвязно, то $\pi(\widetilde{\mathbb{M}}/\mathbb{G}, x) \simeq \mathbb{G}$. □

Это следствие позволяет в ряде случаев найти фундаментальную группу многообразия.

Пример 12.2.1. Нетрудно проверить, что окружность $\mathbb{S}^1$ гомеоморфна пространству орбит $\mathbb{R}/\mathbb{Z}$, где группа $\mathbb{Z}$ действует на $\mathbb{R}$ сдвигами на постоянное число. Действие группы $\mathbb{Z}$ на $\mathbb{R}$ свободно и собственно разрывно. Поскольку вещественная прямая $\mathbb{R}$ односвязна, то отображение $p:\ \mathbb{R}\to\mathbb{R}/\mathbb{Z}$ – универсальное накрытие. Следовательно $\pi(\mathbb{S}^1)\simeq\mathbb{Z}$. □

Пример 12.2.2. Тор является факторпространством, $\mathbb{T}^n\approx\mathbb{R}^n/\mathbb{Z}^n$. Поскольку отображение $p:\ \mathbb{R}^n\to\mathbb{R}^n/\mathbb{Z}^n$ – универсальное накрытие, то $\pi(\mathbb{T}^n)\simeq\mathbb{Z}^n$. □

Следующая теорема дает достаточное условие существования дифференцируемой структуры на пространстве орбит $\mathbb{M}/\mathbb{G}$.

Теорема 12.2.4. *Пусть группа преобразований $\mathbb{G}$ действует на многообразии $\widetilde{\mathbb{M}}$ собственно разрывно и свободно, тогда факторпространство $\widetilde{\mathbb{M}}/\mathbb{G}$ с фактор топологией имеет структуру дифференцируемого многообразия такую, что проекция $p:\ \widetilde{\mathbb{M}}\to\widetilde{\mathbb{M}}/\mathbb{G}$ дифференцируема.*

Доказательство. Условие 1) в определении собственно разрывной группы преобразований эквивалентно хаусдорфовости факторпространства $\widetilde{\mathbb{M}}/\mathbb{G}$. Действительно, окрестностью орбиты $x=\tilde{x}\mathbb{G}$ является множество орбит $\mathbb{U}=\widetilde{\mathbb{U}}\mathbb{G}$, где $\widetilde{\mathbb{U}}$ – окрестность точки $\tilde{x}$, а условие 1) можно переписать в виде $\widetilde{\mathbb{U}}_1\mathbb{G}\cap\widetilde{\mathbb{U}}_2\mathbb{G}=\emptyset$, т.к. орбиты либо не имеют общих точек, либо совпадают. Это и есть условие хаусдорфовости пространства орбит с фактортопологией.

В условиях теоремы группа $\mathbb{G}$ действует свободно и условие 2) в определении собственно разрывной группы преобразований выполняется автоматически, т.к. группа изотропии любой точки состоит только из одного элемента – единицы.

Пусть $\mathbb{U}=\widetilde{\mathbb{U}}\mathbb{G}\subset\widetilde{\mathbb{M}}/\mathbb{G}$ – достаточно малая окрестность орбиты $x\in\widetilde{\mathbb{M}}/\mathbb{G}$. Условие 3) вместе с условием 2) значит, что каждая точка x в пространстве орбит имеет окрестность $\mathbb{U}$ такую, что прообраз $\widetilde{\mathbb{U}}=p^{-1}(\mathbb{U})$ состоит не более, чем из счетного числа компонент, $\widetilde{\mathbb{U}}=\bigcup_i\widetilde{\mathbb{U}}_i$ и проекция p каждой связной компоненты $\widetilde{\mathbb{U}}_i$ на $\mathbb{U}$ есть гомеоморфизм, т.к. множество орбит можно параметризовать точками из какой либо окрестности $\mathbb{U}_i$. Зафиксируем связную компоненту $\widetilde{\mathbb{U}}_i$ в $p^{-1}(\mathbb{U})$. Тогда p – гомеоморфизм $\widetilde{\mathbb{U}}_i$ на $\mathbb{U}$ и существует обратное непрерывное отображение $p^{-1}:\ \mathbb{U}\to\widetilde{\mathbb{U}}_i$. Выбрав $\mathbb{U}$ достаточно малым, можно считать, что имеется допустимая карта $(\widetilde{\mathbb{U}}_i,\varphi)$, где $\varphi:\ \widetilde{\mathbb{U}}_i\to\mathbb{R}^n$, многообразия $\widetilde{\mathbb{M}}$. Теперь можно ввести дифференцируемую структуру в пространстве орбит, рассматривая $(\mathbb{U},\psi)$, где $\psi=\varphi\circ p^{-1}$, как допустимую карту. □

Последняя теорема является достаточным, но не необходимым условием того, что пространство орбит является многообразием.

Пример 12.2.3. Рассмотрим вращения евклидовой плоскости на $\mathbb{R}^2$ вокруг начала координат на фиксированный угол $\alpha=2\pi/n$, где n – одно из натуральных чисел $2,3,\ldots$. При этом мы отождествляем поворот на 2π с единичным элементом группы. Эта группа $\mathbb{G}$ абелева и состоит из n элементов. Действие группы является эффективным. Начало координат является неподвижной точкой и одной из орбит группы. Ее группа изотропии совпадает с $\mathbb{G}$. На остальной части плоскости $\mathbb{R}^2\setminus\{0\}$ группа $\mathbb{G}$ действует свободно. Нетрудно проверить, что группа $\mathbb{G}$ действует на плоскость собственно разрывно. Пространство орбит $\mathbb{R}^2/\mathbb{G}$ представляет собой конус с углом дефицита $2\pi/n-2\pi$ (знак минус означает, что угол вырезается, а не вставляется). Конус является гладким двумерным многообразием, т.е. пространство орбит допускает

гладкую структуру, несмотря на то, что действие группы не является свободным. Вершина конуса – это "дефект" не многообразия, а вложения. В этом примере каноническая проекция $p:\ \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{G}$ не является накрытием, так как прообраз $p^{-1}(\mathbb{U}_0)$ окрестности $\mathbb{U}_0 \subset \mathbb{R}^2/\mathbb{G}$, содержащей начало координат, состоит из одного связного листа, который не гомеоморфен самой окрестности (нет взаимной однозначности). Фундаментальные группы плоскости $\mathbb{R}^2$ и конуса $\mathbb{R}^2/\mathbb{G}$ тривиальны и совпадают. □

Если пара $(\widetilde{\mathbb{M}}, \mathbb{G})$ – группа преобразований, причем группа $\mathbb{G}$ действует на многообразии $\widetilde{\mathbb{M}}$ свободно и собственной разрывно, то одним из способов изучения фактор пространства $\widetilde{\mathbb{M}}/\mathbb{G}$ является построение фундаментальной области.

Определение. Подмножество $\mathbb{D} \subset \widetilde{\mathbb{M}}$ называется *фундаментальной областью* многообразия $\widetilde{\mathbb{M}}$ для группы преобразований $\mathbb{G}$, действующей свободно и собственно разрывно, если выполнены следующие условия:

1) $\mathbb{D}$ является замкнутым подмножеством в $\widetilde{\mathbb{M}}$;
2) орбита $\mathbb{D}\mathbb{G}$ совпадает со всем многообразием $\widetilde{\mathbb{M}}$;
3) покрытие $\widetilde{\mathbb{M}}$ множествами $\mathbb{D}a$, $a \in \mathbb{G}$, таково, что с достаточно малой окрестностью произвольной точки $\widetilde{\mathbb{M}}$ пересекается лишь конечное число множеств вида $\mathbb{D}a$;
4) образ множества всех внутренних точек фундаментальной области, $(\,\mathsf{int}\,\mathbb{D})a$, при действии любого преобразования $a \in \mathbb{G}$, отличного от единичного, не пересекается с множеством внутренних точек фундаментальной области, $(\,\mathsf{int}\,\mathbb{D})a \cap \mathsf{int}\,\mathbb{D} = \emptyset$. □

Фундаментальная область $\mathbb{D}$ всегда является многообразием с краем той же размерности, что и само $\widetilde{\mathbb{M}}$.

Пример 12.2.4. Группа трансляций $\mathbb{G}$ евклидовой плоскости $\mathbb{R}^2$ на всевозможные векторы с целочисленными компонентами

$$\mathbb{R}^2 \ni (x, y) \mapsto (x + m, y + n) \in \mathbb{R}^2, \qquad \forall m, n \in \mathbb{Z},$$

действует гладко, свободно и собственно разрывно. Фактор пространство $\mathbb{R}^2/\mathbb{G}$ представляет собой тор $\mathbb{T}^2$. В качестве фундаментальной области можно выбрать единичный квадрат

$$\mathbb{D} = \{(x, y) \in \mathbb{R}^2 :\ \ 0 \le x \le 1,\ 0 \le y \le 1\}.$$

Тор можно представить в виде квадрата на плоскости, у которого отождествлены противоположные стороны. □

Из определения фундаментальной области следует, что отображение $\widetilde{\mathbb{M}} \to \mathsf{int}\,\mathbb{D}$ всегда представляет собой накрытие. Если фундаментальная область известна, то пространство орбит $\widetilde{\mathbb{M}}/\mathbb{G}$ получается из фундаментальной области путем склеивания граничных точек.

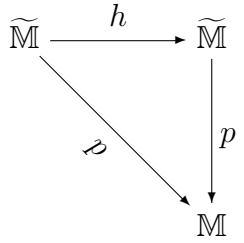
12.3 Группа скольжений и существование накрытий

В предыдущем разделе было показано, что отображение многообразия в пространство орбит для группы преобразований, действующей свободно и собственно разрывно, является накрытием. Теперь мы рассмотрим обратный вопрос о том, в каком случае заданное накрытие можно представить, в виде отображения накрывающего

пространства в пространство орбит относительно действия некоторой группы преобразований и какова эта группа.

Введем новое понятие, которое дает возможность описать произвол, существующий при построении накрывающего пространства, если база задана.

Определение. *Группой скольжений* накрытия $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ называется группа всех гомеоморфизмов $h:\ \widetilde{\mathbb{M}} \to \widetilde{\mathbb{M}}$, при которых $p \circ h = p$, т.е. диаграмма



коммутативна. Эта группа обозначается $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. □

Теорема 12.3.1. *Если топологическое пространство $\widetilde{\mathbb{M}}$ связно и локально линейно связно, то действие группы скольжений $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$ на $\widetilde{\mathbb{M}}$ свободно и собственно разрывно.*

Доказательство. См., например, [**?**]. □

В частности, действие группы скольжений на связном многообразии $\widetilde{\mathbb{M}}$ является свободным и собственно разрывным. В этом случае группа скольжений $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$ конечна или счетна. Поэтому пара $\left(\widetilde{\mathbb{M}}, \mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})\right)$ является группой преобразований.

Теорема 12.3.2. *Пусть $p:\ \widetilde{\mathbb{M}} \to \mathbb{M}$ – накрытие и накрывающее топологическое пространство $\widetilde{\mathbb{M}}$ связно и локально линейно связно. Если индуцированная группа $p_*\pi(\widetilde{\mathbb{M}}, \tilde{x})$ является нормальной подгруппой $\pi(\mathbb{M}, x)$, то база $\mathbb{M}$ гомеоморфна пространству орбит $\widetilde{\mathbb{M}}/\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$.*

Доказательство. См., например, [**?**]. □

Следствие. Пусть пространство $\widetilde{\mathbb{M}}$ связно и локально линейно связно. Если $p_*\pi(\widetilde{\mathbb{M}}, \tilde{x})$ – нормальная подгруппа $\pi(\mathbb{M}, x)$, то $\pi(\mathbb{M}, x)/p_*\pi(\widetilde{\mathbb{M}}, \tilde{x}) \simeq \mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. □

Доказательство. Прямое следствие теорем 12.3.2 и 12.2.3. □

Следствие. Если накрывающее пространство $\widetilde{\mathbb{M}}$ односвязно и локально линейно связно, то $\pi(\mathbb{M}, x) \simeq \mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. □

По своей сути две предыдущие теоремы и следствия означают, что произвольное накрытие можно представить, как отображение многообразия $\widetilde{\mathbb{M}}$ в пространство орбит $\mathbb{M} = \widetilde{\mathbb{M}}/\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. При этом роль группы преобразований играет группа скольжений $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. Если накрывающее пространство $\widetilde{\mathbb{M}}$ – многообразие, то действие группы скольжений сводится к перестановке листов накрытия.

Теперь обсудим вопрос о существовании универсального накрытия для заданной базы $\mathbb{M}$. Универсальное накрытие существует, если на топологическое пространство $\mathbb{M}$ наложен ряд условий. Чтобы не вводить новых понятий для их формулировки, мы ограничимся многообразиями, для которых эти условия выполняются.

Теорема 12.3.3. *Произвольное связное n-мерное многообразие $\mathbb{M}$ имеет универсальное накрытие $p: \widetilde{\mathbb{M}} \to \mathbb{M}$, где накрывающее пространство $\widetilde{\mathbb{M}}$ – также n-мерное многообразие. Универсальное накрывающее пространство определено с точностью до действия группы скольжений $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$, которая действует на $\widetilde{\mathbb{M}}$ свободно и собственно разрывно.*

Доказательство. См., например, [**?**]. □

Эта теорема очень важна, т.к. позволяет разделить задачу классификации многообразий на два этапа: сначала описать все односвязные многообразия, а затем найти все группы преобразований, действующие на них свободно и собственно разрывно. Эта задача решена для поверхностей. Для многообразий размерности три и выше вопрос остается открытым.

Глава 13

Главные и ассоциированные расслоения

Теория расслоений или расслоенных пространств играет важнейшую роль в современной математической физике. Достаточно отметить, что в основе общей теории относительности лежит расслоение реперов, которое является главным расслоением со структурной группой $\mathbb{GL}(n,\mathbb{R})$. В основе калибровочных моделей элементарных частиц лежит главное расслоение с полупростой структурной группой, которой обычно является некоторая подгруппа унитарной группы $\mathbb{U}(n)$. В настоящем разделе мы дадим определения и рассмотрим основные свойства главных и ассоциированных расслоений.

13.1 Главные расслоения

В дифференциальной геометрии важнейшую роль играют главные расслоения. Фактически, они лежат в основе многих геометрических конструкций.

Определение. *Главным расслоением* называется четверка $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$, где $\mathbb{P}$ и $\mathbb{M}$ – многообразия, $\mathbb{G}$ – группа Ли, $\pi:\ \mathbb{P}\to\mathbb{M}$ – отображение, которые удовлетворяют следующим условиям:

1) определено свободное дифференцируемое действие группы $\mathbb{G}$ на $\mathbb{P}$ справа:

$$\mathbb{P}\times\mathbb{G}\ni(p,a)\mapsto pa\in\mathbb{P}; \tag{13.1}$$

2) $\mathbb{M}$ есть факторпространство для $\mathbb{P}$ по отношению эквивалентности, индуцированному действием группы $\mathbb{G}$, и каноническая проекция $\pi:\ \mathbb{P}\to\mathbb{M}=\mathbb{P}/\mathbb{G}$ дифференцируема;

3) каждая точка $x\in\mathbb{M}$ имеет окрестность $\mathbb{U}_x$ такую, что существует диффеоморфизм

$$\chi:\ \pi^{-1}(\mathbb{U}_x)\ni p\mapsto\chi(p)=\Big(\pi(p),\varphi(p)\Big)\in\mathbb{U}_x\times\mathbb{G} \tag{13.2}$$

такой, что отображение $\varphi:\ \pi^{-1}(\mathbb{U}_x)\to\mathbb{G}$ удовлетворяет условию $\varphi(pa)=\Big(\varphi(p)\Big)a$ для всех $p\in\pi^{-1}(\mathbb{U}_x)$ и $a\in\mathbb{G}$ (локальная тривиальность).

Многообразие $\mathbb{P}$ называется *пространством расслоения*, $\mathbb{M}$ – *базой расслоения*, $\mathbb{G}$ – *структурной группой* и π – *проекцией*. □

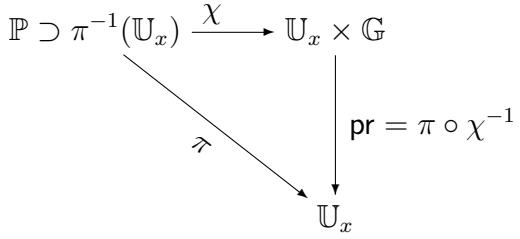
Поскольку отображение (13.2) является диффеоморфизмом, то

$$\mathsf{dim}\,\mathbb{P}=\ \mathsf{dim}\,\mathbb{M}+\ \mathsf{dim}\,\mathbb{G}.$$

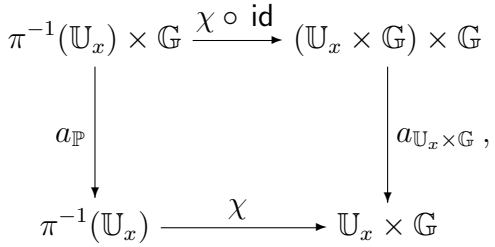
Замечание. Каждое главное расслоение является расслоением в смысле определения, данного в разделе 3.4. В дополнение к общему определению расслоения мы зафиксировали типичный слой, предположив, что им является группа Ли $\mathbb{G}$, и добавили действие этой группы на пространстве расслоения $\mathbb{P}$ так, чтобы оно было согласовано с проекцией. Дифференцируемая структура на $\mathbb{P}$ согласована с дифференцируемыми структурами на $\mathbb{M}$ и $\mathbb{G}$, поскольку отображение χ, по предположению, является диффеоморфизмом. На самом деле, можно было бы потребовать только непрерывность отображения χ, а затем с его помощью перенести дифференцируемую структуру с $\mathbb{M}$ и $\mathbb{G}$ на пространство расслоения $\mathbb{P}$. □

Замечание. Пара $(\mathbb{P}, \mathbb{G})$ является группой преобразований, определенной в разделе 10. Однако, не всякая группа преобразований есть главное расслоение. Напомним, что в общем случае пространство орбит $\mathbb{M}/\mathbb{G}$ группы преобразований $(\mathbb{M}, \mathbb{G})$ может оказаться нехаусдорфовым топологическим пространством и на нем нельзя ввести дифференцируемую структуру. □

Замечание. Условие локальной тривиальности главного расслоения можно изобразить в виде коммутативной диаграммы



где pr – естественная проекция прямого произведения $\mathbb{U}_x \times \mathbb{G} \xrightarrow{\mathsf{pr}} \mathbb{U}_x$ на первый сомножитель. Групповое действие на пространстве расслоения описывается эквивариантной диаграммой



где $a_{\mathbb{P}}$ и $a_{\mathbb{U}_x \times \mathbb{G}}$ обозначают действие элемента группы $a \in \mathbb{G}$ на пространстве расслоения и прямом произведении $\mathbb{M} \times \mathbb{G}$. □

Определение. Для каждой точки базы $x \in \mathbb{M}$ множество $\pi^{-1}(x)$ есть замкнутое подмногообразие в пространстве расслоения $\mathbb{P}$, которое называется *слоем* над x. *Сечением* или *глобальным сечением* главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется дифференцируемое отображение $\sigma : \mathbb{M} \to \mathbb{P}$ такое, что $\pi \circ \sigma = \mathsf{id}_{\mathbb{M}}$. Дифференцируемое отображение $\sigma : \mathbb{U} \to \mathbb{P}$, где $\mathbb{U} \subset \mathbb{M}$ – некоторая окрестность базы, называется *локальным сечением*, если $\pi \circ \sigma = \mathsf{id}_{\mathbb{U}}$. □

Каждый слой является орбитой $p\mathbb{G}$ какой либо точки $p \in \pi^{-1}(x)$. То, что каждый слой представляет собой замкнутое подмногообразие в $\mathbb{P}$ является следствием предложения 10.1.3. Поскольку действие группы Ли $\mathbb{G}$ в каждом слое свободно и транзитивно, то между точками слоя и структурной группы имеется взаимно однозначное

соответствие. То есть каждый слой $\pi^{-1}(x)$ диффеоморфен $\mathbb{G}$. Этот диффеоморфизм осуществляет функция φ в отображении (13.2) при фиксированном $x \in \mathbb{M}$.

При каждом фиксированном $p \in \mathbb{P}$ отображение (13.1) дифференцируемо. Поэтому каждой точке главного расслоения соответствует диффеоморфизм структурной группы на типичный слой в данной точке

$$p: \quad \mathbb{G} \to \pi^{-1}(x), \tag{13.3}$$

где $x = \pi(p)$. В дальнейшем мы иногда будем рассматривать точку p именно в этом смысле, как отображение.

Сечение σ, если оно существует, не может быть сюръективным отображением, т.к. размерность базы в общем случае меньше размерности расслоения. Для накрытий размерность базы совпадает с размерностью главного расслоения. Если накрытие многолистно, то сечение также не является сюръективным отображением. Сечение сюръективно в одном случае, когда структурная группа состоит из единственного элемента – единицы. Сужение проекции π на образ базы $\sigma(\mathbb{M})$ в $\mathbb{P}$ является дифференцируемым отображением, которое обратно к сечению σ. Поэтому пара $(\sigma, \mathbb{M})$ является подмногообразием в $\mathbb{P}$. Если $\dim \mathbb{G} \geq 1$, то это подмногообразие замкнуто в $\mathbb{P}$ (см. раздел 3.10).

Пример 13.1.1. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{M}$ – многообразие. Определим действие группы $\mathbb{G}$ справа на прямом произведении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$:

$$\mathbb{G} \ni b: \quad \mathbb{M} \times \mathbb{G} \ni (x, a) \mapsto (x, ab) \in \mathbb{M} \times \mathbb{G}.$$

Тогда четверка $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\pi: \ \mathbb{M} \times \mathbb{G} \to \mathbb{M}$ – каноническая проекция, является главным расслоением, которое является тривиальным. Каждое тривиальное главное расслоение имеет глобальное сечение $\sigma: \ \mathbb{M} \to \mathbb{P}$, $\pi \circ \sigma = \mathsf{id}_{\mathbb{M}}$. Например, $\sigma: \ \mathbb{M} \ni x \mapsto (x, e) \in \mathbb{M} \times \mathbb{G}$. □

Пример 13.1.2. Расслоение реперов $\mathbb{L}(\mathbb{M}) = \mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{GL}(n, \mathbb{R})\Big)$, рассмотренное в разделе 6.4, является главным расслоением со структурной группой $\mathbb{GL}(n, \mathbb{R})$. □

Пример 13.1.3. Пусть $\mathbb{G}$ – группа Ли и $\mathbb{H} \subset \mathbb{G}$ – ее замкнутая подгруппа. Согласно теореме 10.1.1 на пространстве правых смежных классов $\mathbb{H}a \in \mathbb{G}/\mathbb{H}$, где $a \in \mathbb{G}$, можно задать дифференцируемую структуру. Тогда $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ – главное расслоение с базой $\mathbb{G}/\mathbb{H}$, структурной группой $\mathbb{H}$ и проекцией $\pi: \ \mathbb{G} \to \mathbb{G}/\mathbb{H}$. □

Пример 13.1.4. Накрытия $\widetilde{\mathbb{M}}\Big(\mathbb{M}, \pi, \mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})\Big)$, где $p: \ \widetilde{\mathbb{M}} \to \mathbb{M}$, рассмотренные в разделе 12, являются главными расслоениями с 0-мерной структурной группой Ли, которой является группа скольжений $\mathbb{G}(\widetilde{\mathbb{M}}, p, \mathbb{M})$. □

Пример 13.1.5. Множество вещественных чисел без нуля образует абелеву группу относительно умножения, которую мы обозначим $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$. Пусть эта группа действует в евклидовом пространстве без начала координат $\mathbb{R}^n \setminus \{0\}$ посредством умножения декартовых координат

$$\mathbb{R}_0 \ni a: \quad (x^1, \ldots, x^n) \mapsto (ax^1, \ldots, ax^n).$$

Это действие дифференцируемо, свободно и транзитивно на орбитах. Поэтому $\mathbb{R}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{RP}^{n-1}$ – главное расслоение со структурной группой $\mathbb{R}_0$, базой которого является вещественное проективное пространство $\mathbb{RP}^{n-1} = \Big(\mathbb{R}^n \setminus \{0\}\Big)/\mathbb{R}_0$. □

Пример 13.1.6. Если в предыдущем примере группу $\mathbb{R}_0$ заменить на группу положительных чисел $\mathbb{R}_+$ по умножению, то получим главное расслоение $\mathbb{R}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{S}^{n-1}$ со структурной группой $\mathbb{R}_+$, базой которого является сфера $\mathbb{S}^{n-1} = \Big(\mathbb{R}^n \setminus \{0\}\Big)/\mathbb{R}_+$. □

Пример 13.1.7. Пусть $\mathbb{C}_0 = \mathbb{C} \setminus \{0\}$ – группа комплексных чисел по отношению к умножению, которая действует в n-мерном комплексном пространстве без начала координат $\mathbb{C}^n \setminus \{0\}$ умножением,

$$\mathbb{C}_0 \ni a: \quad (z^1, \dots, z^n) \mapsto (az^1, \dots, az^n).$$

Тогда $\mathbb{C}^n \setminus \{0\} \xrightarrow{\pi} \mathbb{CP}^{n-1}$ – главное расслоение со структурной группой $\mathbb{C}_0$, базой которого является комплексное проективное пространство $\mathbb{CP}^{n-1} = \Big(\mathbb{C}^n \setminus \{0\}\Big)/\mathbb{C}_0$. □

В примере 13.1.1 мы отметили, что произвольное тривиальное расслоение имеет глобальное сечение. Справедливо и обратное утверждение.

Предложение 13.1.1. *Если главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ имеет глобальное сечение $\sigma: \mathbb{M} \to \mathbb{P}$, то оно изоморфно тривиальному, т.е. существует диффеоморфизм*

$$f_{\mathbb{P}}: \quad \mathbb{P} \ni p \mapsto \Big(\pi(p), \varphi(p)\Big) \in \mathbb{M} \times \mathbb{G} \tag{13.4}$$

такой, что следующая диаграмма коммутативна:

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{\;f_{\mathbb{P}}\;} & \mathbb{M} \times \mathbb{G} \\ & \searrow & \downarrow \mathsf{pr}_{\mathbb{M}} \\ & & \mathbb{M} \end{array} \tag{13.5}$$

где $\mathsf{pr}_{\mathbb{M}}$ – проекция на первый сомножитель, и справедливо равенство

$$f_{\mathbb{P}}(pa) = \Big(\pi(p), \varphi(p)a\Big) \qquad \forall\ p \in \mathbb{P},\ a \in \mathbb{G}. \tag{13.6}$$

Доказательство. Пусть

$$\sigma: \quad \mathbb{M} \ni x \mapsto \sigma(x) \in \mathbb{P}, \qquad \pi \circ \sigma = \mathsf{id}_{\mathbb{M}},$$

– глобальное сечение главного расслоения. Для каждой точки слоя $p \in \pi^{-1}(x)$ определим единственный элемент группы $a \in \mathbb{G}$, для которого $p = \sigma(x)a(p)$. Таким образом определено отображение $p \mapsto f_{\mathbb{P}}(p) = \Big(\pi(p), a(p)\Big)$. Легко проверить, что это отображение удовлетворяет требуемым свойствам. □

Замечание. Изоморфизм расслоений будет определен позже в разделе 13.3. □

Следствие. Нетривиальные главные расслоения имеют только локальные сечения. □

Диаграмму (13.5) вместе с условием (13.6) можно изобразить в виде коммутативной диаграммы

$$\begin{array}{ccc} \mathbb{P} \times \mathbb{G} & \xrightarrow{\;\chi \circ \mathsf{id}\;} & (\mathbb{M} \times \mathbb{G}) \times \mathbb{G} \\ \downarrow a_{\mathbb{P}} & & \downarrow a_{\mathbb{M}\times\mathbb{G}}\ , \\ \mathbb{P} & \xrightarrow{\;\chi\;} & \mathbb{M} \times \mathbb{G} \end{array}$$

где $a_{\mathbb{P}}$ и $a_{\mathbb{M}\times\mathbb{G}}$ обозначает действие элемента $a \in \mathbb{G}$ соответственно на $\mathbb{P}$ и $\mathbb{M}\times\mathbb{G}$. Это означает, что отображение (13.4) является эквивариантным (см. раздел 10.4).

Определение. Главное расслоение $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ называется *тривиальным*, если оно изоморфно главному расслоению вида $\mathbb{M}\times\mathbb{G} \xrightarrow{\mathsf{pr}_{\mathbb{M}}} \mathbb{M}$. □

Теорема 13.1.1. *Если база главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ диффеоморфна евклидову пространству, $\mathbb{M}\approx\mathbb{R}^n$, то главное расслоение изоморфно тривиальному, $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})\simeq\mathbb{M}\times\mathbb{G}$.*

Доказательство. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{R}^n,\pi,\mathbb{G})$, базой которого является евклидово пространство. Согласно теореме 14.1.2 любое главное расслоение допускает связность. Предположим, что на $\mathbb{P}(\mathbb{R}^n,\pi,\mathbb{G})$ задана какая либо связность. Рассмотрим в $\mathbb{R}^n$ сферическую систему координат. Тогда каждая отличная от начала координат точка $x\in\mathbb{R}^n$ параметризуется парой (r,n), где r – расстояние от начала координат и n – единичный вектор, определяющий направление луча, выходящего из начала координат и на котором лежит точка x. Пусть $p_0\in\pi^{-1}(0)$ – произвольная точка из слоя над началом координат. Согласно предложению 14.3.1 для каждого луча существует его единственный горизонтальный лифт в пространство расслоения $\mathbb{P}$ с началом в точке p_0. Обозначим через $p(x)$ единственную точку на горизонтальном лифте луча, которая лежит над x, т.е. $\pi(p)=x$. Таким образом, $p(x)$ – это глобальное сечение главного расслоения. Следовательно, по предложению 13.1.1 главное расслоение тривиально $\mathbb{P}=\mathbb{R}^n\times\mathbb{G}$. Если база диффеоморфна $\mathbb{R}^n$, то сферическая система координат просто переносится на $\mathbb{M}$ с помощью диффеоморфизма, и построение глобального сечения повторяется. □

При рассмотрении многообразий в разделе 3.1 были введены функции склейки, с помощью которых осуществляется преобразование координат в двух пересекающихся картах. Обобщением этого понятия на случай главных расслоений являются функции перехода, которые вводятся следующим образом.

Определение. В силу условия 3) в определении главного расслоения на базе можно выбрать такое координатное покрытие, $\mathbb{M}=\bigcup_i\mathbb{U}_i$, что

$$\chi_i:\quad \pi^{-1}(\mathbb{U}_i)\ni p\mapsto\Big(\pi(p),\varphi_i(p)\Big)\in\mathbb{U}_i\times\mathbb{G},$$

причем $\varphi_i(pa)=\varphi_i(p)a$.

Пусть две карты пересекаются, $\mathbb{U}_i\cap\mathbb{U}_j\neq\emptyset$. Если $p\in\pi^{-1}(\mathbb{U}_i\cap\mathbb{U}_j)$, то

$$\varphi_j(pa)\circ\varphi_i(pa)^{-1}=\varphi_j(p)\circ\varphi_i(p)^{-1},$$

где φ_i рассматривается как отображение фиксированного слоя $\pi^{-1}\Big(\pi(p)\Big)$ в группу $\mathbb{G}$. Следовательно, отображение $\varphi_j(p)\circ\varphi_i(p)^{-1}$ зависит только от точки базы $x=\pi(p)$. Поэтому определено отображение

$$a_{ji}:\quad \mathbb{U}_i\cap\mathbb{U}_j\ni x=\pi(p)\mapsto a_{ji}(x)=\varphi_j(p)\circ\varphi_i(p)^{-1}\in\mathbb{G}. \tag{13.7}$$

Эти функции на $\mathbb{M}$ со значениями в $\mathbb{G}$ называются *функциями перехода* или *функциями склейки* главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$, соответствующими координатному покрытию $\mathbb{M}=\bigcup_i\mathbb{U}_i$. Набор функций склейки называется *склеивающим коциклом* главного расслоения $\mathbb{P}$. □

Функции перехода показывают насколько в различных картах "сдвинуты" образы слоя в структурной группе над фиксированной точкой $x \in \mathbb{M}$.

Из определения (13.7) следует, что функции склейки обладают следующим свойством

$$a_{ij}(x) = \Big(a_{ji}(x)\Big)^{-1} \qquad x \in \mathbb{U}_i \cap \mathbb{U}_j. \tag{13.8}$$

Кроме того, нетрудно проверить, что для трех пересекающихся карт выполнено равенство

$$a_{ij}(x)a_{jk}(x)a_{ki}(x) = e, \qquad x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k. \tag{13.9}$$

Поэтому определение функций перехода корректно.

Пример 13.1.8. Пусть $\mathbb{M}$, $\dim \mathbb{M} = n$, – многообразие. Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{M})$ (см. раздел 6.4). Пусть $\mathbb{U}_i$ и $\mathbb{U}_j$ – две пересекающиеся карты с координатами x^α и $x^{\alpha'}$, $\alpha, \alpha' = 1, \ldots, n$. Репер в этих картах имеет компоненты $e^\alpha{}_a \in \mathbb{GL}(n, \mathbb{R})$ и $e^{\alpha'}{}_a \in \mathbb{GL}(n, \mathbb{R})$, $a = 1, \ldots, n$, которые связаны между собой преобразованием

$$e^{\alpha'}{}_a = e^\alpha{}_a \partial_\alpha x^{\alpha'}.$$

Таким образом, функциями перехода для расслоения реперов являются матрицы Якоби преобразования координат. Эти матрицы, как легко проверить, удовлетворяют условиям (13.8) и (13.9). □

Таким образом, для каждого главного расслоения можно однозначно построить семейство функций перехода, соответствующих заданному координатному покрытию базы, и эти функции перехода удовлетворяют равенству (13.9). Справедливо также обратное утверждение.

Теорема 13.1.2. *Пусть $\mathbb{M}$ – многообразие с координатным покрытием $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и $\mathbb{G}$ – группа Ли. Если заданы отображения $a_{ji} : \mathbb{U}_i \cap \mathbb{U}_j \to \mathbb{G}$ для всех непустых пересечений $\mathbb{U}_i \cap \mathbb{U}_j$ такие, что выполнены условия (13.8) и (13.9) во всех областях пересечения карт, то существует единственное с точностью до изоморфизма главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с функциями перехода a_{ji}.*

Доказательство. Рассмотрим несвязное объединение $\mathbb{Q} := \sqcup_i (\mathbb{U}_i \times \mathbb{G})$. Введем на этом множестве отношение эквивалентности:

$$\begin{aligned} &(x, a) \sim \Big(x, aa_{ij}(x)\Big), && \forall \ \ (x, a) \in \mathbb{U}_i \times \mathbb{G}, \ \ \Big(x, aa_{ij}(x)\Big) \in \mathbb{U}_j \times \mathbb{G}, \\ &(x, a) \sim (x, a), && \forall \ \ (x, a) \in \mathbb{U}_i \times \mathbb{G}. \end{aligned}$$

Из свойств (13.8), (13.9) следует, что это действительно отношение эквивалентности. Обозначим фактор пространство $\mathbb{Q}/\sim$ через $\mathbb{P}$ и введем на нем естественную дифференцируемую структуру. Пусть $p := \langle x, a\rangle$ – точка $\mathbb{P}$, т.е. класс эквивалентности пары (x, a). Определим действие группы $\mathbb{G}$ на $\mathbb{P}$ формулой

$$\mathbb{P} \times \mathbb{G} \ni \langle x, a\rangle, b \mapsto \langle x, ab\rangle \in \mathbb{P}.$$

Определим также проекцию

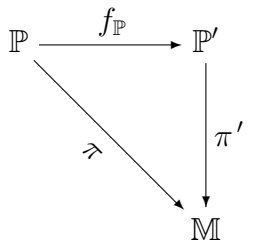
$$\pi : \quad \mathbb{P} \ni \langle x, a\rangle \mapsto \pi(\langle x, a\rangle) := x \in \mathbb{M}.$$

Нетрудно проверить, что все свойства главного расслоения для четверки $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ выполнены.

Теперь докажем единственность построенного главного расслоения с точностью до изоморфизма. Пусть множество функций склеек $\{a_{ji}(x)\}$ построено для некоторого главного расслоения $\mathbb{P}'(\mathbb{M},\pi',\mathbb{G})$ с фиксированным покрытием базы $\mathbb{M}=\cup_i\mathbb{U}_i$. Построим гомеоморфизм $f_{\mathbb{P}}:\ \mathbb{P}\to\mathbb{P}'$, где $f_{\mathbb{P}}$ совпадает с χ_i^{-1} на каждом $\mathbb{U}_i\times\mathbb{G}$. Для корректности этого определения нужно убедиться, что отображения χ_i^{-1} и χ_j^{-1} совпадают на общей области определения. Действительно, если $x\in\mathbb{U}_i\cap\mathbb{U}_j$ и $(x,a)\in\mathbb{U}_i\times\mathbb{G}$, то по построению пары (x,a) и (x,aa_{ji}) определяют одну и ту же точку $\langle x,a\rangle\in\mathbb{P}$. Из определения функций склейки $a_{ji}(x)$ следует

$$\chi_i^{-1}(x,a)=\chi_j^{-1}\circ\chi_j\circ\chi_i^{-1}(x,a)=\chi_j^{-1}(x,aa_{ji}).$$

Эквивариантность отображения $f_{\mathbb{P}}$ и коммутативность диаграммы



(13.10)

следуют из построения. Это и означает, что главные расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и $\mathbb{P}'(\mathbb{M},\pi',\mathbb{G})$ изоморфны. (Подробнее изоморфизм расслоений рассматривается в разделе 13.3. Там будет показано, что построенный изоморфизм $f_{\mathbb{P}}$ относится к классу вертикальных автоморфизмов.). □

Склеивающий коцикл определяется расслоением не однозначно. Он зависит от выбора координатного покрытия базы $\mathbb{M}=\bigcup_i\mathbb{U}_i$ и тривиализаций $\{\chi_i\}$.

Определение. Два коцикла $\{a_{ji}\}$ и $\{a'_{ji}\}$, соответствующие данному координатному покрытию базы $\mathbb{M}=\bigcup_i\mathbb{U}_i$, называются *эквивалентными*, если существуют отображения $b_i:\ \mathbb{U}_i\to\mathbb{G}$ такие, что выполнены равенства

$$a'_{ji}=b_ja_{ji}b_i^{-1},\qquad\forall x\in\mathbb{U}_i\cap\mathbb{U}_j.$$ □

Теорема 13.1.3. *Два главных расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и $\mathbb{P}'(\mathbb{M},\pi',\mathbb{G})$ изоморфны тогда и только тогда, когда их склеивающие коциклы $\{a_{ji}\}$ и $\{a'_{ji}\}$, соответствующие некоторому координатному покрытию базы $\mathbb{M}=\bigcup_i\mathbb{U}_i$, эквивалентны.*

Доказательство. См., например, [**?**]. □

Пример 13.1.9 (Расслоение Хопфа). Реализуем трехмерную сферу $\mathbb{S}^3$ в двумерном комплексном пространстве $\mathbb{C}^2$ с помощью вложения

$$|z^1|^2+|z^2|^2=1,\qquad(z^1,z^2)\in\mathbb{C}^2.\tag{13.11}$$

Каждое уравнение $az^1+bz^2=0$, где $(a,b)\in\mathbb{C}^2$ и по крайней мере одно из комплексных чисел a,b отлично от нуля, задает *комплексную прямую*, комплексной размерности один, проходящую через начало координат. Обратно. Комплексная прямая определяет пару комплексных чисел $(a,b)\in\mathbb{C}^2$ с точностью до отношения эквивалентности $(a,b)\sim(a',b')$, если и только если $|a'|=\lambda|a|$, $|b'|=\lambda|b|$, $\lambda\in\mathbb{R}_+$, или $a'=a\,\mathsf{e}^{it}$, $b'=b\,\mathsf{e}^{it}$, $t\in[0,2\pi]$. Множество всех комплексных прямых, проходящих через начало координат, является одномерным комплексным проективным пространством $\mathbb{CP}^1$ и называется *комплексной проективной прямой*. Каждая комплексная

прямая, проходящая через начало координат, пересекает трехмерную сферу $\mathbb{S}^3$ по большой окружности $\mathbb{S}^1$, называемой *окружностью Хопфа*, которая задается одной из двух систем уравнений:

$$|z^1|^2 = \frac{1}{1+|k|^2}, \qquad z^2 = kz^1, \qquad k := -\frac{a}{b}, \quad b \neq 0,$$
$$|z^2|^2 = \frac{1}{1+|k|^2}, \qquad z^1 = kz^2, \qquad k := -\frac{b}{a}, \quad a \neq 0.$$

Поскольку ровно одна окружность Хопфа проходит через каждую точку $\mathbb{S}^3$, эти окружности заполняют всю трехмерную сферу. При этом окружности Хопфа взаимно однозначно соответствуют комплексным прямым в $\mathbb{C}^2$, проходящим через начало координат. Две точки (z^1, z^2) и (z'^1, z'^2) лежат на одной окружности Хопфа тогда и только тогда, когда их координаты отличаются на фазовый множитель $z'^1 = z^1\,\mathsf{e}^{it}$, $z'^2 = z^2\,\mathsf{e}^{it}$. Выбрав подходящим образом параметр t, всегда можно добиться, например, чтобы $\mathsf{im}\, z^2 = 0$. Тогда каждой прямой, проходящей через начало координат, будет соответствовать одна и только одна точка на окружности Хопфа. Множество таких точек образует двумерную сферу $\mathbb{S}^2$, которая задается в $\mathbb{C}^2$ уравнением

$$(x^1)^2 + (y^1)^2 + (x^2)^2 = 1, \qquad y^2 = 0, \tag{13.12}$$

где $z^1 = x^1 + iy^1$, $z^2 = x^2 + iy^2$. Таким образом, комплексная проективная прямая диффеоморфна двумерной сфере, $\mathbb{CP}^1 \approx \mathbb{S}^2$. Соответствующая проекция $\mathbb{S}^3 \xrightarrow{\pi} \mathbb{S}^2$ определяет главное расслоение с базой $\mathbb{S}^2$ и типичным слоем $\mathbb{G} = \mathbb{U}(1) \approx \mathbb{S}^1$, которое называется *расслоением Хопфа*. Структурная группа действует на пространстве расслоения $\mathbb{S}^3$ умножением комплексных координат на e^{it}, $t \in [0, 2\pi]$.

Отображение $g_t : \mathbb{S}^3 \to \mathbb{S}^3$, $t \in \mathbb{R}$, задаваемое умножением комплексных координат в $\mathbb{C}^2$ на e^{it}, является изометрией для метрики, индуцированной вложением (13.11). Отсюда следует, что у трехмерной сферы $\mathbb{S}^3$ существуют изометрии, не имеющие выделенной оси: в каждой точке это движение выглядит точно так же, как и в любой другой. Это показывает, что трехмерная сфера $\mathbb{S}^3$ "более круглая", чем ее двумерный аналог $\mathbb{S}^2$. Однопараметрическая группа преобразований g_t называется *потоком Хопфа*.

Очевидно, что метрика проективной прямой $\mathbb{CP}^1 \approx \mathbb{S}^2$, индуцированная вложением (13.12), является стандартной метрикой сферы.

На расслоение Хопфа можно взглянуть с другой точки зрения. Вспомним, что трехмерная сфера может быть оснащена групповой структурой, $\mathbb{S}^3 \approx \mathbb{SU}^2$. (см. раздел 2.8). Двумерные унитарные матрицы с единичным определителем могут быть параметризованы двумя комплексными числами

$$U = \begin{pmatrix} z^1 & z^2 \\ -\bar{z}^2 & \bar{z}^1 \end{pmatrix} \in \mathbb{SU}(2),$$

где z^1 и z^2 лежат на сфере (13.11). Группа $\mathbb{SU}(2)$ содержит подгруппу $\mathbb{U}(1)$, которую можно реализовать в виде диагональных матриц

$$\begin{pmatrix} \mathsf{e}^{it} & 0 \\ 0 & \mathsf{e}^{-it} \end{pmatrix} \in \mathbb{U}(1).$$

Эта подгруппа не является нормальной. Рассмотрим множество правых смежных классов $\mathbb{SU}(2)/\mathbb{U}(1)$. Два элемента группы $U, U' \in \mathbb{SU}(2)$ принадлежат одному смежному классу тогда и только тогда, когда $z'^1 = z^1\,\mathsf{e}^{it}$ и $z'^2 = z^2\,\mathsf{e}^{it}$. Поэтому расслоение Хопфа можно представить в эквивалентном виде $\mathbb{SU}(2) \xrightarrow{\pi} \mathbb{M} := \mathbb{SU}(2)/\mathbb{U}(1)$.

Действия группы $\mathbb{SU}(2)$ слева и справа на сферу $\mathbb{S}^3$, которую мы отождествим с $\mathbb{SU}(2)$,

$$U \mapsto a^{-1}Ub, \qquad a, b \in \mathbb{SU}(2)$$

коммутируют и являются изометриями. Тем самым мы имеем гомоморфизм $\mathbb{SU}(2) \times \mathbb{SU}(2) \to \mathbb{SO}(4)$, где $\mathbb{SO}(4)$ является группой движений евклидова пространства $\mathbb{R}^4 \approx \mathbb{C}^2$ и вложенной трехмерной сферы. При этом действие группы $\mathbb{SU}(2)$ слева и справа является транзитивным, а группой изотропии являются преобразования вида $\pm(1,1) \in \mathbb{SU}(2) \times \mathbb{SU}(2)$. Тем самым мы имеем сюрьективный гомоморфизм с ядром $\mathbb{Z}_2$ или изоморфизм

$$\mathbb{SO}(4) \simeq \frac{\mathbb{SU}(2) \times \mathbb{SU}(2)}{\mathbb{Z}_2}. \tag{13.13}$$

Теперь дадим координатное описание расслоения Хопфа. Покроем базу $\mathbb{S}^2 = \overline{\mathbb{C}}$, которую мы отождествим с расширенной комплексной плоскостью, двумя картами следующим образом. Введем координату "окрестности нуля" $z := z^2/z^1$ при $z_1 \neq 0$ и "окрестности бесконечности" $w := z^1/z^2$, при $z^2 \neq 0$. Тогда база $\mathbb{S}^2 = \overline{\mathbb{C}}$ покрывается двумя картами:

$$\mathbb{U}_0 := \{z : \ z \in \mathbb{C}\}, \quad \text{и} \quad \mathbb{U}_\infty := \{w : \ w \in \mathbb{C}\}.$$

Тривиализация расслоения Хопфа задается двумя отображениями:

$$\chi_0 : \ \mathbb{S}^3 \ni (z^1, z^2) \mapsto (z, \, \mathsf{e}^{i \arg z^1}) \in \mathbb{C} \times \mathbb{S}^1,$$
$$\chi_\infty : \ \mathbb{S}^3 \ni (z^1, z^2) \mapsto (w, \, \mathsf{e}^{i \arg z^2}) \in \mathbb{C} \times \mathbb{S}^1.$$

В области пересечения карт $\mathbb{U}_0 \cap \mathbb{U}_\infty$ координаты связаны преобразованием $z \mapsto w = 1/z$. Следовательно, расслоение Хопфа голоморфно. В области пересечения карт задано отображение:

$$\chi_\infty \circ \chi_0^{-1} : \quad \mathbb{C} \times \mathbb{S}^1 \ni (z, \, \mathsf{e}^{i \arg z^1}) \mapsto (z, \, \mathsf{e}^{i \arg z^2}) \in (\mathbb{C} \setminus \{0'rbrace) \times \mathbb{S}^1,$$

которое мы записали в координате z. Поскольку

$$\mathsf{e}^{i \arg z^2} = \mathsf{e}^{\arg z z^1} = \mathsf{e}^{i \arg z^1} \, \mathsf{e}^{i \arg z},$$

то функция склейки имеет вид $a_{\infty 0} = \mathsf{e}^{i \arg z}$. Таким образом, расслоение Хопфа имеет следующее координатное описание:

$$\mathbb{U}_0 = \mathbb{C}, \qquad \mathbb{U}_\infty = \overline{\mathbb{C}} \setminus \{0\},$$
$$a_{\infty 0} \ \ \mathbb{C} \setminus \{0\} \ni z \mapsto \mathsf{e}^{i \arg z} \in \mathbb{U}(1) \approx \mathbb{S}^1. \qquad \square$$

В заключение настоящего раздела поясним соотношение между связностью (топологической) пространства главного расслоения $\mathbb{P}$ и связностями базы $\mathbb{M}$ и структурной группы $\mathbb{G}$.

Предложение 13.1.2. *Если база $\mathbb{M}$ и структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ связны, то пространство расслоения $\mathbb{P}$ также связно.*

Доказательство. Пусть $\mathbb{M} = \bigcup_i \mathbb{U}_i$ покрытие базы открытыми множествами такое, что для каждого i существует тривиализация $\chi_i : \ \pi^{-1}(\mathbb{U}_i) \to \mathbb{U}_i \times \mathbb{G}$. Поскольку $\mathbb{U}_i$ и $\mathbb{G}$ связны, то $\pi^{-1}(\mathbb{U}_i)$ также связно. Поэтому $\mathbb{P}$ тоже связно, как объединение пересекающихся связных множеств. $\square$

Это предложение дает достаточное, но не необходимое условие связности пространства главного расслоения.

Пример 13.1.10. Рассмотрим главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}^2)$, базой которого является окружность, а структурной группой $\mathbb{Z}_2 = \{\pm 1\}$, см. пример 13.2.2. Структурная группа не является связной, однако пространство расслоения связно. □

13.2 Ассоциированные расслоения

Дать определение расслоения, ассоциированного с заданным главным расслоением, в нескольких предложениях довольно затруднительно. Ассоциированные расслоения, хотя и просты по своему содержанию, требуют некоторой конструкции, которую мы сейчас опишем.

Определение. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение, и $(\mathbb{F}, \mathbb{G})$ – группа преобразований, т.е. задано дифференцируемое отображение

$$\mathbb{F} \times \mathbb{G} \ni (v, a) \mapsto va \in \mathbb{F}. \tag{13.14}$$

Сейчас мы построим расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$, которое ассоциировано с главным расслоением $\mathbb{P}$, и типичным слоем которого является многообразие $\mathbb{F}$. Во-первых, определим действие группы $\mathbb{G}$ на прямом произведении многообразий $\mathbb{P} \times \mathbb{F}$ по формуле

$$\mathbb{G} \ni a: \quad \mathbb{P} \times \mathbb{F} \ni (p, v) \mapsto (pa, va) \in \mathbb{P} \times \mathbb{F}.$$

Факторпространство для $\mathbb{P} \times \mathbb{F}$ относительно такого действия группы обозначается $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} \mathbb{F}$. Немного позже мы введем на $\mathbb{E}$ дифференцируемую структуру, а пока $\mathbb{E}$ – всего лишь множество. Проекция π в главном расслоении $\mathbb{P}$ определяет отображение, которое мы обозначим той же буквой

$$\pi: \quad \mathbb{P} \times \mathbb{F} \ni (p, v) \mapsto x = \pi(p) \in \mathbb{M}.$$

Это отображение индуцирует отображение факторпространства,

$$\pi_{\mathbb{E}}: \quad \mathbb{E} \ni u \mapsto x = \pi_{\mathbb{E}}(u) = \pi(p) \in \mathbb{M},$$

где $u = (p\mathbb{G}, v\mathbb{G})$ – орбита точки $(p, v) \in \mathbb{P} \times \mathbb{F}$, которое называется проекцией $\mathbb{E}$ на $\mathbb{M}$. Для каждой точки базы $x \in \mathbb{M}$ множество $\pi_{\mathbb{E}}^{-1}(x)$ называется слоем в $\mathbb{E}$ над x. Поскольку главное расслоение $\mathbb{P}$ локально тривиально, то каждая точка $x \in \mathbb{M}$ имеет окрестность $\mathbb{U} \subset \mathbb{M}$ такую, что прообраз $\pi^{-1}(\mathbb{U})$ диффеоморфен прямому произведению $\mathbb{U} \times \mathbb{G}$. Отождествляя $\pi^{-1}(\mathbb{U})$ с $\mathbb{U} \times \mathbb{G}$, мы видим, что действие структурной группы $\mathbb{G}$ на произведении $\pi^{-1}(\mathbb{U}) \times \mathbb{F}$ задается отображением

$$\mathbb{G} \ni b: \quad \mathbb{U} \times \mathbb{G} \times \mathbb{F} \ni (x, a, v) \mapsto (x, ab, vb) \in \mathbb{U} \times \mathbb{G} \times \mathbb{F}.$$

При этом орбита произвольной точки $(p, v) \in \mathbb{P} \times \mathbb{F}$ имеет вид

$$(x, a\mathbb{G}, v\mathbb{G}) = (x, \mathbb{G}, v\mathbb{G}), \qquad \forall\ a \in \mathbb{G},\ v \in \mathbb{F}.$$

На каждой орбите можно выбрать по одному представителю, соответствующему, например, единице группы $e \in \mathbb{G}$. Тогда множество орбит прямого произведения $\mathbb{U} \times \mathbb{G}$ будет параметризовано парой элементов (x, e, v), $x \in \mathbb{M}$, $v \in \mathbb{F}$. Отсюда следует, что

диффеоморфизм $\pi^{-1}(\mathbb{U}) \approx \mathbb{U}\times\mathbb{G}$ индуцирует изоморфизм $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \approx \mathbb{U}\times\mathbb{F}$ (локальная тривиальность). Поэтому на множестве $\mathbb{E}$ можно ввести дифференцируемую структуру. А именно, мы потребуем, чтобы множество $\pi_{\mathbb{E}}^{-1}(\mathbb{U})$ было открытым подмногообразием в $\mathbb{E}$, диффеоморфным прямому произведению $\mathbb{U}\times\mathbb{F}$ относительно изоморфизма $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \approx \mathbb{U}\times\mathbb{F}$. Таким образом построено отображение $\chi_{\mathbb{E}}:\ \pi_{\mathbb{E}}^{-1}(\mathbb{U}) \to \mathbb{U}\times\mathbb{F}$, которое является диффеоморфизмом. При этом проекция $\pi_{\mathbb{E}}$ будет дифференцируемым отображением $\pi_{\mathbb{E}}:\ \mathbb{E}\to\mathbb{M}$. Назовем объект $\mathbb{E}(\mathbb{M},\pi_{\mathbb{E}},\mathbb{F},\mathbb{G},\mathbb{P})$ *расслоением* с базой $\mathbb{M}$, проекцией $\pi_{\mathbb{E}}$, типичным слоем $\mathbb{F}$ и структурной группой $\mathbb{G}$, *ассоциированным* с главным расслоением $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$. □

Поскольку отображение $\chi_{\mathbb{E}}$ – диффеоморфизм, то

$$\mathsf{dim}\,\mathbb{E} = \mathsf{dim}\,\mathbb{M} + \mathsf{dim}\,\mathbb{F}.$$

Легко проверить, что расслоение, ассоциированное с главным расслоением, является расслоением в смысле определения, данного в разделе 3.4. В отличии от данного ранее определения на ассоциированном расслоении определено действие структурной группы $\mathbb{G}$ справа. Действительно, согласно построению у каждой точки базы $x\in\mathbb{M}$ существует окрестность $\mathbb{U}$ такая, что прообраз $\pi_{\mathbb{E}}^{-1}(\mathbb{U})$ диффеоморфен прямому произведению $\mathbb{U}\times\mathbb{F}$. Отождествляя $\pi_{\mathbb{E}}^{-1}(\mathbb{U})$ с $\mathbb{U}\times\mathbb{F}$, действие структурной группы справа задано равенством

$$\mathbb{G}\ni a:\quad \mathbb{U}\times\mathbb{F}\ni(x,v)\mapsto(x,va)\in\mathbb{U}\times\mathbb{F},$$

где произведение va было определено в самом начале (13.14). Таким образом, пара $(\mathbb{E},\mathbb{G})$ представляет собой группу преобразований. Структурная группа действует внутри каждого слоя и не перемешивает слои между собой.

В терминах представителей действие структурной группы на $\mathbb{E}$ описывается следующим образом. Если $(p,v)\in\mathbb{P}\times\mathbb{F}$ – представитель точки $u\in\mathbb{E}$, то представителем точки $ua\in\mathbb{E}$ является пара $(p,va)\in\mathbb{P}\times\mathbb{F}$.

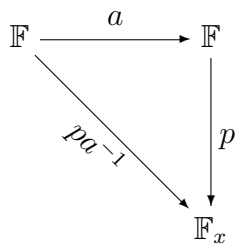
Замечание. В моделях математической физики наиболее распространенным примером типичного слоя $\mathbb{F}$ является векторное пространство, в котором задано представление структурной группы $\mathbb{G}$. В калибровочных моделях Янга–Миллса поля материи, которым соответствуют скалярные или спинорные поля в пространстве-времени, имеют дополнительный изотопический индекс. Этот индекс нумерует компоненты вектора в некотором векторном пространстве (типичном слое $\mathbb{F}$), в котором задано представление калибровочной группы (структурной группы $\mathbb{G}$). □

Замечание. В дальнейшем часть аргументов у ассоциированного расслоения $\mathbb{E}(\mathbb{M},\pi_{\mathbb{E}},\mathbb{F},\mathbb{G},\mathbb{P})$ мы, для краткости, будем опускать, оставляя лишь те, которые наиболее существенны в данный момент. □

Предложение 13.2.1. *Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение и $(\mathbb{F},\mathbb{G})$ – группа преобразований. Пусть $\mathbb{E}$ – расслоение, ассоциированное с $\mathbb{P}$. Для каждого $p\in\mathbb{P}$ и каждого $v\in\mathbb{F}$ обозначим через $p(v)$ образ элемента $(p,v)\in\mathbb{P}\times\mathbb{F}$ при действии естественной проекции $\mathbb{P}\times\mathbb{F}\to\mathbb{E}$. Тогда каждая точка $p\in\mathbb{P}$ определяет отображение типичного слоя $\mathbb{F}$ в слой $\mathbb{F}_x=\pi_{\mathbb{E}}^{-1}(x)$,*

$$p:\quad \mathbb{F}\ni v\mapsto p(v)\in\mathbb{F}_x, \tag{13.15}$$

где $x = \pi(p)$, *и* $(pa^{-1})(v) = p(va)$ *для всех* $p \in \mathbb{P}$, $a \in \mathbb{G}$ *и* $v \in \mathbb{F}$, *т.е. диаграмма*



коммутативна.

Доказательство. Следует прямо из определения, если учесть, что точки (pa, va) и (p, v) из произведения $\mathbb{P} \times \mathbb{F}$ проектируются в одну и ту же точку $\mathbb{E}$. □

Под изоморфизмом слоя $\mathbb{F}_x = \pi_{\mathbb{E}}^{-1}(x)$ на другой слой $\mathbb{F}_y = \pi_{\mathbb{E}}^{-1}(y)$, где $x, y \in \mathbb{M}$, мы понимаем диффеоморфизм, который представим в виде $q \circ p^{-1}$, где

$$\mathbb{F}_x \ni p: \ \mathbb{F} \to \mathbb{F}_x,$$
$$\mathbb{F}_y \ni q: \ \mathbb{F} \to \mathbb{F}_y.$$

В частности, автоморфизм слоя $\mathbb{F}_x$ – это отображение вида $q \circ p^{-1}$, где $p, q \in \mathbb{F}_x$. В этом случае $q = pa$ для некоторого $a \in \mathbb{G}$, так что любой автоморфизм слоя $\mathbb{F}_x$ представим в виде $p \circ a \circ p^{-1}$, где p – произвольная фиксированная точка в том же слое $\mathbb{F}_x$. Поэтому группа автоморфизмов каждого слоя $\mathbb{F}_x$ изоморфна структурной группе $\mathbb{G}$.

Опишем несколько примеров ассоциированных расслоений.

Пример 13.2.1. Рассмотрим группу преобразований $(\mathbb{G}, \mathbb{G})$, когда $\mathbb{G}$ действует на себе справа. Тогда ассоциированное расслоение изоморфно исходному главному расслоению, $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}, \mathbb{G}, \mathbb{P}) \simeq \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. В общем случае они могут не совпадать, т.к. существует произвол в параметризации фактор пространства $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} \mathbb{F}$. Таким образом произвольное главное расслоение ассоциировано само с собой. Если пару $(\mathbb{G}, \mathbb{G})$ рассматривать, как группу преобразований, действующую слева, то ассоциированное расслоение также будет изоморфно главному расслоению $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, но структурная группа будет действовать на $\mathbb{P}$ слева. □

Пример 13.2.2 (Цилиндр, лист Мёбиуса, бутылка Клейна). Рассмотрим главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}^2)$, где пространство расслоения реализовано в виде единичной окружности $\mathbb{P} \approx \mathbb{S}^1$ на евклидовой плоскости $\mathbb{R}^2$ и структурная группа которого состоит из двух элементов $\mathbb{Z}_2 = \{1, -1\}$ с обычным умножением. Действие структурной группы $\mathbb{Z}_2$ на пространстве расслоения определено отражением относительно центра окружности,

$$\mathbb{Z}_2 \ni \pm 1: \quad \mathbb{S}^1 \ni x \mapsto \pm x \in \mathbb{S}^1, \qquad x = (x^1, x^2) \in \mathbb{R}^2.$$

База расслоения $\pi(\mathbb{S}^1) = \mathbb{S}^1$ также является окружностью. Это ни что иное, как двулистное накрытие окружности окружностью, рассмотренное в примере 12.1.5. Построенное главное расслоение $\mathbb{S}^1(\mathbb{S}^1, \pi, \mathbb{Z}_2)$ нетривиально, т.к. не диффеоморфно прямому произведению $\mathbb{S}^1 \times \mathbb{Z}_2$.

Построим для этого главного расслоения три ассоциированных расслоения (см. рис.13.1).

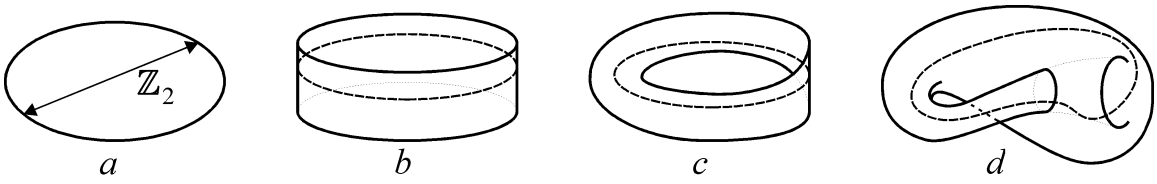


Рис. 13.1: Примеры ассоциированных расслоений. (a) Главное расслоение. (b) Цилиндр. (c) Лист Мёбиуса. (d) Бутылка Клейна.

1). Пусть группа $\mathbb{Z}_2$ действует на отрезок $\mathbb{F} = [-1, 1]$ тривиально,

$$\mathbb{Z}_2 \ni \pm 1 : \quad [-1, 1] \ni v \mapsto v \in [-1, 1].$$

Тогда ассоциированное расслоение представляет собой *цилиндр*, который является прямым произведением $\mathbb{S}^1 \times \mathbb{F}$.

2). Пусть группа $\mathbb{Z}_2$ действует на отрезок $\mathbb{F} = [-1, 1]$ нетривиально,

$$\mathbb{Z}_2 \ni \pm 1 : \quad [-1, 1] \ni v \mapsto \pm v \in [-1, 1].$$

Тогда ассоциированное расслоение представляет собой *лист Мёбиуса*. Лист Мёбиуса является нетривиальным расслоением, т.к. не имеет вида прямого произведения $\mathbb{S}^1 \times \mathbb{F}$.

3). Пусть группа $\mathbb{Z}_2$ действует на окружность $\mathbb{F} = \mathbb{S}^1$, отображая точки окружности относительно диаметра,

$$\mathbb{Z}_2 \ni \pm 1 : \quad \mathbb{S}^1 \ni \mathsf{e}^{iv} \mapsto \mathsf{e}^{\pm iv} \in \mathbb{S}^1,$$

где окружность $\mathbb{S}^1$ вложена в $\mathbb{R}^2$. Тогда ассоциированное расслоение представляет собой *бутылку Клейна*.

В первых двух примерах типичным слоем является многообразие с краем. Все три ассоциированных расслоения имеют глобальные сечения, несмотря но то, что во втором и третьем случаях ассоциированные расслоения не имеют вида прямого произведения. На рисунках они показаны пунктирными линиями. □

Приведенные примеры показывают, что многие хорошо известные примеры многообразий можно рассматривать, как ассоциированные расслоения. Приведем еще два примера, которые играют важную роль в дифференциальной геометрии.

Пример 13.2.3 (Касательное расслоение). Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{M})$, $\mathsf{dim}\,\mathbb{M} = n$, определенное в разделе 6.4. Оно состоит из упорядоченных наборов (реперов) $p = \{x, e_a\}$, $a = 1, \ldots, n$, где $x \in \mathbb{M}$ и $\{e_a\}$ – упорядоченный набор линейно независимых векторов касательного пространства, т.е. базис в $\mathbb{T}_x(\mathbb{M})$. При этом проекция определена отображением $\pi(p) := x \in \mathbb{M}$. Общая линейная группа $\mathbb{GL}(n, \mathbb{R})$ действует в $\mathbb{L}(\mathbb{M})$ справа так. Если $a = \{S_a{}^b\} \in \mathbb{GL}(n, \mathbb{R})$ и $p = \{x, e_a\}$ – репер, то преобразованный репер $pa = \{x, e'_a\}$ есть, по определению, репер, состоящий из векторов $e'_a = S_a{}^b e_b$. Мы записываем правое действие структурной группы как матричное умножение слева из-за принятого нами соглашения суммирования "с десяти до четырех". Отметим, что порядок записи является условным, $S_a{}^b e_b = e_b S_a{}^b$, так как суммирование проводится по одному нижнему и одному верхнему индексу. Группа $\mathbb{GL}(n, \mathbb{R})$ действует на расслоении реперов $\mathbb{L}(\mathbb{M})$ свободно и сохраняет слои, так

как $\pi(p) = \pi(q)$, $p, q \in \mathbb{L}(\mathbb{M})$ тогда и только тогда, когда $p = qa$ для некоторого $a \in \mathbb{GL}(n, \mathbb{R})$.

Построим теперь ассоциированное расслоение. Пусть $\Big(\mathbb{R}^n, \mathbb{GL}(n, \mathbb{R})\Big)$ – группа преобразований. Группа $\mathbb{GL}(n, \mathbb{R})$, по определению, действует в $\mathbb{R}^n$ следующим образом. Пусть b_a, $a = 1, \ldots, n$, – некоторый базис в $\mathbb{R}^n$. Если $a = \{S_a{}^b\} \in \mathbb{GL}(n, \mathbb{R})$ – матрица преобразования, то $b_a a = S_a{}^b b_b \in \mathbb{R}^n$. Покажем, что ассоциированное расслоение $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{R}^n, \mathbb{GL}(n, \mathbb{R}), \mathbb{L}\Big)$ можно отождествить с касательным расслоением $\mathbb{T}(\mathbb{M})$, определенным в разделе 3.6.2. Базы расслоений совпадают, а слои изоморфны как векторные пространства одинаковой размерности. Покажем, что слои ассоциированного расслоения $\pi_{\mathbb{E}}^{-1}(x)$ для всех $x \in \mathbb{M}$ можно рассматривать, как касательные пространства $\mathbb{T}_x(\mathbb{M})$. Пусть $p = \{x, e_a\}$ – репер в точке $x \in \mathbb{M}$. Он же является точкой в главном расслоении $p \in \mathbb{L}(\mathbb{M})$. Согласно предложению 13.2.1, каждому реперу ставится в соответствие невырожденное линейное отображение. Выберем такой репер, что

$$p: \quad \mathbb{R}^n \ni b_a \mapsto e_a \in \pi_{\mathbb{E}}^{-1}(x).$$

С другой стороны, по определению, набор векторов $\{e_a\}$ образует базис касательного пространства $\mathbb{T}_x(\mathbb{M})$. Поэтому $\pi_{\mathbb{E}}^{-1}(x) = \mathbb{T}_x(\mathbb{M})$.

В компонентах это отождествление выглядит следующим образом. Произвольный касательный вектор $X \in \mathbb{T}_x(\mathbb{M})$ в точке $x \in \mathbb{M}$ можно разложить по координатному базису и реперу: $X = X^\alpha \partial_\alpha = X^a e_a$, где $e_a = e^\alpha{}_a \partial_\alpha$. Компоненты вектора X связаны линейным преобразованием $X^a = X^\alpha e_\alpha{}^a$, где $e_\alpha{}^a$ – компоненты обратного репера. Структурная группа $\mathbb{GL}(n, \mathbb{R})$ действует на латинские индексы, не меняя касательного вектора, $X = X^a e_a = X'^a e'_a$, где $X'^a = X^b S_b^{-1a}$ и $e'_a = S_a{}^b e_b$. При преобразовании координат преобразуются греческие индексы с помощью взаимно обратных матриц Якоби, $X = X^\alpha \partial_\alpha = X^{\alpha'} \partial_{\alpha'}$, где $X^{\alpha'} = X^\alpha \partial_\alpha x^{\alpha'}$ и $\partial_{\alpha'} = \partial_{\alpha'} x^\alpha \partial_\alpha$. В каждой точке $x \in \mathbb{M}$ матрицы Якоби также принадлежат общей линейной группе $\mathbb{GL}(n, \mathbb{R})$. □

Пример 13.2.4 (Тензорные расслоения). Пусть $\mathbb{T}^r_s$ – тензорное пространство типа (r, s) над векторным пространством $\mathbb{R}^n$. Общая группа линейных преобразований $\mathbb{GL}(n, \mathbb{R})$ действует в $\mathbb{T}^r_s$ обычным образом:

$$e_{a_1} \otimes \ldots \otimes e_{a_r} \otimes e^{b_1} \otimes \ldots \otimes e^{b_s} \mapsto$$
$$\mapsto S_{a_1}{}^{c_1} e_{c_1} \otimes \ldots \otimes S_{a_r}{}^{c_r} e_{c_r} \otimes e^{d_1} S^{-1 b_1}_{d_1} \otimes \ldots \otimes e^{d_s} S^{-1 b_s}_{d_s},$$

где e_a – базис в $\mathbb{R}^n$, e^a – базис в сопряженном пространстве $\mathbb{R}^{n*}$ и $\{S_a{}^b\} \in \mathbb{GL}(n, \mathbb{R})$. $a, b, ... = 1, \ldots, n$. Расслоение $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{T}^r_s, \mathbb{GL}(n, \mathbb{R}), \mathbb{L}\Big)$, ассоциированное с расслоением реперов $\mathbb{L}(\mathbb{M})$, так же как и касательное расслоение отождествляется с тензорным расслоением $\mathbb{T}^r_s(\mathbb{M})$ типа (r, s), определенным в разделе 3.7. □

В разделе 6.1 было дано определение векторного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(n, \mathbb{R}), \mathbb{P}\Big)$, $\mathsf{dim}\, \mathbb{V} = n$, ассоциированного с расслоением реперов $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{GL}(n, \mathbb{R})\Big)$. Дадим координатное описание векторных расслоений. Рассмотрим координатное покрытие базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$. Каждая область предполагается односвязной, и, следовательно, над каждой областью $\mathbb{U}_i$ расслоение является тривиальным, т.е. существуют диффеоморфизмы (тривиализации)

$$\chi_i: \quad \pi_{\mathbb{E}}^{-1}(\mathbb{U}_i) \to \mathbb{U}_i \times \mathbb{V}.$$

Отсюда следует, что определено отображение

$$f_{ji} := \chi_j \circ \chi_i^{-1}: \quad \mathbb{V} \to \mathbb{V},$$

которое называется *функциями склейки*. Если базис векторного пространства $\mathbb{V}$, $\dim \mathbb{V} = n$, фиксирован, то функции склейки принимают значения в группе невырожденных матриц:

$$f_{ji}: \quad \mathbb{U}_i \cap \mathbb{U}_j \to \mathbb{GL}(n, \mathbb{R}). \tag{13.16}$$

Теорема 13.2.1. *Пусть задано векторное расслоение* $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(n, \mathbb{R}), \mathbb{P}\Big)$, *координатное покрытие базы* $\mathbb{M} = \bigcup_i \mathbb{U}_i$ *и соответствующие тривиализации* χ_i. *Тогда однозначно определен набор функций склеек (13.16), обладающих свойствами:*

$$f_{ij} = f_{ji}^{-1}, \qquad \forall \;\; x \in \mathbb{U}_i \cap \mathbb{U}_j, \tag{13.17}$$

$$f_{ij} f_{jk} f_{ki} = \mathsf{id}\,, \qquad \forall \;\; x \in \mathbb{U}_i \cap \mathbb{U}_j \cap \mathbb{U}_k. \tag{13.18}$$

Обратно. По любому координатному покрытию $\mathbb{M} = \bigcup_i \mathbb{U}_i$ *и по любому набору отображений (13.16) со свойствами (13.17), (13.18) Можно построить векторное расслоение* $\mathbb{E}'\Big(\mathbb{M}, \pi'_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(n, \mathbb{R}), \mathbb{P}\Big)$. *Если набор функций склеек был построен по векторному расслоению* $\mathbb{E}$, *то расслоения* $\mathbb{E}'$ *и* $\mathbb{E}$ *изоморфны.*

Доказательство. Аналогично доказательству теоремы 13.1.2 для главных расслоений. □

Следствие. Главное расслоение $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{G}(n, \mathbb{R})\Big)$ и ассоциированные с ним векторные расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}(n, \mathbb{R}), \mathbb{P}\Big)$ определяют эквивалентные склеивающие коциклы.

Доказательство. См., например, [**?**]. □

Следующая конструкция показывает, что коциклы можно использовать для построения новых ассоциированных расслоений.

Пример 13.2.5. Векторное расслоение $\mathbb{E}(\mathbb{M}, \pi, \mathbb{V})$, рассмотренное в разделе 6.1, ассоциировано с расслоением реперов $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{GL}(n, \mathbb{R})\Big)$, где $n = \dim \mathbb{V}$. □

Предложение 13.2.2. *Пусть дано векторное расслоение* $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V})$, *ассоциированное с расслоением реперов* $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{GL}(n, \mathbb{R})\Big)$, *где* $n = \dim \mathbb{V}$, *с коциклом* $\{f_{ji}\}$, *соответствующим некоторому координатному покрытию базы* $\mathbb{M} = \bigcup_i \mathbb{U}_i$. *Тогда существует* детерминантное расслоение $|\mathbb{E}|(\mathbb{M}, \pi_{|\mathbb{E}|}, \mathbb{R})$, *определенное коциклом* $\{\det f_{ji}\}$, *которое ассоциировано с тем же расслоением реперов* $\mathbb{P}$.

Доказательство. Проверка свойств (13.17), (13.18). Кроме того, отображение $f_{ji} \mapsto \det f_{ji}$ определяет одномерное представление группы $\mathbb{GL}(n, \mathbb{R})$ в $\mathbb{R}$. □

Пример 13.2.6. Скалярные плотности степени 1, введенные в разделе 3.5, являются сечениями детерминантного расслоения. □

Вернемся к общему случаю. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение, и $\mathbb{H}$ – замкнутая подгруппа в $\mathbb{G}$. Пара $(\mathbb{G}/\mathbb{H}, \mathbb{G})$, где $\mathbb{G}/\mathbb{H}$ – пространство правых смежных классов, является группой преобразований. Поэтому определено ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ со стандартным слоем $\mathbb{G}/\mathbb{H}$. С другой стороны, группа $\mathbb{H}$, будучи подгруппой в структурной группе $\mathbb{G}$, действует на $\mathbb{P}$ справа. Поэтому определено фактор пространство $\mathbb{P}/\mathbb{H}$. Между построенными ассоциированным расслоением и фактор пространством существует связь.

Предложение 13.2.3. *Расслоение* $\mathbb{E} = \mathbb{P} \times_{\mathbb{G}} (\mathbb{G}/\mathbb{H})$*, ассоциированное с* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$*, со стандартным слоем* $\mathbb{G}/\mathbb{H}$ *можно отождествить с* $\mathbb{P}/\mathbb{H}$ *так. Пусть* $(p, \mathbb{H}a) \in \mathbb{P} \times (\mathbb{G}/\mathbb{H})$*, где* $a \in \mathbb{G}$*, – представитель элемента ассоциированного расслоения* $\mathbb{E}$*. Отождествим его с элементом* $pa \in \mathbb{P}$*, который является представителем некоторого элемента из фактор пространства* $\mathbb{P}/\mathbb{H}$*.*

Доказательство. Каждый представитель однозначно определяет элемент соответствующего пространства. Поэтому отождествление представителей приводит к отождествлению точек ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ и фактор пространства $\mathbb{P}/\mathbb{H}$. Легко проверить, что это отождествление не зависит от выбора представителей. □

Следствие. Четверка $\mathbb{P}(\mathbb{E}, \nu, \mathbb{H})$, где $\nu : \mathbb{P} \to \mathbb{P}/\mathbb{H}$ – естественная проекция, является главным расслоением, базой которого является ассоциированное расслоение $\mathbb{E}$, а структурной группой $\mathbb{H} \subset \mathbb{G}$.

Доказательство. Из предложения 13.2.3 следует, что фактор пространство $\mathbb{P}/\mathbb{H}$ является многообразием и поэтому может являться базой некоторого расслоения. Действие группы $\mathbb{H}$ на $\mathbb{P}$ не двигает точки базы $\mathbb{E}$ по построению. Локальная тривиальность $\mathbb{P}(\mathbb{E}, \nu, \mathbb{H})$ следует из локальной тривиальности расслоений $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ и $\mathbb{G}(\mathbb{G}/\mathbb{H}, \mathbb{H})$. Действительно, пусть $\mathbb{U}$ – окрестность в $\mathbb{M}$ такая, что $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times (\mathbb{G}/\mathbb{H})$ и $\mathbb{V}$ – окрестность в $\mathbb{G}/\mathbb{H}$, для которой $\pi_{\mathbb{G}}^{-1}(\mathbb{V}) \approx \mathbb{V} \times \mathbb{H}$, где $\pi_{\mathbb{G}} : \mathbb{G} \to \mathbb{G}/\mathbb{H}$ – естественная проекция. Пусть $\mathbb{W}$ – открытое подмножество в $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \subset \mathbb{E}$, которое соответствует прямому произведению $\mathbb{U} \times \mathbb{V}$ при отождествлении $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times (\mathbb{G}/\mathbb{H})$. Тогда $\nu^{-1}(\mathbb{W}) \approx \mathbb{W} \times \mathbb{H}$. □

Выпишем связь размерностей многообразий, которые участвовали в приведенной выше конструкции:

$$\begin{aligned} \dim \mathbb{P}(\mathbb{M}, \pi, \mathbb{G}) &= \dim \mathbb{M} + \dim \mathbb{G}, \\ \dim \mathbb{G} &= \dim \mathbb{G}/\mathbb{H} + \dim \mathbb{H}, \\ \dim \mathbb{E} &= \dim \mathbb{M} + \dim \mathbb{G}/\mathbb{H}, \\ \dim \mathbb{P}(\mathbb{E}, \nu, \mathbb{H}) &= \dim \mathbb{M} + \dim \mathbb{G}/\mathbb{H} + \dim \mathbb{H} = \dim \mathbb{P}(\mathbb{M}, \pi, \mathbb{G}). \end{aligned}$$

Теперь обсудим возможность продолжения сечения ассоциированного расслоения, которое задано на некотором подмножестве базы $\mathbb{N} \subset \mathbb{M}$. Для формулировки теоремы нам понадобится

Определение. Отображение $f : \mathbb{N} \to \mathbb{E}$ подмножества $\mathbb{N}$ многообразия $\mathbb{M}$ в многообразие $\mathbb{E}$ называется *дифференцируемым на* $\mathbb{N}$, если для каждой точки $x \in \mathbb{N}$ существует дифференцируемое отображение $f_x : \mathbb{U}_x \to \mathbb{E}$, где $\mathbb{U}_x \subset \mathbb{M}$ – некоторая окрестность точки $x \in \mathbb{M}$, такое, что $f_x = f$ на пересечении $\mathbb{U}_x \cap \mathbb{N}$. Если задано ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ и некоторое подмножество базы $\mathbb{N} \subset \mathbb{M}$, то *сечением на* $\mathbb{N}$ называется дифференцируемое отображение $\sigma : \mathbb{N} \to \mathbb{E}$ такое, что $\pi_{\mathbb{E}} \circ \sigma = \mathsf{id}_{\mathbb{M}}$. □

Замечание. В данном определении подмножество $\mathbb{N} \subset \mathbb{M}$ может быть произвольным и в общем случае не является подмногообразием в $\mathbb{M}$. □

Пример 13.2.7. Пусть $f : \mathbb{U} \to \mathbb{E}$ – дифференцируемое отображение некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ в $\mathbb{E}$. Тогда сужение $f|_{\mathbb{N}}$ на любое подмножество $\mathbb{N} \subset \mathbb{U}$ является дифференцируемым на $\mathbb{N}$. □

Теорема 13.2.2. *Пусть* $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ *– ассоциированное расслоение такое, что типичный слой диффеоморфен евклидову пространству,* $\mathbb{F} \approx \mathbb{R}^n$*. Пусть* $\mathbb{N}$ *– замкнутое подмножество (возможно, пустое) в базе* $\mathbb{M}$*. Тогда любое сечение* $\sigma : \mathbb{N} \to \mathbb{E}$*, определенное на* $\mathbb{N}$*, может быть продолжено до глобального сечения, определенного на всем* $\mathbb{M}$*. В частности, если подмножество* $\mathbb{N}$ *пусто, то на ассоциированном расслоении* $\mathbb{E}$ *с типичным слоем* $\mathbb{F} \approx \mathbb{R}^n$ *существует глобальное сечение* $\sigma : \mathbb{M} \to \mathbb{E}$*.*

Доказательство. Явно строится сечение, при этом используется паракомпактность многообразия. Детали можно найти, например, в [**?**]. □

Следствие. На любом многообразии $\mathbb{M}$ существует векторное поле, например, тождественно равное нулю. Более того, если векторное поле задано на произвольном замкнутом подмножестве $\mathbb{N} \subset \mathbb{M}$, то его всегда можно продолжить до векторного поля на всем многообразии $\mathbb{M}$. Замкнутость подмножества $\mathbb{N}$ существенна. Пример 3.6.5 показывает, что, если векторное поле задано на открытом подмножестве, то продолжение может не существовать. □

Замечание. Данная теорема показывает, что ассоциированные расслоения существенно отличаются от главных. Согласно предложению 13.1.1 главное расслоение $\mathbb{P}$ допускает глобальное сечение тогда и только тогда, когда оно тривиально. Если ассоциированное расслоение имеет вид прямого произведения, $\mathbb{E} = \mathbb{M} \times \mathbb{F}$, то оно, очевидно, также допускает глобальные сечения. Последняя теорема утверждает, что ассоциированное расслоение $\mathbb{E}$ имеет глобальные сечения даже если оно не имеет вида прямого произведения $\mathbb{M} \times \mathbb{F}$. Однако, взамен мы требуем, чтобы типичный слой был диффеоморфен евклидову пространству, $\mathbb{F} \approx \mathbb{R}^n$. □

Пример 13.2.8. Заменим единичный отрезок $[-1, 1]$ для листа Мёбиуса из примера 13.2.2 на вещественную прямую $\mathbb{R}$. Тогда мы попадаем в условия теоремы 13.2.2, и, следовательно лист Мёбиуса имеет глобальное сечение. □

Определение. Ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ называется *тривиальным*, если соответствующее главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ тривиально. □

Из данного определения следует, что тривиальное ассоциированное расслоение всегда имеет вид прямого произведения, $\mathbb{E} \approx \mathbb{M} \times \mathbb{F}$, т.е. оно тривиально как расслоение. Обратное утверждение неверно.

Пример 13.2.9. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – нетривиальное главное расслоение и $\mathbb{F}$ – произвольное многообразие, на котором структурная группа $\mathbb{G}$ действует тривиально. Тогда ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{F}, \mathbb{G}, \mathbb{P})$ тривиально как расслоение, т.е. $\mathbb{E} \approx \mathbb{M} \times \mathbb{F}$, однако оно не является тривиальным ассоциированным расслоением. □

В комплексном случае справедливо

Предложение 13.2.4. *Векторное расслоение* $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{GL}(n, \mathbb{R}), \mathbb{P}\Big)$*,* $\mathsf{dim}_{\,\mathbb{C}}\mathbb{V} = n$*, имеет вид прямого произведения* $\mathbb{E} \simeq \mathbb{M} \times \mathbb{V}$ *тогда и только тогда, когда оно имеет* n *сечений (того же класса гладкости), линейно независимых над каждой точкой базы.*

Доказательство. См., например, [**?**]. □

13.3 Отображение расслоений

Рассмотрим два главных расслоения $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ и $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$. Оба расслоения являются группами преобразований $(\mathbb{P}_1, \mathbb{G}_1)$ и $(\mathbb{P}_2, \mathbb{G}_2)$. Поэтому при определении гомоморфизма расслоений используется понятие эквивариантного отображения, введенного в разделе 10.4.

Определение. *Гомоморфизмом* расслоений $f : \ \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1) \to \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ называется эквивариантное отображение $f = f_{\mathbb{P}} \times f_{\mathbb{G}}$ (пара отображений),

$$f_{\mathbb{P}} : \ \mathbb{P}_1 \to \mathbb{P}_2,$$
$$f_{\mathbb{G}} : \ \mathbb{G}_1 \to \mathbb{G}_2,$$

где $f_{\mathbb{P}}$ – дифференцируемое отображение пространств расслоений и $f_{\mathbb{G}}$ – гомоморфизм структурных групп, таких, что $f_{\mathbb{P}}(pa) = f_{\mathbb{P}}(p) f_{\mathbb{G}}(a)$ для всех $p \in \mathbb{P}_1$ и $a \in \mathbb{G}_1$, т.е. диаграмма

$$\begin{array}{ccc} \mathbb{P}_1 \times \mathbb{G}_1 & \xrightarrow{f_{\mathbb{P}} \times f_{\mathbb{G}}} & \mathbb{P}_2 \times \mathbb{G}_2 \\ \Big\downarrow a & & \Big\downarrow f_{\mathbb{G}}(a) \\ \mathbb{P}_1 & \xrightarrow{\qquad f_{\mathbb{P}} \qquad} & \mathbb{P}_2 \end{array}$$

коммутативна.

Расслоения $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ и $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ *изоморфны*, если $f_{\mathbb{P}}$ – диффеоморфизм пространств расслоений и $f_{\mathbb{G}}$ – изоморфизм структурных групп. Если $\mathbb{P}_1 = \mathbb{P}_2 = \mathbb{P}$, то изоморфизм расслоений называется *автоморфизмом* расслоения $\mathbb{P}$. □

Из приведенной диаграммы следует, что отображение $f_{\mathbb{P}} : \ \mathbb{P}_1 \to \mathbb{P}_2$ отображает каждый слой $p\mathbb{G}_1 \subset \mathbb{P}_1$ расслоения $\mathbb{P}_1$ в слой $f_{\mathbb{P}}(p)\mathbb{G}_2$ из $\mathbb{P}_2$ и поэтому индуцирует дифференцируемое отображение баз

$$f_{\mathbb{M}} : \ \ \mathbb{M}_1 \ni \pi_1(p) \mapsto \pi_2\Big(f_{\mathbb{P}}(p)\Big) \in \mathbb{M}_2$$

Легко проверить, что это отображение не зависит от выбора точки в слое $p \in \pi_1^{-1}(x)$.

Если $f : \ \mathbb{P}_1 \to \mathbb{P}_2$ – изоморфизм расслоений, то индуцированное отображение баз $f_{\mathbb{M}}$ является диффеоморфизмом и существует обратный изоморфизм $f^{-1} : \ \mathbb{P}_2 \to \mathbb{P}_1$.

Пример 13.3.1 (Вертикальный автоморфизм). Пусть $a(x)$ – произвольная дифференцируемая функция на базе $\mathbb{M}$ со значениями в структурной группе $\mathbb{G}$. Умножим каждую точку главного расслоения $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ на $a\Big(\pi(p)\Big)$, что соответствует повороту каждого слоя. В результате получим главное расслоение $\mathbb{P}'(\mathbb{M}, \pi, \mathbb{G})$, которое изоморфно исходному. Этот автоморфизм задается функциями:

$$f_{\mathbb{P}} : \ \mathbb{P} \ni p \mapsto pa\Big(\pi(p)\Big) \in \mathbb{P}'$$
$$f_{\mathbb{G}} = \mathsf{id}_{\,\mathbb{G}} : \ \mathbb{G} \ni b \mapsto b \in \mathbb{G}.$$

При этом отображение баз является тождественным преобразованием, $f_{\mathbb{M}} = \mathsf{id}$. □

Определение. Гомоморфизм расслоений называется *вложением* или *инъекцией*, если индуцированное отображение баз $f_{\mathbb{M}} : \ \mathbb{M}_1 \to \mathbb{M}_2$ есть вложение многообразий и

$f_{\mathbb{G}}: \mathbb{G}_1 \to \mathbb{G}_2$ – мономорфизм (инъективный гомоморфизм). Отождествляя пространство первого расслоения $\mathbb{P}_1$ с его образом $f_{\mathbb{P}}(\mathbb{P}_1)$, $\mathbb{G}_1$ с $f_{\mathbb{G}}(\mathbb{G}_1)$ и $\mathbb{M}_1$ с $f_{\mathbb{M}}(\mathbb{M}_1)$, мы говорим, что $\mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1)$ есть *подрасслоение* для $\mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$. Если, кроме того, $\mathbb{M}_1 = \mathbb{M}_2 = \mathbb{M}$ и индуцированное отображение баз $f_{\mathbb{M}} = \mathsf{id}_{\mathbb{M}}$ есть тождественное отображение, то гомоморфизм расслоений $f: \mathbb{P}_1(\mathbb{M}, \pi_1, \mathbb{G}_1) \to \mathbb{P}_2(\mathbb{M}, \pi_2, \mathbb{G}_2)$ называется *редукцией* структурной группы $\mathbb{G}_2$ главного расслоения $\mathbb{P}_2(\mathbb{M}, \pi_2, \mathbb{G}_2)$ к подгруппе $\mathbb{G}_1$. Само расслоение $\mathbb{P}_1(\mathbb{M}, \pi_1, \mathbb{G}_1)$ называется *редуцированным расслоением.* Если задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и подгруппа Ли $\mathbb{H}$ в $\mathbb{G}$, то мы говорим, что *структурная группа* $\mathbb{G}$ *редуцируема к* $\mathbb{H}$, если существует редуцированное главное расслоение $\mathbb{P}'(\mathbb{M}, \pi', \mathbb{H})$. □

Замечание. Поскольку при вложении главных расслоений $f_{\mathbb{M}}$ – вложение и $f_{\mathbb{G}}$ – мономорфизм, то

$$\mathsf{dim}\,\mathbb{M}_1 \leq \mathsf{dim}\,\mathbb{M}_2, \quad \text{и} \quad \mathsf{dim}\,\mathbb{G}_1 \leq \mathsf{dim}\,\mathbb{G}_2.$$

Мы не требуем, чтобы подгруппа $\mathbb{H}$ была замкнутой в $\mathbb{G}$. Эта общность необходима в теории связностей. □

Теорема 13.3.1. *Структурная группа* $\mathbb{G}$ *главного расслоения* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *редуцируема к подгруппе Ли* $\mathbb{H}$ *тогда и только тогда, когда существует координатное покрытие базы* $\mathbb{M} = \bigcup_i \mathbb{U}_i$ *такое, что все функции перехода* a_{ji} *принимают значение в* $\mathbb{H}$.

Доказательство. Доказательство проводится путем построения редуцированного расслоения. См., например, [**?**]. □

Приведем еще один критерий редуцируемости расслоений.

Теорема 13.3.2. *Структурная группа* $\mathbb{G}$ *главного расслоения* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *редуцируема к подгруппе* $\mathbb{H}$ *тогда и только тогда, когда ассоциированное расслоение* $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P}) = \mathbb{P}/\mathbb{H}$ *(см. предложение 13.2.3) допускает глобальное сечение* $\sigma: \mathbb{M} \to \mathbb{E}$. *Между редуцированными главными расслоениями* $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ *и сечениями* σ *существует естественное взаимно однозначное соответствие.*

Доказательство. См., например, [**?**]. □

Следствие. Структурная группа $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ редуцируема к единичному элементу $e \in \mathbb{G}$ тогда и только тогда, когда оно имеет глобальное сечение и, следовательно, тривиально, $\mathbb{P} \simeq \mathbb{M} \times \mathbb{G}$.

Замечание. Если структурная группа состоит только из единичного элемента, то пространство редуцированного главного расслоения $\mathbb{P}'(\mathbb{M}, \pi, e)$ естественным образом отождествляется с базой, $\mathbb{P}' = \mathbb{M}$. Поэтому существование отображения $f_{\mathbb{P}}: \mathbb{P}' = \mathbb{M} \to \mathbb{P}$ для редуцированного расслоения эквивалентно существованию глобального сечения для $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. □

Теорема о существовании глобальных сечений на ассоциированном расслоении позволяет по новому доказать и взглянуть на существование римановой метрики на произвольном многообразии. Для формулировки результата нам понадобится дополнительное утверждение.

Из разложения Ивасавы [**?**] следует, что любая связная группа Ли диффеоморфна прямому произведению, $\mathbb{G} \approx \mathbb{H} \times \mathbb{R}^n$, где $\mathbb{H}$ – максимальная компактная подгруппа в $\mathbb{G}$ и $\mathbb{R}^n$ – евклидово пространство размерности $n = \mathsf{dim}\,\mathbb{G} - \mathsf{dim}\,\mathbb{H}$ [].

Предложение 13.3.1. *Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение со связной структурной группой Ли $\mathbb{G}$. Тогда структурная группа редуцируема к максимальной компактной подгруппе $\mathbb{H}\subset\mathbb{G}$.*

Доказательство. Рассмотрим ассоциированное расслоение $\mathbb{E}(\mathbb{M},\pi_{\mathbb{E}},\mathbb{G}/\mathbb{H},\mathbb{G},\mathbb{P})=\mathbb{P}/\mathbb{H}$, где $\mathbb{H}$ – максимальная компактная подгруппа в $\mathbb{G}$. Поскольку типичный слой $\mathbb{G}/\mathbb{H}$ диффеоморфен евклидову пространству (разложение Ивасавы), то, согласно теореме 13.2.2, ассоциированное расслоение допускает глобальное сечение. Поэтому из теоремы 13.3.2 вытекает, что структурная группа главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ редуцируема к максимальной компактной подгруппе $\mathbb{H}$. □

Пример 13.3.2 (Существование римановой метрики). Пусть $\mathbb{L}(\mathbb{M})$ – расслоение линейных реперов над многообразием $\mathbb{M}$, $\dim\mathbb{M}=n$. Известно, что максимальной компактной подгруппой в группе общих линейных преобразований $\mathbb{GL}(n,\mathbb{R})$ является группа вращений $\mathbb{O}(n)$ размерности $\frac{1}{2}n(n-1)$. В силу разложения Ивасавы однородное пространство $\mathbb{GL}(n,\mathbb{R})/\mathbb{O}(n)$ диффеоморфно евклидову пространству $\mathbb{R}^d$ размерности

$$d=n^2-\frac{n(n-1)}{2}=\frac{n(n+1)}{2}.$$

Из теоремы 13.2.2 следует, что ассоциированное расслоение $\mathbb{E}=\mathbb{L}(\mathbb{M})/\mathbb{O}(n)$ со слоем $\mathbb{GL}(n,\mathbb{R})/\mathbb{O}(n)$ имеет глобальное сечение. Поэтому структурная группа $\mathbb{GL}(n,\mathbb{R})$ может быть редуцирована к группе вращений $\mathbb{O}(n)$ как следствие теоремы 13.3.2.

Пусть в евклидовом пространстве $\mathbb{R}^n$ в декартовых координатах задано скалярное произведение

$$(\widetilde{X},\widetilde{Y})=\widetilde{X}^a\widetilde{Y}^b\delta_{ab},\qquad \widetilde{X},\widetilde{Y}\in\mathbb{R}^n, \tag{13.19}$$

где δ_{ab} – евклидова метрика (2.8). Это скалярное произведение инвариантно относительно вращений евклидова пространства $\mathbb{O}(n)$.

Покажем, что каждая редукция структурной группы $\mathbb{GL}(n,\mathbb{R})$ расслоения реперов $\mathbb{L}(\mathbb{M})$ порождает риманову метрику g на $\mathbb{M}$. Пусть $\mathbb{O}(\mathbb{M}):=\mathbb{O}\big(\mathbb{M},\pi,\mathbb{O}(n)\big)$ – редуцированное подрасслоение для расслоения реперов $\mathbb{L}(\mathbb{M})$. Типичный слой редуцированного расслоения состоит из реперов вида $p=\{x,S_a{}^be_b\}$, где $S_a{}^b\in\mathbb{O}(n)$ – произвольная матрица вращений. Если мы рассматриваем репер $p\in\mathbb{L}(\mathbb{M})$ как линейный изоморфизм из $\mathbb{R}^n$ на касательное пространство $\mathbb{T}_x(\mathbb{M})$, где $x=\pi(p)$, то каждый репер из редуцированного расслоения $p\in\mathbb{O}(\mathbb{M})$ определяет скалярное произведение в касательном пространстве $\mathbb{T}_x(\mathbb{M})$ по формуле

$$g(X,Y)=(p^{-1}X,p^{-1}Y),\qquad X,Y\in\mathbb{T}_x(\mathbb{M}). \tag{13.20}$$

Инвариантность скалярного произведения (13.19) относительно вращений $\mathbb{O}(n)$ влечет за собой независимость $g(X,Y)$ от выбора репера $p\in\mathbb{Q}$. Таким образом, мы доказали, что каждое многообразие допускает риманову метрику.

В компонентах определение (13.20) имеет хорошо знакомый вид

$$g(X,Y)=X^\alpha Y^\beta g_{\alpha\beta}=\widetilde{X}^a\widetilde{Y}^b\delta_{ab},\qquad g_{\alpha\beta}=e_\alpha{}^ae_\beta{}^b\delta_{ab},$$

где $\widetilde{X}^a=X^\alpha e_\alpha{}^a$ и $\widetilde{Y}^a=Y^\alpha e_\alpha{}^a$. В таком виде скалярное произведение уже встречалось в разделе 7.9 (без знаков тильды).

Верно также обратное утверждение. Пусть на $\mathbb{M}$ задана риманова метрика g. Пусть $\mathbb{O}(\mathbb{M})$ – подмножество в расслоении реперов $\mathbb{L}(\mathbb{M})$, состоящее из реперов $p=\{x,e_a\}$, которые ортонормальны относительно g, т.е.

$$(e_a,e_b)=e^\alpha{}_ae^\beta{}_bg_{\alpha\beta}=\delta_{ab}.$$

Если репер $p \in \mathbb{L}(\mathbb{M})$ рассматривается как изоморфизм из $\mathbb{R}^n$ в $\mathbb{T}_x(\mathbb{M})$, то p принадлежит редуцированному подрасслоению $\mathbb{O}(\mathbb{M})$ тогда и только тогда, когда $(\widetilde{X}, \widetilde{Y}) = g(p\widetilde{X}, p\widetilde{Y})$ для всех $\widetilde{X}, \widetilde{Y} \in \mathbb{R}^n$. Легко проверить, что $\mathbb{O}(\mathbb{M})$ образует редуцированное подрасслоение в $\mathbb{L}(\mathbb{M})$ над базой $\mathbb{M}$ со структурной группой $\mathbb{O}(n)$. Расслоение $\mathbb{O}(\mathbb{M})$ называется *расслоением ортонормальных реперов* над $\mathbb{M}$. Каждый элемент из $\mathbb{O}(\mathbb{M})$ есть *ортонормальный репер.*

Чтобы подчеркнуть принадлежность вектора евклидову пространству, в настоящем примере был использован знак тильды. Поскольку репер устанавливает изоморфизм $\mathbb{R}^n$ и $\mathbb{T}_x(\mathbb{M})$, то эти пространства мы отождествим и опустим знак тильды, считая, что X^α и X^a – это компоненты одного касательного вектора $X \in \mathbb{T}_x(\mathbb{M})$ относительно координатного базиса и репера. Чтобы их не путать, мы используем буквы греческого и латинского алфавитов. □

В рассмотренном примере мы не только доказали существование римановой метрики на произвольном многообразии, но и следующее утверждение.

Теорема 13.3.3. *Существует взаимно однозначное соответствие между римановыми метриками на многообразии* $\mathbb{M}$ *и редукциями структурной группы* $\mathbb{GL}(n, \mathbb{R})$ *расслоения реперов* $\mathbb{L}(\mathbb{M})$ *к группе вращений* $\mathbb{O}(n)$

В начале настоящего раздела мы установили, что гомоморфизм расслоений индуцирует дифференцируемое отображение баз расслоений. Теперь мы рассмотрим обратную задачу: в какой степени отображение некоторого многообразия в базу заданного расслоения может быть сопоставлено некоторому гомоморфизму расслоений ? Ответ на этот вопрос дает

Теорема 13.3.4. *Пусть дано главное расслоение* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *и дифференцируемое отображение многообразий* $f_{\mathbb{N}} : \mathbb{N} \to \mathbb{M}$. *Тогда существует единственное с точностью до изоморфизма главное расслоение* $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ *с гомоморфизмом* $f : \mathbb{Q} \to \mathbb{P}$, *индуцирующим отображение баз* $f_{\mathbb{N}} : \mathbb{N} \to \mathbb{M}$ *и соответствующим тождественному автоморфизму структурной группы* $\mathbb{G}$.

Доказательство. В прямом произведении $\mathbb{N} \times \mathbb{P}$ рассмотрим подмножество $\mathbb{Q}$, состоящее из точек $(y, p) \in \mathbb{N} \times \mathbb{P}$ таких, что $f_{\mathbb{N}}(y) = \pi(p)$. Определим действие структурной группы $\mathbb{G}$ справа на построенном множестве $\mathbb{Q}$:

$$\mathbb{G} \ni a : \quad \mathbb{Q} \ni (y, p) \mapsto (y, p)a := (y, pa) \in \mathbb{Q}.$$

Это отображение не зависит от точки слоя $\pi^{-1}\big(\pi(p)\big)$. Нетрудно проверить, что $\mathbb{G}$ действует свободно на $\mathbb{Q}$ и что множество $\mathbb{Q}$ представляет собой главное расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ с базой $\mathbb{N}$, структурной группой $\mathbb{G}$ и проекцией $\pi_{\mathbb{Q}} : (y, p) \mapsto y$.

Единственность. Пусть $\mathbb{Q}'(\mathbb{N}, \pi'_{\mathbb{Q}}, \mathbb{G})$ – другое главное расслоение с базой $\mathbb{N}$ и структурной группой $\mathbb{G}$ и $f' : \mathbb{Q}' \to \mathbb{P}$ – гомоморфизм, индуцирующий заданное отображение баз $f_{\mathbb{N}} : \mathbb{N} \to \mathbb{M}$ и соответствующий тождественному автоморфизму структурной группы $\mathbb{G}$. Определим отображение $\mathbb{Q}'$ на $\mathbb{Q}$

$$\tilde{f} : \quad \mathbb{Q}' \ni p' \mapsto \Big(\pi'_{\mathbb{Q}}(p'), f'(p')\Big) \in \mathbb{Q}.$$

Тогда отображение $\tilde{f} : \mathbb{Q}' \to \mathbb{Q}$ – изоморфизм расслоений, индуцирующий тождественное преобразование базы $\mathbb{N}$ и соответствующий тождественному автоморфизму структурной группы $\mathbb{G}$. □

Определение. Расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ в утверждении теоремы 13.3.4 называется *расслоением, индуцированным отображением* $f_{\mathbb{N}} : \mathbb{N} \to \mathbb{M}$ *из главного расслоения* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, или просто *индуцированным расслоением*. Оно обозначается $f_{\mathbb{N}}^{-1}\mathbb{P}$. □

Замечание. Если отображение $f_{\mathbb{N}}$ является вложением, то главное расслоение $\mathbb{Q}(\mathbb{N}, \pi_{\mathbb{Q}}, \mathbb{G})$ есть подрасслоение для $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. □

Пример 13.3.3. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и некоторую окрестность базы $\mathbb{U} \subset \mathbb{M}$ тогда $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$ есть подрасслоение в $\mathbb{P}$. Оно же является индуцированным расслоением $f_{\mathbb{U}}^{-1}\mathbb{P}$, где $f : \mathbb{U} \to \mathbb{P}$ – вложение. □

Глава 14

Связности на главных и ассоциированных расслоениях

Теория связностей на расслоениях играет исключительно важную роль в моделях математической физики, т.к. позволяет определить ковариантную производную. В свою очередь ковариантная производная используется для записи ковариантных уравнений. В настоящем разделе мы определим связность на главном и ассоциированном расслоении. Покажем, что калибровочные поля Янга–Миллса представляют собой компоненты локальной формы связности. Кроме того, будет рассмотрена группа голономии, которая является одной из важнейших характеристик связности.

14.1 Связность на главном расслоении

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с базой $\mathbb{M}$, $\dim \mathbb{M} = n$, и структурной группой Ли $\mathbb{G}$, $\dim \mathbb{G} = \text{N}$. Пара $(\mathbb{P}, \mathbb{G})$ представляет собой группу преобразований. При этом группа Ли $\mathbb{G}$ действует на многообразии $\mathbb{P}$ свободно. Каждый слой $\pi^{-1}(x)$, где $x \in \mathbb{M}$, диффеоморфен структурной группе и ее действие на нем транзитивно. Из предложения 10.2.1 следует, что действие структурной группы $\mathbb{G}$ на $\mathbb{P}$ индуцирует гомоморфизм μ алгебры Ли $\mathfrak{g}$ в алгебру Ли векторных полей $\mathcal{X}(\mathbb{P})$. Этот гомоморфизм является мономорфизмом, т.к. действие группы свободно и, следовательно, эффективно.

Определение. Векторное поле $X^* = \mu(X) \in \mathcal{X}(\mathbb{P})$, где $X \in \mathfrak{g}$ – произвольный элемент алгебры Ли структурной группы Ли $\mathbb{G}$, называется *фундаментальным векторным полем, соответствующим* $X \in \mathfrak{g}$. □

Поскольку действие структурной группы на $\mathbb{P}$ отображает каждый слой в себя, то каждое фундаментальное векторное поле X^*_p касается слоя $\pi^{-1}\big(\pi(p)\big)$ в каждой точке $p \in \mathbb{P}$. Поскольку действие группы свободно, то по предложению 10.2.1 фундаментальные векторные поля нигде не обращаются в нуль. Так как размерность каждого слоя равна размерности алгебры Ли $\mathfrak{g}$, то отображение

$$\mathfrak{g} \ni X \mapsto X^*_p \in \mathbb{T}_p(\mathbb{P})$$

есть линейный мономорфизм алгебры Ли $\mathfrak{g}$ в касательное пространство в точке $p \in \mathbb{P}$ к слою, проходящему через p.

Определение. Образ алгебры Ли $\mathfrak{g}$ в касательном пространстве $\mathbb{T}_p(\mathbb{P})$ называется *вертикальным подпространством* в $\mathbb{T}_p(\mathbb{P})$ и обозначается $\mathbb{V}_p(\mathbb{P})$. □

Пусть $\{L_{\text{A}}\}$, $\text{A} = 1, \ldots, \text{N} = \dim \mathbb{G}$, – базис алгебры Ли $\mathfrak{g}$ с коммутационными соотношениями $[L_{\text{A}}, L_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} L_{\text{C}}$, где $f_{\text{AB}}{}^{\text{C}}$ – структурные константы группы Ли $\mathbb{G}$. Тогда индуцированные фундаментальные векторные поля L^*_{A} удовлетворяют тем же коммутационным соотношениям, $[L^*_{\text{A}}, L^*_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} L^*_{\text{C}}$. Множество фундаментальных векторных полей $\{L^*_{\text{A}}\}$ задает дифференцируемое инволютивное распределение вертикальных подпространств $p \mapsto \mathbb{V}_p(\mathbb{P})$ на пространстве главного расслоения $\mathbb{P}$ (см., раздел 3.11). Согласно теореме Фробениуса 3.11.1 через каждую точку $p \in \mathbb{P}$ проходит интегральное подмногообразие этого распределения. Это интегральное подмногообразие есть ни что иное, как слой $\pi^{-1}\big(\pi(p)\big)$.

Предложение 14.1.1. *Пусть X^* – фундаментальное векторное поле, соответствующее элементу алгебры Ли $X = X_0^{\text{A}} L_{\text{A}} \in \mathfrak{g}$, $X_0^{\text{A}} = \mathsf{const}$ для всех* A. *Тогда для каждого элемента группы $a \in \mathbb{G}$ векторное поле $r_{a*}X^*$, где r_{a*} – дифференциал отображения, индуцированного действием элемента a справа, является фундаментальным векторным полем, которое соответствует левоинвариантному векторному полю $\mathsf{ad}\,(a^{-1})X = X_0^{\text{B}} S^{-1}{}_{\text{B}}{}^{\text{A}}(a) L_{\text{A}} \in \mathfrak{g}$, где ad обозначает присоединенное представление $\mathbb{G}$ в $\mathfrak{g}$ и $S_{\text{A}}{}^{\text{A}}(a)$ – матрица присоединенного представления.*

Доказательство. Фундаментальное векторное поле X^* индуцируется однопараметрической группой преобразований $r_{b(t)}$, где $b(t) = \mathsf{exp}\,(tX)$. Векторное поле $r_{a*}X^*$ индуцируется однопараметрической группой преобразований $r_a \circ r_{b(t)} \circ r_{a^{-1}} = r_{a^{-1}b(t)a}$ по предложению 10.2.2. Утверждение предложения следует из того, что преобразования вида $a^{-1}b(t)a$ представляют собой однопараметрическую группу преобразований, порожденную элементом алгебры Ли $\mathsf{ad}\,(a^{-1})X \in \mathfrak{g}$. □

Пример 14.1.1 (Локальное рассмотрение). Чтобы прояснить абстрактное построение, которое было проведено выше, повторим его в компонентах. В настоящей главе мы будем возвращаться к этому примеру неоднократно. Пусть $\mathbb{U} \subset \mathbb{M}$ – окрестность базы, для которой определено отображение (13.2), с координатами x^α, $\alpha = 1, \ldots, n$. Используя диффеоморфизм $\chi : \pi^{-1}(\mathbb{U}) \to \mathbb{U} \times \widetilde{\mathbb{G}}$, мы отождествим подрасслоение $\mathbb{Q} = \pi^{-1}(\mathbb{U})$ с прямым произведением $\mathbb{U} \times \widetilde{\mathbb{G}}$. Мы отметили структурную группу в прямом произведении знаком тильды, потому что в дальнейшем нам будет необходимо различать точку главного расслоения и точку структурной группы. То есть $p = (x, a)$, где $x \in \mathbb{U}$, $a \in \widetilde{\mathbb{G}}$, для всех $p \in \mathbb{Q}$. Действие структурной группы $\mathbb{G}$ на $\mathbb{Q}$ имеет вид

$$\mathbb{G} \ni a : \quad \mathbb{Q} \ni p = (x, b) \mapsto pa = (x, ba) \in \mathbb{Q}.$$

Фактически, данный пример относится к произвольному тривиальному главному расслоению, база которого покрывается одной картой, $\mathbb{M} \approx \mathbb{R}^n$.

Любой элемент алгебры Ли $\mathfrak{g}$ (левоинвариантное векторное поле X на группе Ли $\mathbb{G}$) имеет постоянные компоненты X_0^{A} относительно левоинвариантного базиса, $X = X_0^{\text{A}} L_{\text{A}}$, где L_{A} – базис алгебры Ли $\mathfrak{g}$ (см. раздел 9.2). Базису алгебры Ли L_{A} соответствуют фундаментальные векторные поля L^*_{A} на главном расслоении, которые образуют базис вертикальных подпространств. Выберем базис касательных пространств $\mathbb{T}_p(\mathbb{Q})$ в виде $\{\partial_\alpha, L^*_{\text{A}}\}$. Тогда фундаментальное векторное поле, соответствующее произвольному элементу алгебры Ли $X \in \mathfrak{g}$, будет иметь только вертикальные компоненты, $X^* = X_0^{\text{A}} L^*_{\text{A}}$. Утверждение предложения 14.1.1 сводится к равенству

$$r_{a*}X^* = X_0^{\text{A}} S_{\text{A}}{}^{\text{B}}(a^{-1}) L^*_{\text{B}}, \tag{14.1}$$

где $S_{\text{A}}{}^{\text{B}}(a^{-1}) = S^{-1}{}_{\text{A}}{}^{\text{B}}(a)$ – матрица присоединенного представления для обратного элемента $a^{-1} \in \mathbb{G}$. Особенно просто это равенство проверяется вблизи единицы группы, где определена функция композиции. Пусть $X^*(e)$ – значение фундаментального

векторного поля в единице группы. Тогда

$$r_{a*}\Big(X^*(e)\Big) = X_0^{\text{A}} R_{\text{A}}{}^{\text{B}}(a)\partial_{\text{B}}|_a = X_0^{\text{A}} R_{\text{A}}{}^{\text{B}}(a) L^{-1}{}_{\text{B}}^{\text{C}}(a) L^*_{\text{C}}(a) = X_0^{\text{A}} S_{\text{A}}{}^{\text{B}}(a^{-1}) L^*_{\text{B}}(a),$$

где матрицы $R_{\text{A}}{}^{\text{B}}$, $L_{\text{A}}{}^{\text{B}}$ и $S_{\text{A}}{}^{\text{B}}$ были определены в разделах 9.2–9.4.

Произвольное векторное поле $\widetilde{X} \in \mathcal{X}(\mathbb{Q})$ имеет вид

$$\widetilde{X} = \widetilde{X}^\alpha \partial_\alpha + \widetilde{X}^{\text{A}} L^*_{\text{A}}.$$

После правого действия группы r_a, оно преобразуется в новое векторное поле

$$r_{a*}\widetilde{X} = \widetilde{X}^\alpha \partial_\alpha + \widetilde{X}^{\text{B}} S^{-1}{}_{\text{B}}^{\text{A}}(a) L^*_{\text{A}}, \tag{14.2}$$

т.к. дифференциал отображения r_{a*} действует только на вертикальные компоненты по правилу (14.1). □

Продолжим общее построение.

Определение. *Связностью* Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется распределение

$$\Gamma: \quad \mathbb{P} \ni p \mapsto \mathbb{H}_p(\mathbb{P}) \subset \mathbb{T}_p(\mathbb{P}),$$

которое каждой точке $p \in \mathbb{P}$ ставит в соответствие подпространство $\mathbb{H}_p(\mathbb{P})$ в касательном пространстве $\mathbb{T}_p(\mathbb{P})$ такое, что

1) в каждой точке p касательное пространство $\mathbb{T}_p(\mathbb{P})$ разлагается в прямую сумму

$$\mathbb{T}_p(\mathbb{P}) = \mathbb{V}_p(\mathbb{P}) \oplus \mathbb{H}_p(\mathbb{P});$$

2) подпространства $\mathbb{H}_p(\mathbb{P})$ инвариантны относительно правого действия структурной группы

$$r_{a*}\mathbb{H}_p(\mathbb{P}) = \mathbb{H}_{pa}(\mathbb{P}); \tag{14.3}$$

3) $\mathbb{H}_p(\mathbb{P})$ зависит дифференцируемо от p.

Множество касательных векторов $\mathbb{H}_p(\mathbb{P})$ в точке $p \in \mathbb{P}$ называется *горизонтальным подпространством* в $\mathbb{T}_p(\mathbb{P})$. Вектор $\widetilde{X} \in \mathbb{T}_p(\mathbb{P})$ называется *вертикальным* или *горизонтальным*, если он лежит соответственно в $\mathbb{V}_p(\mathbb{P})$ или $\mathbb{H}_p(\mathbb{P})$. □

Из условия 1) следует, что каждый вектор $\widetilde{X} \in \mathbb{T}_p(\mathbb{P})$ может быть единственным образом представлен в виде суммы

$$\widetilde{X} = \mathsf{v}\widetilde{X} + \mathsf{h}\widetilde{X}, \qquad \mathsf{v}\widetilde{X} \in \mathbb{V}_p(\mathbb{P}),\ \mathsf{h}\widetilde{X} \in \mathbb{H}_p(\mathbb{P}),$$

где v и h – проекторы на соответствующие подпространства. Векторы $\mathsf{v}\widetilde{X}$ и $\mathsf{h}\widetilde{X}$ называются соответственно *вертикальной* и *горизонтальной* компонентами касательного вектора $\widetilde{X}$. Ясно, что

$$\mathsf{dim}\,\mathbb{V}_p(\mathbb{P}) = \mathsf{dim}\,\mathbb{G}, \quad \mathsf{dim}\,\mathbb{H}_p(\mathbb{P}) = \mathsf{dim}\,\mathbb{M}, \qquad \forall p \in \mathbb{M}.$$

Если главное расслоение тривиально, то у него существует глобальное сечение. В этом случае условие 2) означает, что связность достаточно задать на каком либо сечении, а затем разнести по всему пространству расслоения $\mathbb{P}$ с помощью группового действия. В дальнейшем мы сформулируем теорему 14.1.1, определяющее связность на главном расслоении общего вида через семейство локальных форм связности, которые заданы на координатном покрытии базы.

Условие 3) обозначает по определению, что, если $\widetilde{X}$ – дифференцируемое векторное поле на $\mathbb{P}$, то таковы же вертикальная и горизонтальная компоненты $\mathsf{v}\widetilde{X}$ и $\mathsf{h}\widetilde{X}$.

Пример 14.1.2. Накрытие $\widetilde{\mathbb{M}} \to \mathbb{M}$ (см. раздел 12) является главным расслоением с 0-мерной структурной группой. В этом случае вертикальные подпространства отсутствуют, а горизонтальное подпространство в точке $p \in \widetilde{\mathbb{M}}$ совпадает с касательным пространством $\mathbb{T}_p(\widetilde{\mathbb{M}})$. Это означает, что связность для накрытий единственна и распределение горизонтальных подпространств совпадает с касательны расслоением $\mathbb{T}(\widetilde{\mathbb{M}})$. □

Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана связность Γ. Для того, чтобы конструктивно описать распределение горизонтальных подпространств построим на главном расслоении $\mathbb{P}$ 1-форму ω со значениями в алгебре Ли $\mathfrak{g}$, т.е. $\omega = \omega^{\text{A}} L_{\text{A}}$, $\omega^{\text{A}} \in \Lambda_1(\mathbb{P})$ для всех A. Как было отмечено в начале настоящего раздела, каждому элементу алгебры Ли $X \in \mathfrak{g}$ соответствует единственное фундаментальное векторное поле X^*. При этом отображение $X \mapsto X^*_p$ представляет собой линейный мономорфизм $\mathfrak{g}$ в $\mathbb{T}_p(\mathbb{P})$ для всех точек главного расслоения $p \in \mathbb{P}$.

Определение. Для каждого касательного вектора $\widetilde{X}_p \in \mathbb{T}_p(\mathbb{P})$ определим значение 1-формы $\omega(\widetilde{X}_p)$ на векторе $\widetilde{X}_p$, как единственный элемент алгебры Ли $X \in \mathfrak{g}$ такой, что фундаментальное векторное поле X^*_p в точке $p \in \mathbb{P}$ совпадает с вертикальной компонентой вектора, $X^*_p = \mathsf{v}\widetilde{X}_p$. 1-форма ω на $\mathbb{P}$ со значениями в алгебре Ли $\mathfrak{g}$ называется *формой связности* для заданной связности Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. □

По построению, $\omega(\widetilde{X}) = 0$ тогда и только тогда, когда векторное поле $\widetilde{X}$ горизонтально. В силу следующего предложения 1-форма ω взаимно однозначно определяет распределение горизонтальных подпространств и, следовательно, связность Γ на $\mathbb{P}$.

Предложение 14.1.2. *Форма связности ω на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ имеет следующие свойства:*

1) $\omega(X^*) = X$ для всех $X \in \mathfrak{g}$;

2) $r_a^*\omega = \mathsf{ad}\,(a^{-1})\omega$, т.е.

$$(r_a^*\omega)(\widetilde{X}) = \mathsf{ad}\,(a^{-1})\omega(\widetilde{X}) = \omega(\widetilde{X})^{\text{B}} S^{-1}{}_{\text{B}}{}^{\text{A}}(a) L_{\text{A}}$$

для всех $a \in \mathbb{G}$ и каждого векторного поля $\widetilde{X}$ на $\mathbb{P}$.

Обратно. Если на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана 1-форма ω со значениями в алгебре Ли $\mathfrak{g}$, удовлетворяющая условиям 1) и 2), то на $\mathbb{P}$ существует единственная связность Γ такая, что ее форма связности есть ω.

Доказательство. Пусть ω – форма связности на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Условие 1) следует немедленно из определения ω. Поскольку каждое векторное поле $\widetilde{X} \in \mathcal{X}(\mathbb{P})$ разлагается на вертикальную и горизонтальную составляющую, то достаточно проверить условие 2) для этих двух компонент. Если $\widetilde{X}$ горизонтально, то $r_{a*}\widetilde{X}$ также горизонтально для всех $a \in \mathbb{G}$, что следует из условия 2) в определении связности. Поэтому $(r_a^*\omega)_p(\widetilde{X}) = \omega_{pa}(r_{a*}\widetilde{X})$ и $\mathsf{ad}\,(a^{-1})\omega_p(\widetilde{X})$ одновременно обращаются в нуль. Если $\widetilde{X}$ вертикально, то его значение в точке $p \in \mathbb{P}$ определяется некоторым фундаментальным векторным полем, $\widetilde{X}_p = X^*_p$, для некоторого $X \in \mathfrak{g}$. Тогда из предложения 14.1.1 следует, что векторное поле $r_{a*}X^*$ соответствует элементу алгебры Ли $\mathsf{ad}\,(a^{-1})X \in \mathfrak{g}$. Поэтому

$$(r_a^*\omega)_p(X^*) = \omega_{pa}(r_{a*}X^*) = \mathsf{ad}\,(a^{-1})X = \mathsf{ad}\,(a^{-1})\Big(\omega_p(X^*)\Big).$$

Обратно. Пусть задана 1-форма ω, удовлетворяющая условиям 1) и 2). Определим распределение горизонтальных подпространств

$$\mathbb{H}_p(\mathbb{P}) := \{\widetilde{X} \in \mathbb{T}_p(\mathbb{P}) : \quad \omega(\widetilde{X}) = 0\} \qquad \forall p \in \mathbb{P}.$$

Теперь нетрудно проверить, что распределение $\Gamma:\ p\mapsto \mathbb{H}_p(\mathbb{P})$ является связностью, для которой ω – форма связности. $\square$

Пусть на главной расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ задано две связности Γ_1 и Γ_2 с формами связности ω_1 и ω_2 соответственно. Нетрудно проверить, что 1-форма $(1-t)\omega_1+t\omega_2$, $t\in[0,1]$, также задает некоторую связность на $\mathbb{P}$. Таким образом, любые две связности, заданные на одном главном расслоении, гомотопны.

Пример 14.1.3 (Локальное рассмотрение). Продолжим локальное построение, начатое в примере 14.1.1. У нас есть левоинвариантный базис $\{\partial_\alpha, L^*_{\text{A}}\}$ касательных пространств $\mathbb{T}_p(\mathbb{Q})$ к главному подрасслоению $\mathbb{Q}(\mathbb{U},\pi,\mathbb{G})\subset\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$. Обозначим дуальный к нему базис 1-форм на $\mathbb{Q}$ через $\{dx^\alpha,\omega^{*\text{A}}\}$. Теперь построим форму связности ω на $\mathbb{Q}$. Произвольную 1-форму на $\mathbb{Q}$ со значениями в алгебре Ли $\mathfrak{g}$ можно разложить по этому базису, $\omega=\Big(dx^\alpha\omega_\alpha{}^{\text{A}}+\omega^{*\text{B}}\omega_{\text{B}}{}^{\text{A}}\Big)L_{\text{A}}$, где $\omega_\alpha{}^{\text{A}}(x,a)$ и $\omega_{\text{B}}{}^{\text{A}}(x,a)$ – некоторые компоненты, зависящие от точки $(x,a)\in\mathbb{Q}$. Свойство 1) предложения 14.1.2 взаимно однозначно определяет компоненту $\omega_{\text{B}}{}^{\text{A}}$,

$$\omega(X^*)=X\quad\Leftrightarrow\quad\omega_{\text{B}}{}^{\text{A}}(x,a)=\delta^{\text{A}}_{\text{B}}.$$

Нетрудно проверить, что 1-форма $\omega^{*\text{A}}L_{\text{A}}$ удовлетворяет свойству 2) предложения 14.1.2. Действительно,

$$r_a^*(\omega^{*\text{A}}L_{\text{A}})(\widetilde{X})=(\omega^{*\text{A}}L_{\text{A}})(r_{a*}\widetilde{X})=\widetilde{X}^{\text{B}}S^{-1}{}_{\text{B}}^{\text{A}}(a)L_{\text{A}},$$

где мы учли действие дифференциала отображения на векторное поле (14.2). Отсюда следует, что форма $\omega^{*\text{A}}L_{\text{A}}$ действительно удовлетворяет условию 2),

$$r_a^*(\omega^{*\text{A}}L_{\text{A}})=\omega^{*\text{B}}S^{-1}{}_{\text{B}}^{\text{A}}(a)L_{\text{A}}.$$

Проведя аналогичные вычисления, получаем, что 1-форма со значениями в алгебре Ли $dx^\alpha\omega_\alpha{}^{\text{A}}L_{\text{A}}$ удовлетворяет свойству 2) предложения 14.1.2 тогда и только тогда, когда ее компоненты имеют вид

$$\omega_\alpha{}^{\text{A}}(x,b)=\overset{\circ}{A}_\alpha{}^{\text{B}}(x)S^{-1}{}_{\text{B}}^{\text{A}}(b),\tag{14.4}$$

где $\overset{\circ}{A}_\alpha{}^{\text{B}}(x)$ – произвольные функции от $x\in\mathbb{U}$. Таким образом, форма связности на $\mathbb{Q}(\mathbb{U},\pi,\mathbb{G})=\mathbb{U}\times\widetilde{\mathbb{G}}$ имеет вид

$$\omega(x,a)=\Big[dx^\alpha\omega_\alpha{}^{\text{A}}(x,a)+\omega^{*\text{A}}(a)\Big]L_{\text{A}},\tag{14.5}$$

где компоненты $\omega_\alpha{}^{\text{A}}$ определены равенством (14.4). Мы видим, что форма связности ω на главном расслоении $\mathbb{Q}=\mathbb{U}\times\widetilde{\mathbb{G}}$ параметризуется $n\times\text{N}$ произвольными функциями $\overset{\circ}{A}_\alpha{}^{\text{B}}(x)$ на координатной окрестности $\mathbb{U}\subset\mathbb{M}$. Функции $\overset{\circ}{A}_\alpha{}^{\text{B}}(x)$ определяют компоненты формы связности на *нулевом локальном сечении*

$$\sigma_0:\quad\mathbb{U}\ni x\mapsto(x,e)\in\mathbb{U}\times\widetilde{\mathbb{G}}.\tag{14.6}$$

Вектор $\widetilde{X}_p=\widetilde{X}^\alpha_p\partial_\alpha+\widetilde{X}^{\text{A}}_pL^*_{\text{A}}\in\mathbb{T}_x(\mathbb{Q})$ горизонтален тогда и только тогда, когда выполнено равенство

$$\omega(\widetilde{X}_p)=\Big(\widetilde{X}^\alpha_p\omega_\alpha{}^{\text{A}}+\widetilde{X}^{\text{A}}_p\Big)L_{\text{A}}=0,$$

где $\omega_\alpha{}^{\text{A}}$ имеет вид (14.4). Поэтому его вертикальная и горизонтальная составляющие равны

$$\begin{aligned} \mathsf{v}\widetilde{X}_p &= \left(\widetilde{X}^{\text{A}}_p + \widetilde{X}^\alpha_p \omega_\alpha{}^{\text{A}}\right) L^*_{\text{A}}, \\ \mathsf{h}\widetilde{X}_p &= \widetilde{X}^\alpha_p \partial_\alpha - \widetilde{X}^\alpha_p \omega_\alpha{}^{\text{A}} L^*_{\text{A}}. \end{aligned} \tag{14.7}$$

Перемешивание компонент $\widetilde{X}^\alpha_p$ и $\widetilde{X}^{\text{A}}_p$ связано с тем, что векторные поля ∂_α в общем случае не являются горизонтальными. Они горизонтальны тогда и только тогда, когда $\omega_\alpha{}^{\text{A}} = 0 \Leftrightarrow \overset{\circ}{A}_\alpha{}^{\text{A}} = 0$. □

Замечание. Из вида формы связности (14.5) следует, что она совпадает с канонической 1-формой θ на группе Ли (9.25), если база расслоения состоит из одной точки. В этом смысле форма связности ω на главном расслоении $\mathbb{P}$ является обобщением канонической 1-формы θ на группе Ли $\mathbb{G}$. □

Проекции $\pi : \mathbb{P} \to \mathbb{M}$ соответствует дифференциал отображения $\pi_* : \mathbb{T}_p(\mathbb{P}) \to \mathbb{T}_x(\mathbb{M})$, который линейно отображает касательное пространство к главному расслоению в каждой точке $p \in \mathbb{P}$ в касательное пространство к базе в точке $x = \pi(p) \in \mathbb{M}$. Ядром дифференциала проекции π_* является вертикальное подпространство $\mathbb{V}_p(\mathbb{P})$. Действительно, любой вертикальный вектор касается некоторой кривой, целиком лежащей в слое. При проекции вся кривая отображается в одну точку x. Поэтому каждый вертикальный вектор отображается в нулевой вектор из $\mathbb{T}_x(\mathbb{M})$. Поскольку дифференциал проекции – это сюръективный гомоморфизм, то отсюда следует, что отображение горизонтального подпространства $\pi_* : \mathbb{H}_p(\mathbb{P}) \to \mathbb{T}_x(\mathbb{M})$ является изоморфизмом.

Определение. *Горизонтальным лифтом* или *подъемом* (или просто *лифтом*) векторного поля X на базе $\mathbb{M}$ называется единственное векторное поле $\widetilde{X}$ на $\mathbb{P}$, которое горизонтально, $\widetilde{X}_p \in \mathbb{H}_p(\mathbb{P})$, и проектируется на X, т.е. $\pi_*(\widetilde{X}_p) = X_{\pi(p)}$ для всех $p \in \mathbb{P}$. □

Предложение 14.1.3. *Если на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана связность Γ и на базе $\mathbb{M}$ задано векторное поле X, то существует единственный горизонтальный лифт $\widetilde{X}$ векторного поля X. Лифт $\widetilde{X}$ инвариантен относительно действия структурной группы, $r_{a*}\widetilde{X}_p = \widetilde{X}_{pa}$, для всех $a \in \mathbb{G}$. Обратно. Каждое горизонтальное векторное поле $\widetilde{X}$ на $\mathbb{P}$, инвариантное относительно действия структурной группы, является горизонтальным лифтом некоторого векторного поля X на базе $\mathbb{M}$.*

Доказательство. См., например, [**?**]. □

Предложение 14.1.4. *Пусть $\widetilde{X}$ и $\widetilde{Y}$ – горизонтальные лифты соответственно векторных полей X и Y, заданных на базе $\mathbb{M}$. Тогда:*

1) *$\widetilde{X} + \widetilde{Y}$ – горизонтальный лифт векторного поля $X + Y$;*
2) *для каждой функции $f \in \mathcal{C}^\infty(\mathbb{M})$ произведение $\tilde{f}\widetilde{X}$ есть горизонтальный лифт для $fX \in \mathcal{X}(\mathbb{M})$, где функция $\tilde{f} \in \mathcal{C}^\infty(\mathbb{P})$ определена равенством $\tilde{f} = f \circ \pi$;*
3) *горизонтальная компонента коммутатора $[\widetilde{X}, \widetilde{Y}]$ есть горизонтальный лифт коммутатора $[X, Y]$.*

Доказательство. Первые два утверждения очевидны. Доказательство третьего просто:

$$\pi_*(\mathsf{h}[\widetilde{X},\widetilde{Y}]) = \pi_*([\widetilde{X},\widetilde{Y}]) = [X,Y]. \qquad \square$$

Пусть x^α, $\alpha = 1,\dots,\dim\mathbb{M}$ – система координат на некоторой окрестности базы $\mathbb{U}\subset\mathbb{M}$. Пусть D_α – горизонтальный лифт векторного поля ∂_α в $\pi^{-1}(\mathbb{U})$ для всех α. Тогда векторные поля $\{D_\alpha\}$ образуют локальный базис распределения $\Gamma:\ p\mapsto\mathbb{H}_p(\mathbb{P})$ в окрестности $\pi^{-1}(\mathbb{U})\subset\mathbb{P}$.

Пример 14.1.4 (Локальное рассмотрение). Продолжим локальное построение, начатое в примере 14.1.1. Пусть $X(x) = X^\alpha_x\partial_\alpha$ – произвольное векторное поле на базе $\mathbb{M}$. Его горизонтальный лифт в $\mathbb{Q}$ имеет вид

$$\widetilde{X}_p = X^\alpha_x\partial_\alpha - X^\alpha_x\omega_\alpha{}^{\mathrm{A}}L^*_{\mathrm{A}},\qquad x=\pi(p),$$

для всех $p\in\mathbb{Q}$. Вблизи единицы группы второе слагаемое имеет вид

$$X^\alpha_x\mathring{A}_\alpha{}^{\mathrm{A}}S^{-1}{}_{\mathrm{A}}{}^{\mathrm{B}}L^*_{\mathrm{B}} = X^\alpha_x\mathring{A}_\alpha{}^{\mathrm{A}}R^*_{\mathrm{A}},$$

где R^*_{A} правоинвариантные векторные поля на группе Ли $\widetilde{\mathbb{G}}$ и мы воспользовались выражением матрицы присоединенного представления через производные от функции композиции (9.35). Отсюда сразу следует, что горизонтальное векторное поле $\widetilde{X}$ инвариантно относительно действия структурной группы справа. Утверждения предложения 14.1.2 становятся тривиальными.

В частности, горизонтальный лифт координатного базиса ∂_α на $\mathbb{M}$ имеет вид

$$D_\alpha = \partial_\alpha - \omega_\alpha{}^{\mathrm{A}}(p)L^*_{\mathrm{A}} = \partial_\alpha - \mathring{A}_\alpha{}^{\mathrm{B}}(x)S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(a)L^*_{\mathrm{A}}$$

для всех точек $p=(x,a)\in\mathbb{P}$. $\square$

Теперь спустимся на базу и определим форму связности на $\mathbb{P}$ через семейство 1-форм на $\mathbb{M}$. Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение с формой связности ω. Пусть $\mathbb{U}\subset\mathbb{M}$ – некоторая координатная окрестность базы, на которой задано локальное сечение $\sigma:\ \mathbb{U}\to\mathbb{P}$. Тогда форма связности ω определяет на $\mathbb{U}$ 1-форму с помощью возврата отображения σ^*.

Определение. 1-форма $\omega_\sigma = \sigma^*\omega$ со значениями в алгебре Ли $\mathfrak{g}$ структурной группы, $\omega_\sigma = \omega^{\mathrm{A}}_\sigma L_{\mathrm{A}}$, $\omega^{\mathrm{A}}_\sigma\in\Lambda_1(\mathbb{U})$ называется *локальной формой связности* на $\mathbb{U}\subset\mathbb{M}$. В компонентах она имеет вид

$$\omega_\sigma = dx^\alpha A_\alpha{}^{\mathrm{A}}L_{\mathrm{A}}, \tag{14.8}$$

где $A_\alpha{}^{\mathrm{A}}(x)$ – некоторые функции на $\mathbb{U}$. $\square$

Напомним, что, по определению, $\omega_\sigma(X) = \omega(\sigma_*X)$ для всех векторных полей на базе $X\in\mathcal{X}(\mathbb{M})$, где σ^* и σ_* – возврат и дифференциал локального сечения.

Каждая форма связности ω на $\mathbb{P}$ однозначно определяет локальную форму связности ω_σ на $\mathbb{U}\subset\mathbb{M}$. Обратное, конечно, неверно, т.к. локальная форма связности ω_σ определена только на окрестности $\mathbb{U}$ и, вдобавок, зависит от сечения σ.

Пример 14.1.5 (Локальное рассмотрение). Рассмотрим зависимость локальной формы связности от сечения. Пусть на главном расслоении $\mathbb{Q} = \mathbb{U}\times\widetilde{\mathbb{G}}$ задано два произвольных сечения σ и σ'. Поскольку две точки одного слоя связаны некоторым

групповым преобразованием, то σ' можно выразить через первое сечение, $\sigma'(x) = \sigma(x)a(x)$, где $a(x):\ \mathbb{U} \to \mathbb{G}$ – некоторая функция. Таким образом

$$\sigma:\ x \mapsto \Big(x, b(x)\Big) \quad \text{и} \quad \sigma':\ x \mapsto \Big(x, b(x)a(x)\Big). \tag{14.9}$$

Рассмотрим случай, когда все элементы структурной группы Ли $\mathbb{G}$ находятся в окрестности единицы группы, соответствующей локальной группе Ли (см. раздел 9.1). То есть каждый элемент группы имеет координаты, $a = \{a^{\mathrm{A}}\} \in \mathbb{G}$, $\mathrm{A} = 1, \ldots, \mathrm{N}$, и задана функция композиции $ba = f = \{f^{\mathrm{A}}(b,a)\}$. Тогда действие дифференциалов сечений на векторное поле $X = X^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M})$ имеет вид

$$\begin{aligned}
\sigma_* X &= X^\alpha \partial_\alpha + X^\alpha \partial_\alpha b^{\mathrm{A}} \partial_{\mathrm{A}} = X^\alpha \partial_\alpha + X^\alpha \partial_\alpha b^{\mathrm{A}} L^{-1}{}_{\mathrm{A}}{}^{\mathrm{B}}(b) L^*_{\mathrm{B}}(b),\\
\sigma'_* X &= X^\alpha \partial_\alpha + X^\alpha \partial_\alpha f^{\mathrm{A}} \partial_{\mathrm{A}} = X^\alpha \partial_\alpha + X^\alpha \partial_\alpha f^{\mathrm{A}} L^{-1}{}_{\mathrm{A}}{}^{\mathrm{B}}(f) L^*_{\mathrm{B}}(f),
\end{aligned}$$

где использован явный вид формы связности (14.5). Локальные формы связности ω_σ и $\omega_{\sigma'}$ являются 1-формами на $\mathbb{U}$ со значениями в алгебре Ли $\mathfrak{g}$. Поэтому они представимы в виде

$$\begin{aligned}
\omega_\sigma &= dx^\alpha A_\alpha{}^{\mathrm{A}}(x) L_{\mathrm{A}},\\
\omega_{\sigma'} &= dx^\alpha A'_\alpha{}^{\mathrm{A}}(x) L_{\mathrm{A}},
\end{aligned} \tag{14.10}$$

где $A_\alpha{}^{\mathrm{A}}(x)$ и $A'_\alpha{}^{\mathrm{A}}(x)$ – компоненты этих форм (некоторые дифференцируемые функции от $x \in \mathbb{U}$). С другой стороны, по определению,

$$\begin{aligned}
\omega_\sigma(X) = \omega(\sigma_* X) &= X^\alpha \left[\overset{\circ}{A}_\alpha{}^{\mathrm{B}} S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(b) + \partial_\alpha b^{\mathrm{B}} L^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(b) \right] L_{\mathrm{A}},\\
\omega_{\sigma'}(X) = \omega(\sigma'_* X) &= X^\alpha \left[\overset{\circ}{A}_\alpha{}^{\mathrm{B}} S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(f) + \partial_\alpha f^{\mathrm{B}} L^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(f) \right] L_{\mathrm{A}}.
\end{aligned} \tag{14.11}$$

Рассмотрим нулевое сечение $\sigma_0 = (x, e)$, где e – единица группы. Тогда

$$\omega_0 = dx^\alpha \overset{\circ}{A}_\alpha{}^{\mathrm{A}} L_{\mathrm{A}}.$$

То есть функции $\overset{\circ}{A}_\alpha{}^{\mathrm{A}}(x)$, параметризующие форму связности в (14.5), представляют собой компоненты локальной формы связности для нулевого сечения. Тогда компоненты локальной формы связности ω_σ связаны с компонентами локальной формы связности для нулевого сечения простым соотношением

$$A_\alpha{}^{\mathrm{A}} = \overset{\circ}{A}_\alpha{}^{\mathrm{B}} S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(b) + \partial_\alpha b^{\mathrm{B}} L^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(b). \tag{14.12}$$

Сравнивая формулы (14.11), получим связь между компонентами локальных форм связности,

$$A'_\alpha{}^{\mathrm{A}} = A_\alpha{}^{\mathrm{B}} S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(a) + \partial_\alpha f^{\mathrm{B}} L^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}(f) - \partial_\alpha b^{\mathrm{B}} L^{-1}{}_{\mathrm{B}}{}^{\mathrm{C}}(b) S^{-1}{}_{\mathrm{C}}{}^{\mathrm{A}}(a), \tag{14.13}$$

где функция $a(x)$ связывает данные сечения $\sigma' = \sigma a = (x, f = ba)$.

Полученная связь компонент двух локальных форм связности (14.13), заданных на одной окрестности базы $\mathbb{U} \subset \mathbb{M}$ неудобна в приложениях, т.к. содержит координатные функции $a^{\mathrm{A}}(x)$, $b^{\mathrm{A}}(x)$, f^{A}, которые в явном виде можно задать в исключительно редких случаях. Чтобы обойти эту трудность выберем присоединенное представление алгебры Ли $\mathsf{ad}\,(\mathfrak{g})$. То есть вместо локальных 1-форм ω_σ со значениями в алгебре Ли $\mathfrak{g}$ будем рассматривать 1-формы связности со значениями в присоединенном представлении алгебры Ли $\mathsf{ad}\,(\mathfrak{g})$. Этому соответствует переход к матрицам

$$A_\alpha{}^{\mathrm{A}} \mapsto A_{\alpha\mathrm{B}}{}^{\mathrm{C}} := -A_\alpha{}^{\mathrm{A}} f_{\mathrm{AB}}{}^{\mathrm{C}},$$

где $f_{\text{AB}}{}^{\text{C}}$ – структурные константы группы Ли $\mathbb{G}$. Компоненты локальной формы связности в присоединенном представлении мы будем обозначать $A_\alpha = \{A_{\alpha\text{B}}{}^{\text{C}}\}$, опуская для краткости матричные индексы. Продифференцируем матрицу присоединенного представления $S_{\text{A}}{}^{\text{B}}(b)$,

$$\partial_\alpha S_{\text{A}}{}^{\text{B}} = \partial_\alpha b^{\text{C}} \partial_{\text{C}} S_{\text{A}}{}^{\text{B}} = \partial_\alpha b^{\text{C}} L^{-1}{}_{\text{C}}{}^{\text{D}} f_{\text{DA}}{}^{\text{E}} S_{\text{E}}{}^{\text{B}},$$

где мы воспользовались формулой дифференцирования (9.39). Умножив это равенство справа на S^{-1}, получим

$$\partial_\alpha S_{\text{A}}{}^{\text{B}} S^{-1}{}_{\text{B}}{}^{\text{C}} = \partial_\alpha b^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{D}} f_{\text{DA}}{}^{\text{C}}.$$

Тогда для компонент локальной формы связности (14.12) со значениями в присоединенном представлении алгебры Ли справедлива формула

$$A_\alpha = S\mathring{A}_\alpha S^{-1} + \partial_\alpha S S^{-1}, \tag{14.14}$$

где мы опустили матричные индексы. Отсюда следует правило преобразования компонент локальной формы связности при переходе от одного сечения $\sigma:\ \big(x, b(x)\big)$ к другому $\sigma':\ \big(x, b(x)a(x)\big)$,

$$A'_\alpha = S A_\alpha S^{-1} + \partial_\alpha S S^{-1}, \tag{14.15}$$

где $S = \{S_{\text{A}}{}^{\text{B}}(a)\}$ – матрица присоединенного представления. Это есть ни что иное, как хорошо известная формула калибровочного преобразования полей Янга–Миллса. Эта формула имеет явные преимущества по сравнению с (14.13), т.к. позволяет проводить вычисления с матрицами присоединенного представления, элементы которых зависят от точки базы $x \in \mathbb{M}$.

Вместо присоединенного можно рассматривать произвольное представление алгебры Ли $\mathfrak{g}$. Если $L_{\text{A}i}{}^j$ – матрицы, соответствующие левоинвариантным векторным полям L_{A}, и $A_{\alpha i}{}^j = A_\alpha{}^{\text{A}} L_{\text{A}i}{}^j$ – компоненты локальной формы связности в данном представлении, то формулы преобразования компонент при переходе между сечениями (14.15) сохраняются, если под S понимать матрицу соответствующего преобразования, $S = \{S_i{}^j(a)\}$. □

Функции $a(x)$ можно рассматривать как функции, задающие вертикальный автоморфизм главного расслоения $\mathbb{Q} = \mathbb{U} \times \tilde{\mathbb{G}}$ (см. пример 13.3.1). Тогда формула (14.15) задает преобразование компонент локальной формы связности при вертикальном автоморфизме.

Отображение $\mathbb{U} \ni x \mapsto S_{\text{A}}{}^{\text{B}}(x) \in \mathsf{ad}\,(\mathfrak{g})$ сопоставляет каждой точке базы матрицу присоединенного представления структурной группы. В формуле (14.15) каждое слагаемое определяет 1-форму на $\mathbb{U}$ и не зависит явно от функции композиции. Поэтому, несмотря на то, что формулы преобразования компонент локальной формы связности были получены в окрестности единицы группы, они справедливы для произвольных локальных сечений на $\mathbb{U} \subset \mathbb{M}$.

Чтобы установить связь рассмотренной конструкции с понятиями, которые широко используются в математической физике, дадим

Определение. *Калибровочным полем* или *полем Янга–Миллса* $A_\alpha{}^{\text{A}}(x)$ на координатной окрестности $\mathbb{U} \subset \mathbb{M}$ называются компоненты локальной формы связности (14.8),

после добавления соответствующих уравнений движения. *Калибровочным преобразованием* называется переход между двумя локальными сечениями σ и σ' на $\mathbb{U}$ или, что эквивалентно, преобразование компонент локальной формы связности при вертикальном автоморфизме. Структурная группа $\mathbb{G}$ называется *калибровочной* группой. В электродинамике структурной группой является абелева группа $\mathbb{U}(1)$, а компоненты локальной формы связности $A_\alpha(x)$ на $\mathbb{U}$, после добавления уравнений Максвелла, называются *электромагнитным потенциалом.* □

Замечание. Форма связности ω на $\mathbb{P}$ определена инвариантным образом и ее компоненты $\omega_\alpha{}^{\text{A}}(p)$ в (14.5) являются тензорными полями на пространстве главного расслоения $\mathbb{P}$. Важное неоднородное слагаемое в калибровочном преобразовании (14.15) для компонент локальной формы связности $A_\alpha{}^{\text{A}}(x)$ возникает только при переходе к локальным сечениям. □

Как уже было отмечено, локальная форма связности ω_σ не определяет форму связности ω и, следовательно, связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Однако, если задано координатное покрытие базы $\mathbb{M}$ и семейство локальных форм на каждой координатной окрестности, то это семейство однозначно определяет связность на $\mathbb{P}$. Опишем соответствующую конструкцию. Пусть задано координатное покрытие базы, $\mathbb{M} = \bigcup_i \mathbb{U}_i$, семейство изоморфизмов

$$\chi_i : \quad \pi^{-1}(\mathbb{U}_i) \to \mathbb{U}_i \times \mathbb{G}$$

и соответствующие функции перехода

$$a_{ji} : \quad \mathbb{U}_i \cap \mathbb{U}_j \to \mathbb{G}.$$

Это необходимо для однозначного определения главного расслоения в терминах функций перехода (теорема 13.1.2). На каждой окрестности $\mathbb{U}_i$ выберем сечение $\sigma_i : \mathbb{U}_i \to \mathbb{P}$, которое соответствует единичному элементу группы $e \in \mathbb{G}$, положив $\sigma_i(x) = \chi_i^{-1}(x, e)$. Пусть θ – (левоинвариантная $\mathfrak{g}$-значная) каноническая 1-форма на $\mathbb{G}$, которая определена в разделе 9.2.

Для каждого непустого пересечения $\mathbb{U}_i \cap \mathbb{U}_j$ определим $\mathfrak{g}$-значную 1-форму на $\mathbb{U}_i \cap \mathbb{U}_j$ с помощью возврата отображения a_{ji},

$$\theta_{ji} = a_{ji}^* \theta. \tag{14.16}$$

Для каждой окрестности $\mathbb{U}_i$ определим локальную форму связности на $\mathbb{U}_i$ с помощью возврата сечения σ_i,

$$\omega_i = \sigma_i^* \omega,$$

где ω – форма связности на $\mathbb{P}$. Тогда справедлива

Теорема 14.1.1. *Формы θ_{ji} и ω_i удовлетворяют условиям:*

$$\omega_j = \mathsf{ad}\,(a_{ji}^{-1})\omega_i + \theta_{ji}, \qquad \textit{на} \quad \mathbb{U}_i \cap \mathbb{U}_j. \tag{14.17}$$

Обратно. Для каждого семейства локальных форм связности $\{\omega_i\}$, заданных на координатном покрытии $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и удовлетворяющих условиям (14.17) во всех пересечениях карт, существует единственная форма связности ω на $\mathbb{P}$, которая порождает семейство 1-форм $\{\omega_i\}$ вышеописанным образом.

Доказательство. См., например, [**?**]. □

Пример 14.1.6 (Локальное рассмотрение). Продолжим локальное рассмотрение, чтобы прояснить содержание последнего утверждения. Допустим, что мы находимся в окрестности единицы структурной группы. Тогда каноническая форма на группе Ли имеет вид $\theta = da^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{A}} L_{\text{A}}$, где L_{A} – левоинвариантный базис алгебры Ли $\mathfrak{g}$. Пусть в пересекающихся картах $\mathbb{U}_i$ и $\mathbb{U}_j$ заданы координаты x^α и $x^{\alpha'}$ соответственно. Пусть

$$X = X^\alpha \partial_\alpha = X^{\alpha'} \partial_{\alpha'} \in \mathcal{X}(\mathbb{U}_i \cap \mathbb{U}_j)$$

– касательный вектор к базе на пересечении окрестностей. Тогда дифференциал функций перехода отображает этот вектор в касательное пространство группы

$$a_{ji*} X = X^\alpha \partial_\alpha a_{ij}^{\text{A}} \partial_{\text{A}} = X^\alpha \partial_\alpha a_{ji}^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{A}} L_{\text{A}}.$$

Такая же формула имеет место в штрихованной системе координат. Поэтому $\mathfrak{g}$-значная 1-форма (14.16) имеет вид

$$\theta_{ji} = dx^\alpha \partial_\alpha a_{ji}^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{A}}(a_{ji}) L_{\text{A}} = dx^{\alpha'} \partial_{\alpha'} a_{ji}^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{A}}(a_{ji}) L_{\text{A}}.$$

Пусть

$$\omega_i = dx^\alpha A_\alpha{}^{\text{A}} L_{\text{A}} \quad \text{и} \quad \omega_j = dx^{\alpha'} A_{\alpha'}{}^{\text{A}} L_{\text{A}}$$

– локальные формы связности соответственно на $\mathbb{U}_i$ и $\mathbb{U}_j$. Тогда формула (14.17) дает

$$A_{\alpha'}{}^{\text{A}} = \partial_{\alpha'} x^\alpha \left[A_\alpha{}^{\text{B}} S^{-1}{}_{\text{B}}{}^{\text{A}}(a_{ji}) + \partial_\alpha a_{ji}^{\text{B}} L^{-1}{}_{\text{B}}{}^{\text{A}}(a_{ji}) \right]. \tag{14.18}$$

Это преобразование калибровочного поля отличается от описанного ранее (14.12) только множителем $\partial_{\alpha'} x^\alpha$, соответствующим преобразованию координат в пересечении $\mathbb{U}_i \cap \mathbb{U}_j$. □

Таким образом, для того, чтобы описать связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ в локальных терминах, необходимо задать координатное покрытие базы $\mathbb{M} = \bigcup_i \mathbb{U}_i$, функции перехода $a_{ji}(x)$ в каждом непустом пересечении $\mathbb{U}_i \cap \mathbb{U}_j$, которые удовлетворяют условию (13.9) в областях пересечения трех карт, и семейство калибровочных полей, заданных на каждой координатной окрестности, такое, что в областях пересечения выполнено условие (14.18).

Замечание. В квантовой теории поля обычно рассматривают тривиальные главные расслоения $\mathbb{P} = \mathbb{R}^{1,3} \times \mathbb{G}$, где $\mathbb{R}^{1,3}$ – четырехмерное пространство Минковского и $\mathbb{G}$ – калибровочная группа. В этом частном случае локальная форма связности $dx^\alpha A_\alpha{}^{\text{A}} L_{\text{A}}$, заданная на всем пространстве-времени $\mathbb{R}^{1,3}$, взаимно однозначно определяет форму связности ω и, следовательно, связность Γ на $\mathbb{P}$. □

В заключение данного раздела обсудим вопрос о существовании связностей. Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $\mathbb{N} \subset \mathbb{M}$ – некоторое подмножество базы. Мы говорим, что связность определена над $\mathbb{N}$, если в каждой точке $p \in \pi^{-1}(\mathbb{N}) \subset \mathbb{P}$ из прообраза $\mathbb{N}$ определено горизонтальное подпространство $\mathbb{H}_p(\mathbb{P})$ таким образом, что выполнены первые два условия (прямая сумма и инвариантность относительно действия структурной группы) в определении связности и распределение $\mathbb{H}_p(\mathbb{P})$ дифференцируемо зависит от точки p в следующем смысле. Для каждой точки $x \in \mathbb{N}$ существует окрестность $\mathbb{U} \ni x$ и связность на $\pi^{-1}(\mathbb{U})$ такая, что ее сужение на $\pi^{-1}(x)$ совпадает с $\mathbb{H}_p(\mathbb{P})$ для всех точек слоя $p \in \pi^{-1}(x)$.

Теорема 14.1.2. *Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение и $\mathbb{N}$ – замкнутое подмножество (возможно, пустое) в базе $\mathbb{M}$. Любая связность, определенная над $\mathbb{N}$ может быть продолжена до связности на всем $\mathbb{P}$. В частности, любое главное расслоение $\mathbb{P}$ допускает связность.*

Доказательство. Доказательство проводится путем явного построения связности. При этом существенно используется паракомпактность многообразия. См., например, [**?**]. □

Пример 14.1.7. Рассмотрим главное расслоение $\mathbb{P}(\mathbb{M},\pi,e)$, структурная группа которого состоит из единственного элемента – единицы. Тогда на нем существует единственная связность. В этом случае вертикальные касательные пространства отсутствуют, а распределение горизонтальных пространств совпадает с касательным расслоением $\mathbb{T}(\mathbb{M})$. Форма связности и форма кривизны при этом тождественно равны нулю. □

Продолжение связности не является единственным.

Пример 14.1.8. Рассмотрим расслоение реперов $\mathbb{L}(\mathbb{R}^n)$ над евклидовым пространством $\mathbb{R}^n$. Поскольку $\mathbb{R}^n$ покрывается одной картой, то форма связности $\omega = dx^\alpha \omega_{\alpha a}{}^b \boldsymbol{e}_b{}^a$, где $\boldsymbol{e}_b{}^a$ – базис алгебры Ли $\mathfrak{gl}(n,\mathbb{R})$, взаимно однозначно определятся своими компонентами $\omega_{\alpha a}{}^b$ на $\mathbb{L}$. Чтобы их задать, достаточно задать компоненты локальной формы связности на каком либо сечении, а затем разнести их по всему пространству расслоения с помощью действия структурной группы. Если компоненты локальной формы связности $\omega_{\alpha a}{}^b(x)$ заданы на замкнутом подмножестве $\mathbb{N}\subset\mathbb{R}^n$ дифференцируемым образом, то их можно дифференцируемо продолжить на все евклидово пространство $\mathbb{R}^n$ многими способами. Тем самым мы определим связность на всем $\mathbb{L}(\mathbb{R}^n)$. □

14.2 Форма кривизны и структурное уравнение

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ со связностью $\Gamma:\ \mathbb{P}\ni p\mapsto\mathbb{H}_p\subset\mathbb{T}_p(\mathbb{P})$. Каждая связность Γ на $\mathbb{P}$ взаимно однозначно определяет $\mathfrak{g}$-значную 1-форму связности $\omega=\omega^{\mathrm{A}}L_{\mathrm{A}}$, где $\omega^{\mathrm{A}}\in\Lambda_1(\mathbb{P})$ для всех $\mathrm{A}=1,\ldots,\mathrm{N}$, которой мы поставим в соответствие единственную 2-форму кривизны $R=R^{\mathrm{A}}L_{\mathrm{A}}$, где $R^{\mathrm{A}}\in\Lambda_2(\mathbb{P})$ для всех A, со значениями в алгебре Ли $\mathfrak{g}$. Форма кривизны R является важнейшей характеристикой заданной связности Γ. Для ее определения нам понадобятся новые понятия.

Пусть задано представление ρ структурной группы Ли $\mathbb{G}$ в конечномерном векторном пространстве $\mathbb{V}$,

$$\rho:\quad \mathbb{G}\ni a\mapsto\rho(a)\in\mathsf{aut}\,(\mathbb{V}).$$

При этом $\rho(ab)=\rho(a)\circ\rho(b)$ для всех $a,b\in\mathbb{G}$.

Определение. r-форма $\varphi\in\Lambda_r(\mathbb{P})$ на главном расслоении $\mathbb{P}$ со значениями в векторном пространстве $\mathbb{V}$ называется *псевдотензориальной r-формой типа* $(\rho,\mathbb{V})$, если выполнено условие

$$r_a^*\varphi|_p=\rho(a^{-1})\varphi|_{pa}\qquad\forall a\in\mathbb{G},\ p\in\mathbb{P},$$

где r_a^* – возврат отображения $r_a:\ p\mapsto pa$. Псевдотензориальная r-форма φ называется *тензориальной* r-формой типа $(\rho,\mathbb{V})$, если она *горизонтальна*, т.е. $\varphi(\widetilde{X}_1,\ldots,\widetilde{X}_r)=$

0, как только один из касательных векторов $\widetilde{X}_i \in \mathcal{X}(\mathbb{P})$ вертикален. Под тензориальной 0-формой типа $(\rho, \mathbb{V})$ мы понимаем функцию $\varphi : \mathbb{P} \to \mathbb{V}$, удовлетворяющую условию $\varphi(p) = \rho(a^{-1})\varphi(pa)$ или, что эквивалентно, $\varphi(pa) = \rho(a)\varphi(p)$. □

Пример 14.2.1. В силу свойства 2) предложения 14.1.2 форма связности ω является псевдотензориальной 1-формой типа $(\mathsf{ad}, \mathfrak{g})$. Она не является тензориальной, т.к. не обращается в нуль на вертикальных векторных полях. Для краткости, будем говорить, что форма типа $(\mathsf{ad}, \mathfrak{g})$ является формой типа $\mathsf{ad}\,\mathbb{G}$. □

Пример 14.2.2. Пусть ρ_0 – тривиальное представление группы $\mathbb{G}$ в $\mathbb{V}$, т.е. $\rho_0(a)$ есть тождественное преобразование $\mathbb{V}$ для всех $a \in \mathbb{G}$. Тогда каждая тензориальная r-форма φ типа $(\rho_0, \mathbb{V})$ может быть взаимно однозначно представлена в виде $\varphi = \pi^*\varphi_{\mathbb{M}}$, где $\varphi_{\mathbb{M}}$ некоторая r-форма на базе $\mathbb{M}$ со значениями в $\mathbb{V}$. Для этого достаточно показать, что тензориальная форма φ однозначно определяет форму на базе $\varphi_{\mathbb{M}}$. Поскольку r-форма φ горизонтальна, то ее значения на произвольном наборе касательных векторных полей определяются только горизонтальными компонентами,

$$\varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r) = \varphi(\mathsf{h}\widetilde{X}_1, \ldots, \mathsf{h}\widetilde{X}_r).$$

Так как r-форма φ правоинвариантна, то ее значения на правоинвариантных векторных полях не зависят от точки слоя $p \in \pi^{-1}(x)$. Поэтому r-форма φ взаимно однозначно определяет r-форму $\varphi_{\mathbb{M}}$ на базе следующим равенством

$$\varphi_{\mathbb{M}}(X_1, \ldots, X_r) = \varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r),$$

где $\widetilde{X}_1, \ldots, \widetilde{X}_r$ – произвольный набор правоинвариантных векторных полей таких, что $X_i = \pi_*\widetilde{X}_i$ для всех $i = 1, \ldots, r$. □

Пример 14.2.3. Пусть $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}, \mathbb{P})$ – ассоциированное с $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ расслоение, типичным слоем которого является векторное пространство $\mathbb{V}$, в котором задано представление ρ структурной группы $\mathbb{G}$. Тогда тензориальную r-форму φ типа $(\rho, \mathbb{V})$ можно рассматривать как сопоставление каждой точке $x \in \mathbb{M}$ мультилинейного антисимметричного отображения

$$\tilde{\varphi}_x : \underbrace{\mathbb{T}_x(\mathbb{M}) \times \ldots \times \mathbb{T}_x}_{r} \to \pi_{\mathbb{E}}^{-1}(x) \tag{14.19}$$

с помощью равенства

$$\tilde{\varphi}_x(X_1, \ldots, X_r) = p\Big(\varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r)\Big), \qquad X_i \in \mathbb{T}_x(\mathbb{M}),$$

где p – любая точка слоя $\pi^{-1}(x)$ и $\widetilde{X}_i$ – произвольный касательный вектор в p, для которого $\pi_*\widetilde{X}_i = X_i$ для каждого i. Подробнее, $\varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r)$ есть тогда элемент стандартного слоя $\mathbb{V}$, а p – линейное отображение из $\mathbb{V}$ на слой $\pi_{\mathbb{E}}^{-1}(x)$. Поэтому $p\Big(\varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r)\Big)$ является элементом слоя $\pi_{\mathbb{E}}^{-1}(x)$. Нетрудно проверить, что этот элемент не зависит от выбора точки p и вектора $\widetilde{X}_i$ в слое $\pi^{-1}(x)$.

Обратно. Пусть задано мультилинейное антисимметричное отображение (14.19) для всех $x \in \mathbb{M}$. Тогда можно определить тензориальную r-форму типа $(\rho, \mathbb{V})$ на $\mathbb{P}$ следующим образом

$$\varphi(\widetilde{X}_1, \ldots, \widetilde{X}_r) = p^{-1}\Big(\tilde{\varphi}_x(\pi_*\widetilde{X}_1, \ldots, \pi_*\widetilde{X}_r)\Big), \qquad \widetilde{X}_i \in \mathbb{T}_p(\mathbb{P}),$$

где $x = \pi(p)$. В частности, тензориальная 0-форма типа $(\rho, \mathbb{V})$, т.е. функция $\varphi : \mathbb{P} \to \mathbb{V}$, удовлетворяющая условию $\varphi(pa) = \rho(a)\varphi(p)$, может быть отождествлена с сечением ассоциированного расслоения $\sigma_{\mathbb{E}} : \mathbb{M} \to \mathbb{E}$. Несколько специальных случаев данного примера будет использовано в дальнейшем. □

Продолжим обсуждение связности Γ на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$. Пусть $\mathbb{V}_p(\mathbb{P})$ и $\mathbb{H}_p(\mathbb{P})$ – вертикальное и горизонтальное подпространства в $\mathbb{T}_p(\mathbb{P})$ и $\mathsf{h}:\ \mathbb{T}_p(\mathbb{P})\to\mathbb{H}_p(\mathbb{P})$ – проекция на горизонтальное подпространство. Введем новое понятие внешней ковариантной производной от псевдотензориальной r-формы, с помощью которой будет определена кривизна.

Предложение 14.2.1. *Пусть φ – псевдотензориальная r-форма на $\mathbb{P}$ типа $(\rho,\mathbb{V})$. Тогда:*

1) r-форма $\varphi\mathsf{h}$, определенная равенством

$$\varphi\mathsf{h}(\widetilde{X}_1,\dots,\widetilde{X}_r)=\varphi(\mathsf{h}\widetilde{X}_1,\dots,\mathsf{h}\widetilde{X}_r),\qquad \widetilde{X}_i\in\mathbb{T}_p(\mathbb{P}),\quad i=1,\dots,r,$$

есть тензориальная форма типа $(\rho,\mathbb{V})$;

2) $d\varphi$ есть псевдотензориальная $(r+1)$-форма типа $(\rho,\mathbb{V})$;

3) $(r+1)$-форма $D\varphi$, определенная как $D\varphi=(d\varphi)\mathsf{h}$, является тензориальной формой типа $(\rho,\mathbb{V})$.

Доказательство. Распределение горизонтальных векторных полей инвариантно относительно действия структурной группы, поэтому $r_a\circ\mathsf{h}=\mathsf{h}\circ r_a$. Отсюда следует, что $\varphi\mathsf{h}$ является псевдотензориальной r-формой типа $(\rho,\mathbb{V})$. Очевидно, что

$$\varphi\mathsf{h}(\widetilde{X}_1,\dots,\widetilde{X}_r)=0,$$

если один из касательных векторов $\widetilde{X}_i$ вертикален, и, следовательно, форма $\varphi\mathsf{h}$ горизонтальна. Второе утверждение следует из равенства $r_a^*\circ d=d\circ r_a^*$ для всех $a\in\mathbb{G}$ (4.44). Третье утверждение является прямым следствием двух первых. □

Определение. Форма $D\varphi=(d\varphi)\mathsf{h}$ называется *внешней ковариантной производной* от псевдотензориальной r-формы φ на $\mathbb{P}$. Оператор D называется *внешним ковариантным дифференцированием.* □

Пример 14.2.4 (Локальное рассмотрение). Выпишем ковариантную производную от тензориальной 0-формы на $\mathbb{Q}=\mathbb{U}\times\widetilde{\mathbb{G}}$ типа $(\rho,\mathbb{V})$ в компонентах. Пусть $\boldsymbol{e}_i$, $i=1,\dots,\mathsf{dim}\,\mathbb{V}$, – базис векторного пространства $\mathbb{V}$ и

$$\rho:\quad \mathbb{G}\ni a\mapsto S_i{}^j(a)\in\mathsf{aut}\,\mathbb{V}$$

– представление структурной группы. Пусть $L_{\mathrm{A}j}{}^i$ – представление генераторов (базиса) L_{A} структурной группы. Тогда справедливы формулы (9.63) и (9.64). Поскольку φ – тензориальная 0-форма типа $(\rho,\mathbb{V})$, то она имеет вид

$$\varphi(p)=\varphi^i(p)\boldsymbol{e}_i=\mathring{\varphi}^j(x)S^{-1}{}_j^i(b)\boldsymbol{e}_i,\qquad p=(x,b)\in\mathbb{Q},$$

где $\mathring{\varphi}(x,0)=\mathring{\varphi}^i(x)\boldsymbol{e}_i$ – значение этой функции на нулевом сечении $\sigma_0(x)=(x,0)\in\mathbb{P}$. Внешний дифференциал от компонент φ^i равен

$$\begin{aligned}d\varphi^i&=dx^\alpha\partial_\alpha\mathring{\varphi}^jS^{-1}{}_j^i+db^{\mathrm{A}}\mathring{\varphi}^j\partial_{\mathrm{A}}S^{-1}{}_j^i=dx^\alpha\partial_\alpha\mathring{\varphi}^jS^{-1}{}_j^i+\omega^{*\mathrm{A}}\mathring{\varphi}^jL_{\mathrm{A}}^*S^{-1}{}_j^i=\\&=dx^\alpha\partial_\alpha\mathring{\varphi}^jS^{-1}{}_j^i+\omega^{*\mathrm{A}}\mathring{\varphi}^jS^{-1}{}_j^kL_{\mathrm{A}k}{}^i.\end{aligned}$$

Поскольку горизонтальная составляющая вектора имеет вид (14.7), то значение внешней ковариантной производной на произвольном векторном поле $\widetilde{X}\in\mathbb{X}(\mathbb{P})$ равно

$$D\varphi^i(\widetilde{X})=d\varphi^i(\mathsf{h}\widetilde{X})=\widetilde{X}^\alpha\mathring{\nabla}_\alpha\mathring{\varphi}^jS^{-1}{}_j^i,\tag{14.20}$$

где

$$\mathring{\nabla}_\alpha \mathring{\varphi}^j = \partial_\alpha \mathring{\varphi}^j + \mathring{\varphi}^k \mathring{A}_{\alpha k}{}^j, \qquad \mathring{A}_{\alpha k}{}^j = -\mathring{A}_\alpha{}^{\text{A}} L_{\text{A}k}{}^j.$$

Теперь спустимся на базу $\mathbb{U}$. Рассмотрим два произвольных сечения

$$\sigma: \quad x \mapsto \Big(x, b(x)\Big), \qquad \text{и} \qquad \sigma': \quad x \mapsto \Big(x, b(x)a(x)\Big),$$

связанных калибровочным преобразованием, которое параметризуется функцией $a(x)$. Обозначим значение компонент $\varphi^i(p)$ на этих сечениях через

$$\varphi^i(x) = \varphi^i|_\sigma, \qquad \text{и} \qquad \varphi'^i(x) = \varphi^i|_{\sigma'}. \tag{14.21}$$

Эти функции на $\mathbb{U}$ связаны калибровочным преобразованием

$$\varphi'^i = \varphi^j S^{-1}{}_j{}^i(a), \qquad a = a(x). \tag{14.22}$$

Внешний ковариантный дифференциал $D\varphi$ после проектирования на базу с помощью возвратов отображений σ^* и σ'^* принимает вид

$$\begin{aligned} \sigma^*(D\varphi^i) &= dx^\alpha \nabla_\alpha \varphi^i, \\ \sigma'^*(D\varphi^i) &= dx^\alpha \nabla'_\alpha \varphi'^i, \end{aligned} \tag{14.23}$$

где

$$\begin{aligned} \nabla_\alpha \varphi^i &= \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i, \\ \nabla'_\alpha \varphi'^i &= \partial_\alpha \varphi'^i + \varphi'^j A'_{\alpha j}{}^i. \end{aligned} \tag{14.24}$$

Выше мы ввели обозначения: $A_{\alpha j}{}^i = -A_\alpha{}^{\text{A}} L_{\text{A}j}{}^i$ и $A'_{\alpha j}{}^i = -A'_\alpha{}^{\text{A}} L_{\text{A}j}{}^i$, где $A_\alpha{}^{\text{A}}$ и $A'_\alpha{}^{\text{A}}$ – компоненты локальных форм связности (14.10). Нетрудно проверить, что при калибровочном преобразовании (14.22), (14.14) ковариантная производная ведет себя ковариантно

$$\nabla'_\alpha \varphi'^i = \nabla_\alpha \varphi^j S^{-1}{}_j{}^i. \tag{14.25}$$

Этого следовало ожидать, так как определение ковариантной производной было дано в инвариантном виде. Таким образом, мы видим, что инвариантное определение внешней ковариантной производной для тензориальной 0-формы типа $(\rho, \mathbb{V})$ совпадает с обычным определением ковариантной производной в теории калибровочных полей. В примере 14.2.3 было показано, что тензориальную 0-форму типа $(\rho, \mathbb{V})$ можно отождествить с сечением ассоциированного расслоения, типичным слоем которого является векторное пространство $\mathbb{V}$. Таким образом, формулы (14.24) определяют ковариантные производные от сечений $\sigma_{\mathbb{E}}: \mathbb{M} \to \mathbb{E}$ ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}, \mathbb{P})$. □

В примере 14.2.1 было отмечено, что форма связности ω на $\mathbb{P}$ есть псевдотензориальная 1-форма типа $\mathsf{ad}\,\mathbb{G}$. Используя предложение 14.2.1, дадим

Определение. Внешняя ковариантная производная $R = D\omega$ от формы связности ω является тензориальной 2-формой на $\mathbb{P}$ типа $\mathsf{ad}\,\mathbb{G}$ и называется *формой кривизны* для формы связности ω. Если L_{A} базис алгебры Ли, то $R = R^{\text{A}} L_{\text{A}}$, где $R^{\text{A}} \in \Lambda_2(\mathbb{P})$ для всех $\text{A} = 1, \ldots, \text{N}$. □

Теорема 14.2.1 (Структурное уравнение). *Пусть ω – форма связности на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и $R = D\omega$ – ее форма кривизны. Тогда*

$$d\omega(\widetilde{X},\widetilde{Y}) = -\frac{1}{2}\Big[\omega(\widetilde{X}),\omega(\widetilde{Y})\Big] + R(\widetilde{X},\widetilde{Y}) \tag{14.26}$$

для всех $\widetilde{X},\widetilde{Y} \in \mathbb{T}_p(\mathbb{P})$ и $p \in \mathbb{P}$.

Доказательство. Каждый вектор в $\mathbb{T}_p(\mathbb{P})$ единственным образом разлагается в сумму вертикального и горизонтального векторов. Так как каждый член в (14.26) билинеен и антисимметричен по $\widetilde{X}$ и $\widetilde{Y}$, то достаточно проверить равенство (14.26) в трех случаях.

1) $\widetilde{X}_p$ и $\widetilde{Y}_p$ горизонтальны для всех $p \in \mathbb{P}$. В этом случае $\omega(\widetilde{X}) = \omega(\widetilde{Y}) = 0$ и равенство (14.26) сводится к определению формы кривизны R, т.к. $D\omega = d\omega$ для горизонтальных векторных полей.

2) $\widetilde{X}_p$ и $\widetilde{Y}_p$ вертикальны для всех $p \in \mathbb{P}$. В этом случае их значения в точке p соответствуют некоторым фундаментальным векторным полям, т.е. $\widetilde{X}_p = X^*_p$ и $\widetilde{Y}_p = Y^*_p$ для некоторых $X, Y \in \mathfrak{g}$. Из формулы (4.36) следует равенство

$$\begin{aligned} 2d\omega(X^*,Y^*) &= X^*\Big(\omega(Y^*)\Big) - Y^*\Big(\omega(X^*)\Big) - \omega\Big([X^*,Y^*]\Big) = \\ &= -[X,Y] = -[\omega(X^*),\omega(Y^*)], \end{aligned}$$

т.к. $\omega(X^*) = X$, $\omega(Y^*) = Y$ и $[X^*,Y^*] = [X,Y]^*$. Поскольку для фундаментальных векторных полей $R(X^*,Y^*) = 0$, то формула (14.26) в рассматриваемом случае имеет место.

3) $\widetilde{X}_p$ горизонтально, $\widetilde{Y}_p$ вертикально для всех $p \in \mathbb{P}$. Продолжим $\widetilde{X}_p \in \mathbb{H}_p(\mathbb{P})$ до горизонтального векторного поля $\widetilde{X}$ на $\mathbb{P}$. Это всегда возможно в силу следствия теоремы 13.2.2. Пусть $\widetilde{Y}_p = Y^*_p$ для некоторого $Y \in \mathfrak{g}$. В рассматриваемом случае правая часть равенства (14.26) равна нулю, поэтому достаточно доказать равенство $d\omega(\widetilde{X},Y^*) = 0$. Из формулы (4.36) следует, что

$$2d\omega(\widetilde{X},Y^*) = \widetilde{X}\Big(\omega(Y^*)\Big) - Y^*\Big(\omega(\widetilde{X})\Big) - \omega\Big([\widetilde{X},Y^*]\Big) = -\omega\Big([\widetilde{X},Y^*]\Big).$$

Теперь достаточно доказать

Лемма 14.2.1. *Пусть $\widetilde{X}$ – горизонтальное векторное поле и Y^* – фундаментальное векторное поле, соответствующее элементу алгебры $Y \in \mathfrak{g}$. Тогда коммутатор $[\widetilde{X},Y^*]$ горизонтален.*

Доказательство. Фундаментальное векторное поле Y^* индуцировано действием $r_{a(t)}$, где $a(t)$ – 1-параметрическая подгруппа в $\mathbb{G}$, порожденная элементом алгебры $Y \in \mathfrak{g}$ (экспоненциальное отображение). Поскольку коммутатор векторных полей совпадает с производной Ли, то из (3.122) следует

$$[\widetilde{X},Y^*] = -\,\mathsf{L}_{Y^*}\widetilde{X} = -\lim_{t\to 0}\frac{\widetilde{X} - r_{a(t)*}\widetilde{X}}{t}.$$

Если векторное поле $\widetilde{X}$ горизонтально, то $r_{a(t)*}\widetilde{X}$ тоже горизонтально. Поэтому коммутатор $[\widetilde{X},Y^*]$ горизонтален. □ □

Следствие. Если $\widetilde{X}$ и $\widetilde{Y}$ – горизонтальные векторные поля, то

$$\omega([\widetilde{X},\widetilde{Y}]) = -2R(\widetilde{X},\widetilde{Y}). \quad \square \tag{14.27}$$

Доказательство. Для горизонтальных векторных полей $\omega(\widetilde{X}) = \omega(\widetilde{Y}) = 0$ и

$$2d\omega(\widetilde{X}, \widetilde{Y}) = -\omega([\widetilde{X}, \widetilde{Y}])$$

как следствие (4.36). □

Структурное уравнение (14.26) называют также *структурным уравнением Картана* и часто для простоты записывают в виде

$$d\omega = -\frac{1}{2}[\omega, \omega] + R. \tag{14.28}$$

Приведем еще одну форму записи структурного уравнения. Пусть L_{A}, $\text{A} = 1, \ldots, \text{N}$, – базис алгебры Ли $\mathfrak{g}$ с коммутационными соотношениями

$$[L_{\text{A}}, L_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} L_{\text{C}},$$

где $f_{\text{AB}}{}^{\text{C}}$ – структурные константы алгебры. Тогда формы связности и кривизны можно разложить по базису, $\omega = \omega^{\text{A}} L_{\text{A}}$ и $R = R^{\text{A}} L_{\text{A}}$, и структурные уравнения принимают вид

$$d\omega^{\text{A}} = -\frac{1}{2}\omega^{\text{B}} \wedge \omega^{\text{C}} f_{\text{BC}}{}^{\text{A}} + R^{\text{A}}. \tag{14.29}$$

Замечание. Структурное уравнение (14.26) отличается от формулы Маурера–Картана для групп Ли (9.21) только слагаемым с кривизной. Для фундаментальных векторных полей $R(X^*, Y^*) = 0$ и формулы просто совпадают. □

Теперь спустимся на базу. Пусть задано локальное сечение $\sigma : \mathbb{U} \to \mathbb{P}$ на некоторой окрестности $\mathbb{U} \subset \mathbb{M}$.

Определение. 2-форма на $\mathbb{U}$ со значениями в алгебре Ли $\mathfrak{g}$, определенная возвратом сечения, $R_\sigma = \sigma^* R$, где R – форма кривизны на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, называется *локальной формой кривизны* формы связности ω. □

В компонентах локальная форма кривизны имеет вид

$$R_\sigma = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}{}^{\text{A}} L_{\text{A}}, \tag{14.30}$$

где $F^{\text{A}} = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}{}^{\text{A}} \in \Lambda_2(\mathbb{U})$ для всех $\text{A} = 1, \ldots, \text{N}$.

Пример 14.2.5 (Локальное рассмотрение). В настоящем примере мы выразим компоненты локальной формы кривизны через компоненты локальной формы связности. Чтобы это сделать, сначала определим компоненты формы кривизны на $\mathbb{Q} = \mathbb{U} \times \widetilde{\mathbb{G}}$ через компоненты формы связности. Внешняя производная от формы связности (14.5) имеет вид

$$\begin{aligned} d\omega &= \left[\frac{1}{2} dx^\alpha \wedge dx^\beta \left(\partial_\alpha \omega_\beta{}^{\text{A}} - \partial_\beta \omega_\alpha{}^{\text{A}}\right) - dx^\alpha \wedge db^{\text{B}} \partial_{\text{B}} \omega_\alpha{}^{\text{A}} + d\omega^{*\text{A}}\right] L_{\text{A}} = \\ &= \left[\frac{1}{2} dx^\alpha \wedge dx^\beta \left(\partial_\alpha \omega_\beta{}^{\text{A}} - \partial_\beta \omega_\alpha{}^{\text{A}}\right) + dx^\alpha \wedge \omega^{*\text{B}} \omega_\alpha{}^{\text{C}} f_{\text{BC}}{}^{\text{A}} - \frac{1}{2} \omega^{*\text{B}} \wedge \omega^{*\text{C}} f_{\text{BC}}{}^{\text{A}}\right] L_{\text{A}}, \end{aligned}$$

где $\omega_\alpha{}^{\text{A}}(x, b) = \overset{\circ}{A}_\alpha{}^{\text{B}}(x) S^{-1}{}_{\text{B}}{}^{\text{A}}(b)$. Кроме того мы воспользовались формулой Маурера–Картана (9.22) для структурной группы и правилом дифференцирования матрицы

присоединенного представления (9.40). Теперь нетрудно вычислить значение формы кривизны на векторных полях

$$D\omega(\widetilde{X},\widetilde{Y}) = d\omega(\mathsf{h}\widetilde{X},\mathsf{h}\widetilde{Y}) = \widetilde{X}^{\alpha}\widetilde{Y}^{\beta}R_{\alpha\beta}{}^{\mathrm{A}}L_{\mathrm{A}},$$

где

$$R_{\alpha\beta}{}^{\mathrm{A}} = \partial_{\alpha}\omega_{\beta}{}^{\mathrm{A}} - \partial_{\beta}\omega_{\alpha}{}^{\mathrm{A}} - \omega_{\alpha}{}^{\mathrm{B}}\omega_{\beta}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}} \tag{14.31}$$

– компоненты формы кривизны. Эти компоненты можно выразить через компоненты, заданные на нулевом сечении

$$R_{\alpha\beta}{}^{\mathrm{A}}(x,a) = \overset{\circ}{F}{}_{\alpha\beta}{}^{\mathrm{B}}(x)S^{-1}{}_{\mathrm{B}}^{\mathrm{C}}(a),$$

где

$$\overset{\circ}{F}{}_{\alpha\beta}{}^{\mathrm{A}} = \partial_{\alpha}\overset{\circ}{A}{}_{\beta}{}^{\mathrm{A}} - \partial_{\beta}\overset{\circ}{A}{}_{\alpha}{}^{\mathrm{A}} - \overset{\circ}{A}{}_{\alpha}{}^{\mathrm{B}}\overset{\circ}{A}{}_{\beta}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}} \tag{14.32}$$

– компоненты тензора кривизны для нулевого сечения. Таким образом, форма кривизны имеет вид

$$R = \frac{1}{2}dx^{\alpha}\wedge dx^{\beta}R_{\alpha\beta}{}^{\mathrm{A}}L_{\mathrm{A}}. \tag{14.33}$$

Теперь спустимся на базу $\mathbb{U}$. Для сечений (14.21) получаем следующие выражения для локальных форм кривизны

$$R_{\sigma} = \sigma^{*}R = \frac{1}{2}dx^{\alpha}\wedge dx^{\beta}F_{\alpha\beta}{}^{\mathrm{A}}L_{\mathrm{A}},$$
$$R_{\sigma'} = \sigma'^{*}R = \frac{1}{2}dx^{\alpha}\wedge dx^{\beta}F'_{\alpha\beta}{}^{\mathrm{A}}L_{\mathrm{A}},$$

где

$$F_{\alpha\beta}{}^{\mathrm{A}} = \partial_{\alpha}A_{\beta}{}^{\mathrm{A}} - \partial_{\beta}A_{\alpha}{}^{\mathrm{A}} - A_{\alpha}{}^{\mathrm{B}}A_{\beta}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}} \tag{14.34}$$

и такое же выражение для $F'_{\alpha\beta}{}^{\mathrm{A}}$ через штрихованные компоненты $A'_{\alpha}{}^{\mathrm{A}}$. Нетрудно проверить, что компоненты локальной формы кривизны преобразуются ковариантным образом,

$$F'_{\alpha\beta}{}^{\mathrm{A}} = F_{\alpha\beta}{}^{\mathrm{B}}S^{-1}{}_{\mathrm{B}}^{\mathrm{A}}, \tag{14.35}$$

при калибровочном преобразовании (**??**).

Переходя к присоединенному представлению

$$F_{\alpha\beta}{}^{\mathrm{A}} \mapsto F_{\alpha\beta\mathrm{B}}{}^{\mathrm{C}} := -F_{\alpha\beta}{}^{\mathrm{A}}f_{\mathrm{AB}}{}^{\mathrm{C}},$$

получаем следующее выражение для локальной формы кривизны

$$F_{\alpha\beta\mathrm{A}}{}^{\mathrm{B}} = \partial_{\alpha}A_{\beta\mathrm{A}}{}^{\mathrm{B}} - \partial_{\beta}A_{\alpha\mathrm{A}}{}^{\mathrm{B}} - A_{\alpha\mathrm{A}}{}^{\mathrm{C}}A_{\beta\mathrm{C}}{}^{\mathrm{B}} + A_{\beta\mathrm{A}}{}^{\mathrm{C}}A_{\alpha\mathrm{C}}{}^{\mathrm{B}}. \tag{14.36}$$

Это выражение совпадает с выражением для локальной формы кривизны в аффинной геометрии (6.56), которое было получено ранее, после замены $A_{\alpha\mathrm{A}}{}^{\mathrm{B}} \mapsto \omega_{\alpha a}{}^{b}$. Тем самым мы показали, что аффинная связность, которая была введена ранее независимым образом, является частным случаем связности на главном расслоении общего вида. Опустив матричные индексы, выражение для компонент локальной формы кривизны (14.36) можно записать в виде

$$F_{\alpha\beta} = \partial_{\alpha}A_{\beta} - \partial_{\beta}A_{\alpha} - [A_{\alpha},A_{\beta}], \tag{14.37}$$

где мы для краткости опустили матричные индексы и квадратные скобки обозначают коммутатор матриц. При калибровочном преобразовании (14.15) компоненты локальной формы кривизны преобразуются ковариантно

$$F'_{\alpha\beta} = SF_{\alpha\beta}S^{-1},$$

как и следовало ожидать.

При проведении вычислений на пространстве главного расслоения $\mathbb{P}$, например, в моделях типа Калуцы–Клейна, в касательном расслоении $\mathbb{T}(\mathbb{P})$ удобно использовать базис $\{D_\alpha, L^*_{\rm A}\}$, состоящий из горизонтальных векторных полей

$$D_\alpha = \partial_\alpha - \omega_\alpha{}^{\rm A} L^*_{\rm A}, \tag{14.38}$$

и фундаментальных векторных полей $L^*_{\rm A}$. Этот базис неголономен,

$$[D_\alpha, D_\beta] = -R_{\alpha\beta}{}^{\rm A} L^*_{\rm A}, \tag{14.39}$$
$$[D_\alpha, L^*_{\rm A}] = 0, \tag{14.40}$$
$$[L^*_{\rm A}, L^*_{\rm B}] = f_{\rm AB}{}^{\rm C} L^*_{\rm C}, \tag{14.41}$$

где $R_{\alpha\beta}{}^{\rm A}$ – компоненты формы кривизны (14.31). Второе коммутационное соотношение (14.40) является следствием инвариантности распределения горизонтальных подпространств относительно действия группы справа (напомним, что левоинвариантные векторные поля генерируют действие группы справа, а правоинвариантные – слева, раздел [**?**]). Заметим, что ковариантную производную $D\varphi$ от тензориальной 0-формы типа $(\rho, \mathbb{V})$ можно записать в виде

$$D\varphi = dx^\alpha (D_\alpha \varphi^i) \boldsymbol{e}_i,$$

где векторное поле D_α действует как дифференцирование. Это следует из определения (14.20). В приведенной формуле $\varphi^i = \varphi^i(x, a)$ в отличии от формул (14.24), где ковариантная производная берется от сечений $\varphi^i = \varphi^i\Big(x, \sigma(x)\Big)$. Кроме того, справедливо равенство

$$[D_\alpha, D_\beta]\varphi^i = \varphi^j R_{\alpha\beta j}{}^i, \tag{14.42}$$

где $R_{\alpha\beta j}{}^i = R_{\alpha\beta}{}^{\rm A} L_{{\rm A}j}{}^i$. Эта формула является аналогом формулы (7.81), полученной в аффинной геометрии. В аффинной геометрии в правой части стоит дополнительное слагаемое с тензором кручения. □

Для связи с моделями математической физики дадим

Определение. Компоненты локальной формы кривизны $F_{\alpha\beta}{}^{\rm A}$ называются *напряженностью калибровочного поля* или *напряженностью поля Янга–Миллса*. В электродинамике калибровочной группой является абелева группа $\mathbb{U}(1)$, а компоненты локальной формы кривизны $F_{\alpha\beta}$ называются *напряженностью электромагнитного поля*. □

Продолжим общее рассмотрение.

Теорема 14.2.2 (Тождества Бианки). *Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ задана форма кривизны ω. Тогда форма кривизны $R = D\omega$ удовлетворяет тождествам Бианки:*

$$DR = 0, \tag{14.43}$$

где D – внешняя ковариантная производная.

Доказательство. Из определения внешней ковариантной производной следует, что $DR(\widetilde{X},\widetilde{Y},\widetilde{Z})=0$, если хотя бы один из векторов $\widetilde{X},\widetilde{Y},\widetilde{Z}$ вертикален. Поэтому достаточно доказать, что $dR(\widetilde{X},\widetilde{Y},\widetilde{Z})=0$, когда все три вектора горизонтальны. Возьмем внешнюю производную от структурного уравнения (14.29):

$$0=dd\omega^{\mathrm{A}}=-\frac{1}{2}d\omega^{\mathrm{B}}\wedge\omega^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}}+\frac{1}{2}\omega^{\mathrm{B}}\wedge d\omega^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}}+dR^{\mathrm{A}}.$$

Поскольку $\omega^{\mathrm{A}}(\widetilde{X})=0$, если вектор $\widetilde{X}$ горизонтален, то

$$dR^{\mathrm{A}}(\widetilde{X},\widetilde{Y},\widetilde{Z})=0$$

если все три вектора горизонтальны. □

Теорема 14.2.3. *Пусть ω – форма связности на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и φ – произвольная тензориальная 1-форма типа* $\mathsf{ad}\,\mathbb{G}$. *Тогда*

$$D\varphi(\widetilde{X},\widetilde{Y})=d\varphi(\widetilde{X},\widetilde{Y})+\frac{1}{2}[\varphi(\widetilde{X}),\omega(\widetilde{Y})]+\frac{1}{2}[\omega(\widetilde{X}),\varphi(\widetilde{Y})]$$

для всех $\widetilde{X},\widetilde{Y}\in\mathbb{T}_p(\mathbb{P})$ и $p\in\mathbb{P}$.

Доказательство. Аналогично доказательству теоремы 14.2.1 о структурном уравнении. См., например, [**?**]. □

Пример 14.2.6 (Локальное рассмотрение). Запишем тождества Бианки в компонентах. Внешняя производная от компонент формы кривизны (14.33) имеет вид

$$\begin{aligned}dR^{\mathrm{A}}&=\frac{1}{2}dx^{\alpha}\wedge dx^{\beta}\wedge dx^{\gamma}\partial_{\alpha}R_{\beta\gamma}{}^{\mathrm{A}}+\frac{1}{2}dx^{\alpha}\wedge dx^{\beta}\wedge\omega^{*\mathrm{B}}L^{*}_{\mathrm{B}}R_{\alpha\beta}{}^{\mathrm{A}}=\\&=\frac{1}{2}dx^{\alpha}\wedge dx^{\beta}\wedge dx^{\gamma}\partial_{\alpha}R_{\beta\gamma}{}^{\mathrm{A}}-\frac{1}{2}dx^{\alpha}\wedge dx^{\beta}\wedge\omega^{*\mathrm{B}}R_{\alpha\beta}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}}.\end{aligned}$$

Ее значения на горизонтальных векторных полях равно

$$dR^{\mathrm{A}}(\mathsf{h}\widetilde{X},\mathsf{h}\widetilde{Y},\mathsf{h}\widetilde{Z})=3\widetilde{X}^{\alpha}\widetilde{Y}^{\beta}\widetilde{Z}^{\alpha}D_{[\alpha}R_{\beta\gamma]}{}^{\mathrm{A}},$$

где квадратные скобки обозначают антисимметризацию по трем индексам, а

$$D_{\alpha}R_{\beta\gamma}{}^{\mathrm{A}}=\partial_{\alpha}R_{\beta\gamma}{}^{\mathrm{A}}+\omega_{\alpha}{}^{\mathrm{B}}R_{\beta\gamma}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}}.$$

Таким образом, тождества Бианки в компонентах имеют вид

$$D_{\alpha}R_{\beta\gamma}{}^{\mathrm{A}}+D_{\beta}R_{\gamma\alpha}{}^{\mathrm{A}}+D_{\gamma}R_{\alpha\beta}{}^{\mathrm{A}}=0. \tag{14.44}$$

Если задано локальное сечение расслоения, то эти тождества можно спустить на базу, используя возврат отображения. Тогда тождества Бианки для компонент локальных форм кривизны и связности примут вид

$$\nabla_{\alpha}F_{\beta\gamma}{}^{\mathrm{A}}+\nabla_{\beta}F_{\gamma\alpha}{}^{\mathrm{A}}+\nabla_{\gamma}F_{\alpha\beta}{}^{\mathrm{A}}=0,$$

где

$$\nabla_{\alpha}F_{\beta\gamma}{}^{\mathrm{A}}=\partial_{\alpha}F_{\beta\gamma}{}^{\mathrm{A}}+A_{\alpha}{}^{\mathrm{B}}F_{\beta\gamma}{}^{\mathrm{C}}f_{\mathrm{BC}}{}^{\mathrm{A}}$$

и напряженность $F_{\alpha\beta}{}^{\mathrm{A}}$ имеет вид (14.34). Именно в таком виде они, как правило, используются в приложениях. □

14.3 Параллельный перенос

Пусть на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ задана связность $\Gamma:\ \mathbb{P}\ni p\mapsto\mathbb{H}_p\subset\mathbb{T}_p(\mathbb{P})$. Определим понятие параллельного переноса слоя $\pi^{-1}(x_0)$ над точкой базы $x_0\in\mathbb{M}$ вдоль произвольной кусочно дифференцируемой кривой

$$\gamma:\quad [0,1]\ni t\mapsto x(t)\in\mathbb{M} \tag{14.45}$$

с началом в точке x_0. Для наших целей достаточно рассматривать кусочно дифференцируемые кривые класса $\mathcal{C}^1$.

Определение. *Горизонтальной кривой* в $\mathbb{P}$ называется кусочно дифференцируемая кривая, все касательные векторы которой горизонтальны. *Горизонтальным лифтом* или *подъемом* (или просто *лифтом*) кривой γ (14.45), заданной на базе $\mathbb{M}$, называется горизонтальная кривая в $\mathbb{P}$,

$$\tilde\gamma:\quad [0,1]\ni t\mapsto p(t)\in\mathbb{P},$$

такая, что $\pi\circ\tilde\gamma=\gamma$. □

Понятие горизонтального лифта кривой соответствует понятию лифта векторного поля. Действительно, если $\widetilde{X}\in\mathcal{X}(\mathbb{P})$ – лифт дифференцируемого векторного поля X, заданного на базе $\mathbb{M}$, то интегральная кривая $\tilde\gamma$ векторного поля $\widetilde{X}$, проходящая через точку $p_0\in\mathbb{P}$, есть горизонтальный лифт интегральной кривой γ поля X, проходящей через точку $x_0=\pi(p_0)$.

Предложение 14.3.1. *Пусть $\gamma=x(t)$, $t\in[0,1]$, – кусочно дифференцируемая кривая класса $\mathcal{C}^1$ в $\mathbb{M}$ с началом в точке $x_0\in\mathbb{M}$. Тогда для произвольной точки слоя $p_0\in\pi^{-1}(x_0)$ существует единственный горизонтальный лифт $\tilde\gamma=p(t)$ кривой γ с началом в точке p_0.*

Доказательство. Состоит в явном построении лифта $\tilde\gamma$. См., например, [**?**]. □

Используя предложение 14.3.1, определим параллельное перенесение слоев следующим образом.

Определение. Пусть $\gamma=x(t)$ – кривая в $\mathbb{M}$ с началом и концом в точках x_0 и x_1. Пусть $\tilde\gamma$ – единственный горизонтальный лифт кривой γ, который начинается в точке p_0, находящейся в слое $\pi^{-1}(x_0)$. Лифт $\tilde\gamma$ имеет конечную точку $p_1\in\mathbb{P}$ такую, что $\pi(p_1)=x_1$. Меняя начальную точку p_0 в слое $\pi^{-1}(x_0)$, мы получаем отображение слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_1)$, которое переводит точку p_0 в p_1. Это отображение называется *параллельным переносом слоя* из точки x_0 в точку x_1 вдоль кривой γ. Параллельный перенос слоев обозначается той же буквой, что и кривая, $\gamma:\ \pi^{-1}(x_0)\to\pi^{-1}(x_1)$. □

Параллельный перенос слоев является изоморфизмом, что вытекает из следующего утверждения.

Предложение 14.3.2. *Параллельный перенос слоя $\gamma:\ \pi^{-1}(x_0)\to\pi^{-1}(x_1)$ вдоль любой кривой перестановочен с действием структурной группы $\mathbb{G}$ на $\mathbb{P}$: $\gamma\circ r_a=r_a\circ\gamma$ для всех $a\in\mathbb{G}$.*

Доказательство. Это следует из того, что правое действие структурной группы r_a отображает каждую горизонтальную кривую в горизонтальную, см. рис.14.1. □

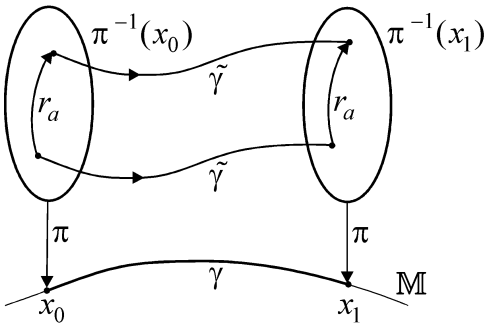


Рис. 14.1: Параллельный перенос слоев вдоль кривой γ.

Параллельный перенос слоя вдоль кривой γ не зависит от выбора параметризации кривой. Кроме того, если слой переносится из точки x_0 в точку x_1 параллельно вдоль кривой γ, то он параллельно переносится вдоль этой же кривой из точки x_0 в любую промежуточную точку $x(t)$, $t \in [0,1]$, на γ.

Замечание. Если точки x_0 и x_1 в базе $\mathbb{M}$ фиксированы, то в общем случае параллельный перенос слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_1)$ зависит от кривой γ, соединяющей эти точки. Для односвязных баз эта зависимость характеризуется формой кривизны R формы связности ω и будет обсуждаться в следующих разделах. □

При рассмотрении фундаментальной группы в разделе 11.3 мы определили произведение путей (кривых) $\gamma_2 \circ \gamma_1$ как последовательный проход вдоль путей γ_1 и γ_2 (11.4) и обратный путь γ^{-1} как путь γ, проходимый в обратном направлении (11.5). Следующее предложение очевидно.

Предложение 14.3.3. *1) Если γ_1 – путь из x_0 в x_1 и γ_2 – путь из x_1 в x_2, то параллельный перенос слоя $\pi^{-1}(x_0)$ в слой $\pi^{-1}(x_2)$ вдоль произведения путей $\gamma_2 \circ \gamma_1$ равен произведению отображений слоев $\gamma_2 \circ \gamma_1 :\ \pi^{-1}(x_0) \to \pi^{-1}(x_2)$.*

2) Если γ^{-1} – обратный путь для пути γ из точки x_0 в точку x_1, то параллельный перенос слоя $\pi^{-1}(x_1)$ в слой $\pi^{-1}(x_0)$ вдоль пути γ^{-1} является обратным отображением $\gamma^{-1} = \gamma^{-1} :\ \pi^{-1}(x_1) \to \pi^{-1}(x_0)$.

14.4 Группы голономии

Пусть задано главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связностью Γ. Используя понятие параллельного переноса, определим группу голономии данной связности Γ.

Обозначим через $\Omega(\mathbb{M}, x)$ множество замкнутых кусочно дифференцируемых кривых (петель) на базе $\mathbb{M}$ с началом и концом в точке $x \in \mathbb{M}$. Подмножество, состоящее из путей, гомотопных постоянному пути в точке x, обозначим $\Omega_0(\mathbb{M}, x) \subset \Omega(\mathbb{M}, x)$. Произведение и обратный путь для всех путей из $\Omega(\mathbb{M}, x)$ были определены в разделе 11.3. В разделе 14.3 было показано, что параллельный перенос слоя $\pi^{-1}(x)$ вдоль замкнутого пути $\gamma \in \Omega(\mathbb{M}, x)$ есть изоморфизм слоя $\pi^{-1}(x)$ на себя. В общем случае этот изоморфизм будет нетривиален, т.к. после параллельного переноса вдоль замкнутого пути слой может повернуться. Множество всех таких изоморфизмов образует группу в силу предложения 14.3.3.

Определение. Группа, состоящая из изоморфизмов слоя $\pi^{-1}(x)$, которые соответствуют параллельным переносам данного слоя вдоль всех замкнутых кусочно дифференцируемых путей $\gamma \in \Omega(\mathbb{M}, x)$, называется *группой голономии* $\Phi(x)$ связности Γ в точке $x \in \mathbb{M}$. Подгруппа $\Phi_0(x) \subset \Phi(x)$, соответствующая параллельным переносам вдоль замкнутых путей, стягиваемых в точку, $\gamma \in \Omega_0(\mathbb{M}, x)$, называется *суженной группой голономии* связности Γ в точке $x \in \mathbb{M}$. □

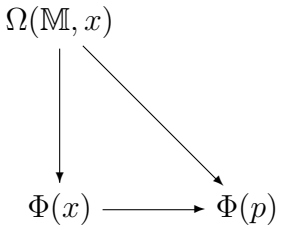
Группу голономии $\Phi(x)$ и суженную группу голономии $\Phi_0(x)$ можно считать подгруппами в структурной группе $\mathbb{G}$ следующим образом.

Определение. Зафиксируем некоторую точку слоя $p \in \pi^{-1}(x)$. После параллельного переноса слоя вдоль пути $\gamma \in \Omega(\mathbb{M}, x)$ эта точка отобразится в некоторую точку $\gamma(p) = pa \in \pi^{-1}(x)$, где $a \in \mathbb{G}$ – некоторый элемент структурной группы. Если задан другой путь $\gamma' \in \Omega(\mathbb{M}, x)$, которому соответствует элемент $b \in \mathbb{G}$, то произведение путей $\gamma' \circ \gamma$ определяет элемент $ba \in \mathbb{G}$, поскольку

$$\gamma' \circ \gamma(p) = \gamma'(pa) = \Big(\gamma'(p)\Big)a = pba.$$

По предложению 14.3.3 множество элементов $a \in \mathbb{G}$, определенных всеми путями $\gamma \in \Omega(x)$, образует группу, которая называется *группой голономии* $\Phi(p)$ связности Γ в точке $p \in \mathbb{P}$. Замкнутым путям $\gamma \in \Omega_0(\mathbb{M}, x)$ соответствует *суженная группа голономии* $\Phi_0(p)$ в точке $p \in \mathbb{P}$. □

Замечание. $\Phi(x)$ есть группа изоморфизмов слоя $\pi^{-1}(x)$ на себя, а $\Phi(p)$ есть подгруппа в $\mathbb{G}$. Выше мы построили единственный изоморфизм из $\Phi(x)$ на $\Phi(p)$, который делает коммутативной следующую диаграмму:



□

Группу голономии $\Phi(p)$ можно определить другим образом. Введем на пространстве главного расслоения отношение эквивалентности $p \sim q$, где $p, q \in \mathbb{P}$, если точки p и q можно можно соединить горизонтальной кривой. При этом точки p и q не обязательно лежат в одном слое. Нетрудно проверить, что это действительно отношение эквивалентности. Тогда группа голономии $\Phi(p)$ совпадает с множеством тех элементов $a \in \mathbb{G}$, для которых $p \sim pa$. Легко проверить, что это множество элементов образует подгруппу в $\mathbb{G}$, т.к. $p \sim q$ влечет за собой $pa \sim qa$.

Предложение 14.4.1. *Пусть дано главное расслоение* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *со связностью* Γ. *Тогда:*

1) *Если* $q = pa$, $a \in \mathbb{G}$, *то* $\Phi(q) = \mathsf{ad}\,(a^{-1})\Phi(p)$, *т.е. группы голономии точек одного слоя* $\Phi(q)$ *и* $\Phi(p)$ *сопряжены в* $\mathbb{G}$. *Аналогично,* $\Phi_0(q) = \mathsf{ad}\,(a^{-1})\Phi_0(p)$.
2) *Если точки* $p, q \in \mathbb{P}$ *можно соединить горизонтальной кривой, т.е.* $p \sim q$, *то* $\Phi(p) = \Phi(q)$ *и* $\Phi_0(p) = \Phi_0(q)$.

Доказательство. 1). Пусть $b \in \Phi(p)$, т.е. $p \sim pb$. Тогда $pa \sim pba$ и, следовательно, $q \sim qa^{-1}ba$. Поэтому $\mathsf{ad}\,(a^{-1})b = a^{-1}ba \in \Phi(q)$. Отсюда вытекает, что $\Phi(q) = \mathsf{ad}\,(a^{-1})\Phi(p)$ и $\Phi_0(q) = \mathsf{ad}\,(a^{-1})\Phi_0(p)$.

2) Отношение $p \sim q$ влечет за собой $pb \sim qb$. Из транзитивности отношения эквивалентности $\sim$ следует, что $p \sim pb$ тогда и только тогда, когда $q \sim qb$, т.е. $b \in \Phi(p)$ тогда и только тогда, когда $b \in \Phi(q)$. Тем самым $\Phi(p) = \Phi(q)$. Чтобы доказать равенство $\Phi_0(p) = \Phi_0(q)$, допустим, что $\tilde{\delta}$ – горизонтальная кривая в $\mathbb{P}$ из p в q. Если $b \in \Phi_0(p)$, то существует горизонтальная кривая $\tilde{\gamma}$ в $\mathbb{P}$ из p в pb такая, что кривая в базе $\pi\tilde{\gamma}$ является замкнутым путем с началом и концом в точке $\pi(p)$, которая гомотопна постоянному пути в точке $\pi(p)$. Тогда композиция $(r_b\tilde{\delta}) \circ \tilde{\gamma} \circ \tilde{\delta}^{-1}$ есть горизонтальная кривая в $\mathbb{P}$ из q в qb и ее проекция на базу $\mathbb{M}$ есть замкнутый путь с началом и концом в точке $\pi(q)$, который гомотопен постоянному пути. Поэтому $b \in \Phi_0(q)$. □

Если база $\mathbb{M}$ связна, то для любой пары точек $p, q \in \mathbb{P}$ найдется элемент $a \in \mathbb{G}$ такой, что $q \sim pa$. Поэтому из предложения 14.4.1 следует, что группы голономии $\Phi(p)$ для всех точек $p \in \mathbb{P}$ сопряжены друг другу в $\mathbb{G}$ и поэтому изоморфны. По тем же причинам все суженные группы голономии $\Phi_0(p)$ также изоморфны друг другу.

Итак, мы определили группу голономии $\Phi(p)$, суженную группу голономии $\Phi_0(p)$ и показали, что с точностью до преобразования подобия они не зависят от точки расслоения $\mathbb{P}$. Теперь сформулируем несколько общих свойств групп голономий.

Теорема 14.4.1. *Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связной базой $\mathbb{M}$. Пусть $\Phi(p)$ и $\Phi_0(p)$ – группа голономии и суженная группа голономии связности Γ в точке $p \in \mathbb{P}$. Тогда:*

1) $\Phi_0(p)$ есть связная подгруппа Ли в $\Phi(p)$;

2) $\Phi_0(p)$ есть нормальная подгруппа в $\Phi(p)$ и фактор группа $\Phi(p)/\Phi_0(p)$ счетна.

Доказательство. Используется паракомпактность базы $\mathbb{M}$. См., например, [?]. □

Следствие. Группа голономии $\Phi(p)$ является подгруппой Ли в структурной группе $\mathbb{G}$ с компонентой единицы $\Phi_0(p)$. В частности, $\dim \Phi(p) = \dim \Phi_0(p)$. □

При определении групп голономий мы не оговорили класс дифференцируемости рассматриваемых кусочно дифференцируемых кривых $\gamma \in \Omega(\mathbb{M}, x)$. Чем ниже класс дифференцируемости, тем больше множество кривых. Поэтому могло бы оказаться так, что группы голономии зависят от класса дифференцируемости кривых. Это оказывается не верно. Пусть $\Omega^k(\mathbb{M}, x)$ – множество замкнутых кусочно дифференцируемых путей в $\mathbb{M}$ с началом и концом в точке $x \in \mathbb{M}$ класса $\mathcal{C}^k$. Обозначим соответствующую группу голономии через $\Phi^k(p)$. Очевидно, что $\Phi^1(p) \supset \Phi^2(p) \supset \ldots \supset \Phi^\infty(p)$. Верны также и обратные включения.

Теорема 14.4.2 (Номидзу, Одзеки). *Все группы голономии $\Phi^k(p)$, $1 \leq k \leq \infty$, совпадают.*

Доказательство. См. [?]. □

Следствие. Все суженные группы голономии $\Phi_0^k(p)$, $1 \leq k \leq \infty$, совпадают. □

Доказательство. Согласно теореме 14.4.1 суженная группа голономии $\Phi_0^k(p)$ есть связная компонента единицы группы $\Phi^k(p)$. □

Замечание. В случае, когда $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ является вещественно аналитическим главным расслоением с аналитической связностью Γ, можно определить группу голономии $\Phi^\omega(p)$, используя только кусочно аналитические кривые. Можно доказать, что $\Phi^\omega(p) = \Phi^1(p)$ и $\Phi_0^\omega(p) = \Phi_0^1(p)$ [?]. □

Поскольку группы голономии не зависят от класса дифференцируемости путей, то в дальнейшем мы не будем его указывать.

14.5 Петля Вильсона

Настоящий раздел посвящен одному из способов вычисления группы голономии. Пусть задано главное расслоение $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ со связностью Г. Рассмотрим окрестность $\mathbb{U}\subset\mathbb{M}$ с координатами x^α, $\alpha=1,\dots,n$, содержащую точку базы $x_0\in\mathbb{M}$, и отождествим подрасслоение $\pi^{-1}(\mathbb{U})$ с прямым произведением $\mathbb{U}\times\mathbb{G}$. Тогда точка подрасслоения задается парой элементов $p=(x,a)\in\mathbb{U}\times\mathbb{G}$. Пусть $\gamma=x(t)$ – произвольная кривая в $\mathbb{U}$ с началом в точке x_0 и $\tilde\gamma=\big(\tilde x(t),a(t)\big)$ – ее единственный горизонтальный лифт с началом в точке $p_0=(x_0,a_0)$. При этом функция $a(t)$ определяет некоторую кривую в структурной группе $\mathbb{G}$ с началом в точке a_0. В инвариантном виде мы пишем

$$\tilde\gamma=\mathring{\gamma}\circ a(t),\tag{14.46}$$

где $\mathring{\gamma}=\big(x(t),e\big)$ – опорная кривая в $\pi^{-1}(\mathbb{U})$, лежащая в нулевом сечении, для которой $\pi\big(\mathring{\gamma}\big)=\gamma$. При этом мы рассматриваем равенство (14.46) как уравнение на $a(t)$ при заданной кривой γ, которая однозначно определяет опорную кривую $\mathring{\gamma}$.

Касательный вектор к горизонтальной кривой $\tilde\gamma$ имеет вид $\dot{\tilde\gamma}=(\dot x,\dot a)$. Уравнение для $a(t)$ получается из условия горизонтальности. Рассмотрим окрестность единицы группы, где форма связности имеет вид (14.5). Горизонтальность касательного вектора к кривой записывается в виде равенства

$$\omega(\dot{\tilde\gamma})=(\dot x^\alpha\omega_\alpha{}^{\text{A}}+\dot a^{\text{B}}L^{-1}{}_{\text{B}}{}^{\text{A}})L_{\text{A}}=0.$$

Отсюда следует система уравнений на $a^{\text{A}}(t)$:

$$\dot x^\alpha\mathring{A}_\alpha{}^{\text{B}}S^{-1}{}_{\text{B}}{}^{\text{A}}+\dot a^{\text{B}}L^{-1}{}_{\text{B}}{}^{\text{A}}=0,\tag{14.47}$$

где $\mathring{A}_\alpha{}^{\text{A}}(x)$ – компоненты локальной формы связности на нулевом сечении (поле Янга–Миллса) и $S_{\text{B}}{}^{\text{A}}(a)$ – матрица присоединенного представления, соответствующая элементу $a\in\mathbb{G}$.

Уравнение (14.47) неудобно для приложений, т.к. содержит матрицу $L_{\text{B}}{}^{\text{A}}(a)$, которая определена только в окрестности единицы группы. Чтобы устранить это неудобство, перейдем к какому либо представлению структурной группы

$$\rho:\quad\mathbb{G}\ni a\mapsto\Big(S^{-1}{}_j{}^i(a)\Big)\in\mathsf{aut}\,\mathbb{V},\qquad i,j=1,\dots,\mathsf{dim}\,\mathbb{V}.$$

Мы выбрали представление в виде обратных матриц, чтобы не менять общепринятого определения хронологического произведения, которое будет дано ниже. Теперь заметим, что

$$\dot S^{-1}{}_i{}^j=\dot a^{\text{A}}\partial_{\text{A}}S^{-1}{}_i{}^j=\dot a^{\text{B}}L^{-1}{}_{\text{B}}{}^{\text{A}}L_{\text{A}}S^{-1}{}_i{}^j=\dot a^{\text{B}}L^{-1}{}_{\text{B}}{}^{\text{A}}S^{-1}{}_i{}^kL_{\text{A}k}{}^j,$$

где $L_{\text{A}i}{}^j$ – представление генераторов L_{A} алгебры Ли $\mathfrak{g}$, и мы воспользовались правилом дифференцирования матриц представления (9.62). Умножив уравнение (14.47) на $S^{-1}{}_i{}^kL_{\text{A}k}{}^j$ и воспользовавшись инвариантностью генераторов $L_{\text{A}i}{}^j$ (9.64), получаем уравнение на матрицу представления

$$\dot S^{-1}{}_i{}^j-\dot x^\alpha\mathring{A}_{\alpha i}{}^kS^{-1}{}_k{}^j=0,$$

где $\mathring{A}_{\alpha i}{}^j=-\mathring{A}_\alpha{}^{\text{A}}L_{\text{A}i}{}^j$. Это уравнение можно переписать в почти ковариантном виде

$$\dot x^\alpha\Big(\partial_\alpha S^{-1}{}_i{}^j-\mathring{A}_{\alpha i}{}^kS^{-1}{}_k{}^j\Big)=0.$$

Для ковариантности не хватает одного слагаемого с калибровочным полем для индекса j. Опустив, для краткости, матричные индексы, получаем уравнение

$$\dot{S}^{-1} = \dot{x}^\alpha \mathring{A}_\alpha S^{-1}, \tag{14.48}$$

которое можно записать в интегральном виде:

$$S^{-1}(t) = S_0^{-1} + \int_0^t ds\, \dot{x}^\alpha \mathring{A}_\alpha S^{-1} = S_0^{-1} + \int_{x(0)}^{x(t)} dx^\alpha \mathring{A}_\alpha S^{-1}.$$

где $S_0^{-1} = S^{-1}(a_0)$, точка обозначает дифференцирование по s и второй интеграл берется вдоль кривой γ от точки $x(0)$ до точки $x(t)$. Решение этого уравнения записывается в виде упорядоченной P-экспоненты

$$S^{-1}(t) = \mathsf{P}\exp\left(\int_0^t ds\, \dot{x}^\alpha \mathring{A}_\alpha\right) S_0^{-1}, \tag{14.49}$$

которая определена разложением в ряд

$$\mathsf{P}\exp\left(\int_0^t ds\, \dot{x}^\alpha \mathring{A}_\alpha\right) = 1 + \int_0^t ds\, \dot{x}^\alpha \mathring{A}_\alpha + \int_0^t ds_1 \int_0^{s_1} ds_2\, \dot{x}^{\alpha_1} \mathring{A}_{\alpha_1} \dot{x}^{\alpha_2} \mathring{A}_{\alpha_2} + \dots, \tag{14.50}$$

где

$$\dot{x}^{\alpha_k} \mathring{A}_{\alpha_k} = \frac{dx^{\alpha_k}(s_k)}{ds_k} \mathring{A}_{\alpha_k}(s_k), \qquad \forall k = 1, 2, \dots .$$

Напомним общее определение P-произведения.

Определение. Пусть задано семейство операторов $A(t)$, непрерывно зависящих от вещественного параметра t, тогда

$$\mathsf{P}[A(t_1)A(t_2)] = \begin{cases} A(t_1)A(t_2), & t_1 \geq t_2, \\ A(t_2)A(t_1), & t_1 < t_2 \end{cases} \tag{14.51}$$

называется P-*произведением* или *упорядоченным произведением* операторов $A(t_1)$ и $A(t_2)$. □

Если операторы для различных точек коммутируют, то P-произведение совпадает с обычным произведением.

Замечание. В квантовой теории поля роль параметра t часто играет время. Поэтому P-произведение называют также *хронологическим произведением.* □

Нетрудно проверить равенство

$$\int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_{k-1}} ds_k A(s_1)A(s_2)\dots A(s_k) = \\ = \frac{1}{k!} \int_0^t ds_1 \int_0^t ds_2 \dots \int_0^t ds_k \mathsf{P}[A(s_1)A(s_2)\dots A(s_k)].$$

Поэтому разложение (14.50) имеет место.

Известно, что ряд (14.50) равномерно сходится в шаре произвольного радиуса.

Продолжим общее построение. Если представление структурной группы является точным, то матрица представления $S^{-1}(a)$ однозначно определяет элемент структурной группы $a = \rho^{-1}(S^{-1})$. В этом случае решение (14.49) определяет кривую $a(t)$ с началом в точке $a_0 \in \mathbb{G}$. Меняя точку a_0, мы получаем отображение

$$\pi^{-1}(x_0) \ni (x_0, a_0) \mapsto \Big(x(t), a(t)\Big) \in \pi^{-1}\Big(x(t)\Big), \qquad \forall a_0 \in \mathbb{G}. \tag{14.52}$$

Поскольку кривая $\tilde{\gamma}$ горизонтальна, то это отображение задает параллельный перенос слоя $\pi^{-1}(x_0)$ главного расслоения $\mathbb{P}$ вдоль пути $\gamma \in \mathbb{M}$.

Таким образом, решение (14.49) определяет параллельный перенос слоев (14.52) через компоненты связности $\overset{\circ}{A}_\alpha$ на опорном нулевом сечении. Отображение (14.52) можно построить для произвольного опорного сечения следующим образом. Пусть задано произвольное сечение $\sigma(x) = \Big(x, b(x)\Big)$. Тогда компоненты локальной формы связности для рассматриваемого представления на этом сечении имеют вид (14.14)

$$A_\alpha = S_b \overset{\circ}{A}_\alpha S_b^{-1} + \partial_\alpha S_b S_b^{-1}, \qquad S_b := S(b).$$

Для горизонтальной кривой $\tilde{\gamma}$ имеем

$$\tilde{\gamma} = \Big(x(t), a(t)\Big) = \Big(x(t), b(t)c(t)\Big), \tag{14.53}$$

где $c(t)$ – некоторая новая кривая в $\mathbb{G}$, связывающая кривую $\Big(x(t), b(t)\Big)$ на сечении σ с горизонтальной кривой $\tilde{\gamma}$. Поскольку $S_a^{-1} = S_b^{-1} S_c^{-1}$, то простые вычисления приводят (14.48) к уравнению

$$\dot{S}_c^{-1} = \dot{x}^\alpha A_\alpha S_c^{-1},$$

определяющему кривую $c(t)$ с начальным условием $c_0 = b_0^{-1} a_0$. Решение этого уравнения также дается упорядоченной экспонентой

$$S_c^{-1}(t) = \mathsf{P}\,\mathsf{exp}\,\left(\int_0^t ds\, \dot{x}^\alpha A_\alpha\right) S_{c(0)}^{-1}.$$

Эта формула также определяет параллельный перенос (14.52), что следует из (14.53).

Поскольку $S_0^{-1} = S_{b(0)}^{-1} S_{c(0)}^{-1}$, то простые вычисления приводят к следующему правилу преобразования упорядоченной экспоненты при изменении сечения $\sigma_0 \mapsto \sigma$

$$\mathsf{P}\,\mathsf{exp}\,\left(\int_0^t ds\, \dot{x}^\alpha \overset{\circ}{A}_\alpha\right) = S(b_0)\mathsf{P}\,\mathsf{exp}\,\left(\int_0^t ds\, \dot{x}^\alpha A_\alpha\right) S^{-1}(b_t). \tag{14.54}$$

Этот закон преобразования похож на тензорный, однако таковым не является, потому что слева и справа стоят матрицы преобразования, взятые в различных точках: в начале и конце пути.

Если задано два произвольных сечения (14.9) то имеет место аналогичная формула.

Рассмотрим замкнутый путь в базе $\gamma \in \Omega(\mathbb{U}, x_0)$. Поскольку мы ограничились координатной окрестностью $\mathbb{U}$, то все пути стягиваемы к точке x_0 (гомотопны постоянному пути в x_0). Для этих путей $S(1) = S(0)$ и P-экспонента при калибровочном преобразовании $\sigma_0 \mapsto \sigma$ преобразуется по тензорному закону

$$\mathsf{P}\,\mathsf{exp}\,\left(\oint_\gamma dx^\alpha \overset{\circ}{A}_\alpha\right) = S(b_0)\mathsf{P}\,\mathsf{exp}\,\left(\oint_\gamma dx^\alpha A_\alpha\right) S^{-1}(b_0). \tag{14.55}$$

При параллельном переносе слоя $\pi^{-1}(x_0)$ вдоль замкнутого пути γ точка $p_0 = (x_0, a_0)$ отобразится в точку $p_1 = (x_0, a_1)$. Пусть $a_0 = e$, тогда $S_0 = \mathbb{1}$. В этом случае

$$\hat{a} = \rho^{-1}\left[\mathsf{P}\,\mathsf{exp}\,\left(\oint_\gamma dx^\alpha \overset{\circ}{A}_\alpha\right)\right] \quad \in \Phi_0\Big(\sigma_0(x_0)\Big) \tag{14.56}$$

– элемент суженной группы голономии $\Phi_0\Big(\sigma_0(x_0)\Big)$. Он определяется калибровочным полем, соответствующим нулевому сечению σ_0, которое проходит через начальную точку кривой (x_0, e). Если $a_0 \neq e$, то элемент группы голономии имеет вид $a_0^{-1}a_1 \in \Phi_0(p_0)$ и равен

$$\rho^{-1}\left[S_0^{-1}\mathsf{P}\,\mathsf{exp}\left(\oint_\gamma dx^\alpha \mathring{A}_\alpha\right)S_0\right] = a_0^{-1}\hat{a}a_0 \quad \in \Phi_0\Big(\sigma(x_0)\Big).$$

То есть он сопряжен элементу $\hat{a}$ в соответствии с утверждением 1) предложения 14.4.1.

Поскольку P-экспонента для замкнутого пути при изменении сечения преобразуется по тензорному закону (14.55), то для произвольного сечения σ P-экспонента определяет элемент суженной группы голономии $\Phi_0\Big(\sigma(x_0)\Big)$,

$$\rho^{-1}\left[\mathsf{P}\,\mathsf{exp}\left(\oint_\gamma dx^\alpha A_\alpha\right)\right] \quad \in \Phi_0\Big(\sigma(x_0)\Big), \tag{14.57}$$

где A_α – калибровочное поле, соответствующее произвольному сечению, проходящему через точку $\sigma(x_0) \in \pi^{-1}(\mathbb{U})$. Таким образом, рассматривая все замкнутые пути $\gamma \in \mathbb{M}$ с началом в точке x_0, можно определить суженную группу голономии.

Определение. След от P-экспоненты

$$W_\gamma[A] := \mathsf{tr}\left[\mathsf{P}\,\mathsf{exp}\left(\oint_\gamma dx^\alpha A_\alpha\right)\right]. \tag{14.58}$$

называется *петлей Вильсона*. □

Предложение 14.5.1. *Петля Вильсона инвариантна относительно калибровочных преобразований.*

Доказательство. Прямое следствие равенства (14.55). □

Замечание. Петля Вильсона играет важную роль в решеточной формулировке квантовых калибровочных моделей. □

Полученное выражение для упорядоченной экспоненты (14.49) не зависит от выбора координат на $\mathbb{U} \subset \mathbb{M}$. Это значит, что понятие упорядоченной экспоненты, которое было получено в одной карте, без труда переносится на произвольные пути в $\mathbb{M}$, которые в общем случае не покрываются одной картой. Для этого весь путь надо разбить на отрезки, каждый из которых покрывается одной картой, и взять сумму интегралов вдоль каждого отрезка.

14.6 Отображение связностей

В разделе 13.3 мы изучили отображение расслоений. В частности, было определено вложение (инъекция) расслоений, редукция структурной группы, а также индуцированное расслоение. Ниже мы изучим вопрос о том, как ведут себя связности и соответствующие им группы голономий при отображении расслоений. В дальнейшем эти результаты будут использованы при изучении групп голономий.

Предложение 14.6.1. *Пусть* $f:\ \mathbb{P}_1(\mathbb{M}_1,\pi_1,\mathbb{G}_1)\to\mathbb{P}_2(\mathbb{M}_2,\pi_2,\mathbb{G}_2)$ *– гомоморфизм главных расслоений, состоящий из дифференцируемого отображения пространств расслоений* $f_{\mathbb{P}}:\ \mathbb{P}_1\to\mathbb{P}_2$ *и гомоморфизма структурных групп* $f_{\mathbb{G}}:\ \mathbb{G}_1\to\mathbb{G}_2$*, такой, что индуцированное отображение баз* $f_{\mathbb{M}}:\ \mathbb{M}_1\to\mathbb{M}_2$ *есть диффеоморфизм. Пусть* Γ_1 *– связность на* $\mathbb{P}_1$ *с формой связности* ω_1 *и формой кривизны* R_1*. Тогда:*

1) *Существует единственная связность* Γ_2 *на* $\mathbb{P}_2$ *такая, что* $f_{\mathbb{P}}$ *отображает горизонтальные подпространства связности* Γ_1 *в горизонтальные подпространства связности* Γ_2*.*
2) *Если* ω_2 *и* R_2 *– формы связности и кривизны для* Γ_2*, то*

$$f^*_{\mathbb{P}}\omega_2=f_{\mathfrak{g}}\omega_1\qquad\text{и}\qquad f^*_{\mathbb{P}}R_2=f_{\mathfrak{g}}R_1,$$

где правые части $f_{\mathfrak{g}}\omega_1$ *и* $f_{\mathfrak{g}}R_1$ *обозначают* $\mathfrak{g}_2$*-значные формы на* $\mathbb{P}_1$*, определенные соотношениями:*

$$(f_{\mathfrak{g}}\omega_1)(\widetilde{X})=f_{\mathfrak{g}}\Big(\omega_1(\widetilde{X})\Big)\qquad\text{и}\qquad(f_{\mathfrak{g}}R_1)(\widetilde{X},\widetilde{Y})=f_{\mathfrak{g}}\Big(R_1(\widetilde{X},\widetilde{Y})\Big)\qquad\forall\widetilde{X},\widetilde{Y}\in\mathcal{X}(\mathbb{P}_1),$$

где $f_{\mathfrak{g}}=f_{\mathbb{G}*}$ *– гомоморфизм алгебр Ли* $\mathfrak{g}_1\to\mathfrak{g}_2$*, индуцированный отображением структурных групп* $f_{\mathbb{G}}:\ \mathbb{G}_1\to\mathbb{G}_2$ *(дифференциал отображения* $f_{\mathbb{G}}$*).*

3) *Если* $p_2=f_{\mathbb{P}}(p_1)\in\mathbb{P}_2$ *– образ точки* $p_1\in\mathbb{P}_1$*, то* $f_{\mathbb{G}}$ *гомоморфно отображает группу голономии* $\Phi(p_1)$ *в точке* p_1 *на* $\Phi(p_2)$ *и ограниченную группу голономии* $\Phi_0(p_1)$ *на* $\Phi_0(p_2)$*.*

Доказательство. Проводится путем явного построения связности Γ_2 на $\mathbb{P}_2$. См., например, [**?**]. □

Определение. В ситуации, описанной в предложении 14.6.1, говорят, что f отображает связность Γ_1 в связность Γ_2. В частном случае, если $\mathbb{P}_1(\mathbb{M}_1,\pi_1,\mathbb{G}_1)$ – редуцированное подрасслоение в $\mathbb{P}_2(\mathbb{M}_2,\pi_2,\mathbb{G}_2)$, т.е. $f_{\mathbb{G}}$ – мономорфизм, $\mathbb{M}_1=\mathbb{M}_2=\mathbb{M}$ и $f_{\mathbb{M}}=\mathsf{id}_{\mathbb{M}}$, то говорят, что связность Γ_2 на $\mathbb{P}_2$ *редуцируема* к связности Γ_1 на $\mathbb{P}_1$. Автоморфизм f главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ называется *автоморфизмом связности* Γ на $\mathbb{P}$, если он отображает Γ в Γ. В этом случае говорят, что связность Γ *инвариантна* относительно f. □

Предложение 14.6.2. *Любая связность* Γ *на главном расслоении* $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ *инвариантна относительно вертикальных автоморфизмов (пример 13.3.1).* □

Доказательство. Прямое следствие свойства 2) в определении связности. □

Следствие. Пусть $\mathbb{Q}(\mathbb{M},\pi,\mathbb{H})$ – подрасслоение в $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$, где $\mathbb{H}$ – подгруппа Ли в $\mathbb{G}$. Пусть Γ - связность на $\mathbb{P}$ с формой связности ω. Тогда связность Γ на $\mathbb{P}$ редуцируема к связности Γ' на $\mathbb{Q}$ тогда и только тогда, когда сужение формы связности ω на $\mathbb{Q}$ является $\mathfrak{h}$-значным, где $\mathfrak{h}$ – подалгебра Ли в $\mathfrak{g}$, соответствующая подгруппе Ли $\mathbb{H}\subset\mathbb{G}$. Если связность Γ редуцируема к Γ', то группы голономии Φ и суженные группы голономии Φ_0 для $\mathbb{Q}$ и $\mathbb{P}$ изоморфны. □

Доказательство. Пусть Γ' – связность на $\mathbb{Q}$. Ее форма связности, по определению, $\mathfrak{h}$-значна и по предложению 14.6.1 продолжается единственным образом до связности Γ на $\mathbb{P}$. Обратно. Если связность Γ на $\mathbb{P}$ редуцируема к связности на $\mathbb{Q}$, то сужение ω на $\mathbb{Q}$ $\mathfrak{h}$-значно. Пусть связность Γ редуцируема к Γ'. Отождествим множество точек $\mathbb{Q}$ с его образом $f_{\mathbb{P}}(\mathbb{Q})$ в $\mathbb{P}$. Тогда любая горизонтальная кривая в $\mathbb{P}$ с началом в произвольной точке $p'\in f_{\mathbb{P}}(\mathbb{Q})$ будет целиком лежать в $f_{\mathbb{P}}(\mathbb{Q})$, т.к. сужение распределения горизонтальных подпространств в $\mathbb{P}$ на $\mathbb{Q}$ совпадает со связностью на

$\mathbb{Q}$. Поскольку группы голономии для всех точек $p \in \mathbb{P}$ изоморфны, то изоморфны также все группы голономии для $\mathbb{Q}$ и $\mathbb{P}$. □

Предложение 14.6.3. *Пусть $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ – подрасслоение в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, где $\mathbb{H}$ – подгруппа Ли в $\mathbb{G}$. Допустим, что алгебра Ли $\mathfrak{g}$ для $\mathbb{G}$ допускает подпространство $\mathfrak{m}$ такое, что $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и $\mathsf{ad}\,(\mathbb{H})\mathfrak{m} = \mathfrak{m}$, где $\mathfrak{h}$ подалгебра Ли для $\mathbb{H}$. Тогда для каждой формы связности ω на $\mathbb{P}$ $\mathfrak{h}$-компонента ω' формы связности ω, суженная на $\mathbb{Q}$ является формой связности на $\mathbb{Q}$.*

Доказательство. См., например, [**?**]. □

Замечание. В силу следствия из предложения 14.6.1 форма связности ω' совпадает с ω на $\mathbb{Q}$. □

В предложении 14.6.1 мы рассматривали отображение связности Γ_1 на $\mathbb{P}_1$ в некоторую связность Γ_2 на $\mathbb{P}_2$. При определенных условиях справедливо также обратное утверждение, и связность с $\mathbb{P}_2$ можно перенести на $\mathbb{P}_1$.

Предложение 14.6.4. *Пусть $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}_1) \to \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G}_2)$ – гомоморфизм главных расслоений такой, что гомоморфизм структурных групп $f_{\mathbb{G}} : \mathbb{G}_1 \to \mathbb{G}_2$ является изоморфизмом. Пусть Γ_2 – связность на $\mathbb{P}_2$ с формой связности ω_2 и формой кривизны R_2. Тогда:*

1) *Существует единственная связность Γ_1 на $\mathbb{P}_1$ такая, что $f_{\mathbb{P}}$ отображает горизонтальные подпространства связности Γ_1 в горизонтальные подпространства связности Γ_2.*
2) *Если ω_1 и R_1 – формы связности и кривизны для Γ_1, то*

$$f^*_{\mathbb{P}}\omega_2 = f_{\mathfrak{g}}\omega_1 \quad \text{и} \quad f^*_{\mathbb{P}}R_2 = f_{\mathfrak{g}}R_1,$$

   где праве части определены в предложении 14.6.1.
3) *Если $p_2 = f_{\mathbb{P}}(p_1) \in \mathbb{P}_2$ – образ точки $p_1 \in \mathbb{P}_1$, то изоморфизм $f_{\mathbb{G}}$ гомоморфно отображает группу голономии $\Phi(p_1)$ в точке p_1 в $\Phi(p_2)$ и ограниченную группу голономии $\Phi_0(p_1)$ в $\Phi_0(p_2)$.*

Доказательство. Проводится путем явного построения связности Γ_1 на $\mathbb{P}_1$. См., например, [**?**]. □

Определение. В ситуации, описанной в предложении 14.6.4, говорят, что связность Γ_1 *индуцирована гомоморфизмом f* из связности Γ_2. □

Следствие. Пусть $f : \mathbb{P}_1(\mathbb{M}_1, \pi_1, \mathbb{G}) \to \mathbb{P}_2(\mathbb{M}_2, \pi_2, \mathbb{G})$ – отображение расслоений с одинаковой структурной группой такое, что $f_{\mathbb{G}} = \mathsf{id}\,_{\mathbb{G}}$ – тождественный автоморфизм. Если ω_2 – форма связности на $\mathbb{P}_2$, то отображение $f_{\mathbb{P}}$ индуцирует связность на $\mathbb{P}_1$: $\omega_1 = f^*_{\mathbb{P}}\omega_2$. В частности, для данного главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и отображения баз $f_{\mathbb{N}} : \mathbb{N} \to \mathbb{M}$ каждая связность на $\mathbb{P}$ индуцирует связность на $f^{-1}_{\mathbb{N}}(\mathbb{P})$. В частном случае, если $\mathbb{U} \subset \mathbb{M}$ – открытое подмножество, то связность Γ на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ индуцирует связность на индуцированном подрасслоении $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$. □

Замечание. В предложении 14.6.1 связность Γ_2 на $\mathbb{P}_2$ строилась таким образом, что гомоморфное отображение групп голономий является сюръективным. В предложении 14.6.4 утверждается, что отображение групп голономий является только гомоморфизмом. □

14.7 Связность на ассоциированном расслоении

Пусть дано главное расслоение $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и ассоциированное с ним расслоение $\mathbb{E}(\mathbb{M},\pi_{\mathbb{E}},\mathbb{F},\mathbb{G},\mathbb{P})$ с типичным слоем $\mathbb{F}$ (см. раздел 13.2). Если на $\mathbb{P}$ задана связность, то она определяет связность на $\mathbb{E}$ следующим образом.

Определение. Пусть $u\in\mathbb{E}$ – произвольная точка ассоциированного расслоения. *Вертикальным подпространством* $\mathbb{V}_u(\mathbb{E})$ в касательном пространстве $\mathbb{T}_u(\mathbb{E})$ называется касательное пространство к слою $\pi_{\mathbb{E}}^{-1}\Big(\pi_{\mathbb{E}}(u)\Big)$, которое лежит в $\mathbb{T}_u(\mathbb{E})$. □

Чтобы определить горизонтальное подпространство, вспомним, что ассоциированное расслоение строилось с помощью естественной проекции на фактор пространство

$$\mathbb{P}\times\mathbb{F}\to\mathbb{E}=\mathbb{P}\times_{\mathbb{G}}\mathbb{F}.$$

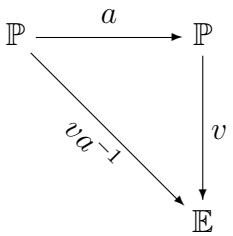
Выберем точку $(p,v)\in\mathbb{P}\times\mathbb{F}$, которая проектируется на $u\in\mathbb{E}$. Теперь зафиксируем точку типичного слоя $v\in\mathbb{F}$ и рассмотрим отображение

$$v:\ \mathbb{P}\ni p\mapsto v(p)=u\in\mathbb{E}, \tag{14.59}$$

отображающее точку $p\in\mathbb{P}$ в $v(p)\in\mathbb{E}$. То есть каждой точке типичного слоя v ставится в соответствие отображение (14.59), которое мы обозначаем той же буквой.

Замечание. В отличии от отображения $p:\ \mathbb{F}\to\mathbb{F}_x$, определенного ранее (13.15), это отображение в общем случае не является диффеоморфизмом, так как размерности главного и ассоциированного расслоения могут отличаться. Даже если размерности совпадают, $\dim\mathbb{G}=\dim\mathbb{F}$, то этого недостаточно для того, чтобы отображение (14.59) было диффеоморфизмом. Действительно, если v_0 – неподвижная точка группы преобразований, то отображение $v_0(p)$ переводит все точки слоя $\pi^{-1}(x)$ в одну фиксированную точку ассоциированного расслоения $u_0\in\pi_{\mathbb{E}}^{-1}(x)$. □

Поскольку точки (p,v) и (pa,va) из $\mathbb{P}\times\mathbb{F}$ проектируются в одну и ту же точку ассоциированного расслоения $u\in\mathbb{E}$, то $(va^{-1})(p)=v(pa)$, т.е. диаграмма



коммутативна. Используя построенное отображение, определим горизонтальные подпространства в ассоциированное расслоении.

Определение. *Горизонтальным подпространством* в точке $u\in\mathbb{E}$ называется образ $\mathbb{H}_u(\mathbb{E})=v_*\Big(\mathbb{H}_p(\mathbb{P})\Big)$, где v_* – дифференциал отображения (14.59). □

Легко видеть, что подпространство $\mathbb{H}_u(\mathbb{E})$ не зависит от выбора точек $(p,v)\in\mathbb{P}\times\mathbb{F}$, которые проектируются в точку $u\in\mathbb{E}$. Действительно, поскольку $(va)(p)=v(pa^{-1})$, то $(va)_*=v_*\circ r_{a*}^{-1}$. Поэтому для точки (pa,va), которая проектируется в ту же точку $u\in\mathbb{E}$, что и (p,v), справедливо равенство

$$\mathbb{H}_u=(va)_*\mathbb{H}_{pa}=v_*\circ r_{a*}^{-1}\circ r_{a*}\mathbb{H}_p=v_*\mathbb{H}_p,$$

где мы использовали инвариантность (14.3) распределения горизонтальных подпространств на $\mathbb{P}$. В следующем примере мы докажем, что касательное пространство к ассоциированному расслоению представляет собой прямую сумму,

$$\mathbb{T}_u(\mathbb{E}) = \mathbb{V}_u(\mathbb{E}) \oplus \mathbb{H}_u(\mathbb{E}). \tag{14.60}$$

Этого достаточно для определения связности на ассоциированном расслоении, т.к. в определение связности на расслоении общего вида **??** инвариантность распределения горизонтальных подпространств относительно групповых преобразований не входит. Тем не менее построенная связность на ассоциированном расслоении обладает дополнительным свойством, которое наследуется из главного расслоения: она инвариантна относительно действия группы справа. По построению, отображение (14.59) перестановочно с групповым действием. Поэтому перестановочны также дифференциалы этих отображений,

$$v_* \circ r_{a*} = r_{a*} \circ v_*.$$

Тогда из правой инвариантности связности на главном расслоении (14.3) следует инвариантность связности на ассоциированном расслоении

$$r_{a*}\mathbb{H}_u(\mathbb{E}) = \mathbb{H}_{ua}(\mathbb{E}). \tag{14.61}$$

Пример 14.7.1 (Локальное рассмотрение). Рассмотрим достаточно малую координатную окрестность на базе, $\mathbb{U} \subset \mathbb{M}$, с координатами x^α, $\alpha = 1, \dots, n$, такую, что выполнены условия локальной тривиализации расслоений, $\pi^{-1}(\mathbb{U}) \approx \mathbb{U} \times \mathbb{G}$ и $\pi_{\mathbb{E}}^{-1}(\mathbb{U}) \approx \mathbb{U} \times \mathbb{F}$. Отождествим $\pi^{-1}(\mathbb{U})$ с $\mathbb{U} \times \mathbb{G}$ и $\pi_{\mathbb{E}}^{-1}(\mathbb{U})$ с $\mathbb{U} \times \mathbb{F}$. Ограничим наше рассмотрение окрестностью единицы группы Ли, $\mathbb{U}_\mathbb{G} \subset \mathbb{G}$, где определены координаты a^{A}, $\mathrm{A} = 1, \dots, \mathrm{N}$, и функция композиции (см. раздел 9.1). Выберем также некоторую координатную окрестность в типичном слое, $\mathbb{U}_\mathbb{F} \subset \mathbb{F}$, где определены координаты v^i, $i = 1, \dots, \dim \mathbb{F}$, и задано отображение $v \mapsto va$ в координатной форме. Тогда точки расслоений будут иметь координаты $p = \{x^\alpha, a^{\mathrm{A}}\} \in \mathbb{P}$ и $u = \{x^\alpha, v^i\} \in \mathbb{E}$. По определению, векторы ∂_i касательны к слою $\pi_{\mathbb{E}}^{-1}(x)$ и, следовательно, образуют базис вертикальных подпространств $\mathbb{V}_u(\mathbb{E})$ для всех $u \in \mathbb{E}$. Пусть $v_0 = \{v_0^i\} \in \mathbb{U}_\mathbb{F}$ – фиксированная точка типичного слоя. Тогда соответствующее этой точке отображение (14.59) в координатах имеет вид

$$v_0 : \quad \mathbb{U} \times \mathbb{U}_\mathbb{G} \ni \{x^\alpha, a^{\mathrm{A}}\} \mapsto \{x^\alpha, v^i(v_0, a)\} \in \mathbb{U} \times \mathbb{U}_\mathbb{F},$$

где $v^i(v_0, a)$ – некоторая функция координат v_0^i и a^{A}. Пусть $D_\alpha = \partial_\alpha - \omega_\alpha{}^{\mathrm{A}} L^*_{\mathrm{A}}$, где $\omega_\alpha{}^{\mathrm{A}}$ есть компоненты формы связности (14.4), – базис горизонтальных подпространств $\mathbb{H}_p(\mathbb{P})$ в главном расслоении, который был построен ранее (14.38). Тогда он отображается в касательное пространство к ассоциированному расслоению,

$$v_{0*} : \quad \mathbb{H}_p(\mathbb{P}) \ni D_\alpha \mapsto \overline{D}_\alpha = \partial_\alpha - \omega_\alpha{}^{\mathrm{A}} L^*_{\mathrm{A}} v^i \partial_i \in \mathbb{H}_u(\mathbb{E}),$$

где L^*_{A} действует на $v^i(v_0, a)$ как дифференцирование по a. Конкретный вид функций $v^i(v_0, a)$ зависит от типичного слоя и действия на нем структурной группы. Независимо от вида функций $v^i(v_0, a)$ векторы $\overline{D}_\alpha$ линейно независимы, т.к. содержат ∂_α, и поэтому образуют базис горизонтального подпространства $\mathbb{H}_u(\mathbb{E})$. Таким образом совокупность векторов $\{\overline{D}_\alpha, \partial_i\}$ образует базис касательного пространства $\mathbb{T}_u(\mathbb{E})$, который соответствует разложению (14.60). Аналогичное построение можно выполнить в окрестности произвольной точки ассоциированного расслоения $u \in \mathbb{E}$. Следовательно, разложение касательного пространства $\mathbb{T}_u(\mathbb{E})$ в прямую сумму (14.60) имеет место в общем случае.

Если типичный слой – это векторное пространство $\mathbb{V}$, на котором задано представление ρ (см., раздел 9.8), то

$$v^i(v_0, a) = v_0^j S^{-1}{}_j^i(a),$$

где $S_j{}^i(a)$ – матрица представления элемента $a \in \mathbb{G}$. В этом случае

$$\omega_\alpha{}^{\text{A}} L^*_{\text{A}} v^i = \omega_\alpha{}^{\text{A}} v_0^j S_j{}^k L_{\text{A}k}{}^i = -v^j \omega_{\alpha j}{}^i,$$

где введено обозначение $\omega_{\alpha j}{}^i = -\omega_\alpha{}^{\text{A}} L_{\text{A}j}{}^i$ и мы воспользовались правилом дифференцирования матриц представления (9.63). Тогда базис горизонтальных векторных полей на ассоциированном расслоении имеет вид

$$\overline{D}_\alpha = \partial_\alpha + v^j \omega_{\alpha j}{}^i \partial_i. \tag{14.62}$$

Этот базис инвариантен относительно действия структурной группы $\mathbb{G}$ на ассоциированном расслоении $\mathbb{E}$ справа. Действительно, действие элемента $b \in \mathbb{G}$ на базисный вектор (14.62) имеет вид

$$r_{b*} \overline{D}_\alpha = \partial_\alpha + v^i \omega_{\alpha i}{}^j(u) S^{-1}{}_j^k(b) \partial_k,$$

т.к. $v^i(v_0, ab) = v^j(v_0, a) S^{-1}{}_j^i(b)$. Поскольку $v^i(ub) = v^j(u) S^{-1}{}_j^i(b)$ и $\omega_\alpha{}^{\text{A}}(pb) = \omega_\alpha{}^{\text{B}} S^{-1}{}_{\text{B}}^{\text{A}}(b)$, то

$$r_{b*} \overline{D}_\alpha|_u = \overline{D}_\alpha|_{ub},$$

где мы воспользовались инвариантностью (9.64) матриц представления генераторов структурной группы. Это соответствует инвариантности распределения горизонтальных подпространств на ассоциированном расслоении (14.61). □

Определения горизонтального лифта и параллельного переноса для ассоциированных расслоений дословно повторяют определения, данные для главных расслоений.

Определение. Кривая γ^{-1} в ассоциированном расслоении $\mathbb{E}$ называется *горизонтальной*, если касательный к ней вектор горизонтален в каждой точке. Если задана кривая γ в базе $\mathbb{M}$, то *горизонтальным лифтом* или просто *лифтом* этой кривой называется такая горизонтальная кривая γ^{-1} в $\mathbb{E}$, что $\pi_{\mathbb{E}}(\gamma^{-1}) = \gamma$. □

Предложение 14.7.1. *Пусть $\gamma = x(t)$, $t \in [0,1]$, – кусочно дифференцируемая кривая класса $\mathcal{C}^1$ в $\mathbb{M}$ с началом в точке $x_0 \in \mathbb{M}$. Тогда для произвольной точки слоя $u_0 \in \pi_{\mathbb{E}}^{-1}(x_0)$ существует единственный горизонтальный лифт $\gamma^{-1} = u(t)$ кривой γ с началом в точке u_0.*

Доказательство. Сначала докажем существование горизонтального лифта. Выберем точку $(p_0, v_0) \in \mathbb{P} \times \mathbb{F}$ такую, что $p_0(v_0) = u_0$, где отображение p определено формулой (13.15). Согласно предложению 14.3.1 существует единственный горизонтальный лифт $\tilde{\gamma}$ кривой γ в главное расслоение $\mathbb{P}$ с началом в точке $p_0 \in \mathbb{P}$. Тогда кривая $\gamma^{-1} = \tilde{\gamma}(v_0)$ является горизонтальным лифтом кривой в базе γ в ассоциированное расслоение $\mathbb{E}$. Действительно, касательный вектор к кривой γ^{-1} лежит в $\mathbb{H}_\gamma(\mathbb{E})$, что сразу следует из определения дифференциала отображения (3.71). Единственность горизонтального лифта следует из единственности решения системы линейных дифференциальных с заданными начальными условиями. □

Для формулировки следующего утверждения нам понадобится естественное

Определение. Локальное сечение ассоциированного расслоения $\sigma : \mathbb{U} \to \mathbb{E}$, определенное на открытом подмножестве $\mathbb{U} \subset \mathbb{M}$, называется *параллельным или горизонтальным*, если образ $\sigma_* \mathbb{T}_x(\mathbb{M})$, где σ_* – дифференциал сечения, горизонтален при всех $x \in \mathbb{U}$, т.е. для любой кривой γ в $\mathbb{U}$, соединяющей точки x_0 и x_1, точка слоя $\sigma(x_0)$ при параллельном переносе слоя вдоль кривой γ переходит в точку $\sigma(x_1)$. □

В предложении 13.2.3 мы отождествили ассоциированное расслоение $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ с фактор пространством $\mathbb{P}/\mathbb{H}$. Затем в теореме 13.3.2 привели критерий редуцируемости структурной группы $\mathbb{G}$ главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ к подгруппе $\mathbb{H}$, который заключается в существовании глобального сечения σ ассоциированного расслоения $\mathbb{E}$. Кроме того, была установлена естественная взаимно однозначная связь между сечениями σ ассоциированного расслоения $\mathbb{E}$ и редуцированными главными расслоениями $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$. В примере 13.3.2 эта теорема была использована для доказательства существования римановой метрики на произвольном многообразии. Возникает вопрос о том, в каком случае связность, заданная на главном расслоении $\mathbb{P}$ редуцируема к связности на редуцированном расслоении $\mathbb{Q}$? Ответ дает следующее утверждение.

Предложение 14.7.2. *Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение и $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{G}/\mathbb{H}, \mathbb{G}, \mathbb{P})$ ассоциированное расслоение со стандартным слоем $\mathbb{G}/\mathbb{H}$, где $\mathbb{H}$ – замкнутая подгруппа в $\mathbb{G}$. Пусть $\sigma : \mathbb{M} \to \mathbb{E}$ – глобальное сечение ассоциированного расслоения и $\mathbb{Q}(\mathbb{M}, \pi, \mathbb{H})$ редуцированное подрасслоение в $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, соответствующее сечению σ. Связность Γ на $\mathbb{P}$ редуцируема к связности Γ' на $\mathbb{Q}$ тогда и только тогда, когда сечение σ параллельно относительно Γ.*

Доказательство. См., например, [**?**]. □

14.8 Свойства групп голономий

Продолжим изучение свойств групп голономий, которое было начато в разделе 14.4.

Теорема 14.8.1 (Теорема редукции). *Пусть $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ – главное расслоение со связностью Γ и p – произвольная точка в $\mathbb{P}$. Обозначим через $\mathbb{P}(p)$ множество точек в $\mathbb{P}$, которые можно соединить с точкой p горизонтальными кусочно дифференцируемыми кривыми. Тогда:*

1) $\mathbb{P}(p)$ – редуцированное главное расслоение со структурной группой $\Phi(p)$;

2) Связность Γ редуцируема к связности на $\mathbb{P}(p)$.

Доказательство. См., например, [**?**]. □

Эта теорема оправдывает следующее

Определение. Главное расслоение $\mathbb{P}(p)$, с базой $\mathbb{M}$, проекцией π и структурной группой $\Phi(p)$, состоящее из множества точек в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, которые можно соединить с точкой p горизонтальными кусочно дифференцируемыми кривыми, называется *расслоением голономии* через p. □

Очевидно, что $\mathbb{P}(p) = \mathbb{P}(q)$ тогда и только тогда, когда точки p и q можно соединить горизонтальной кривой. В разделе 14.4 было введено отношение эквивалентности: $p \sim q$, если p и q можно соединить горизонтальной кривой. Поэтому для каждой

пары точек из главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ либо $\mathbb{P}(p)=\mathbb{P}(q)$, либо $\mathbb{P}(p)\cap\mathbb{P}(q)=\emptyset$. Другими словами, главное расслоение $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ разлагается в объединение,

$$\mathbb{P}(\mathbb{M},\pi,\mathbb{G})=\bigcup_{p\in\mathbb{P}}\mathbb{P}(p),$$

попарно непересекающихся расслоений голономии. Так как каждый элемент $a\in\mathbb{G}$ отображает каждую горизонтальную кривую в горизонтальную, то $r_a\mathbb{P}(p)=\mathbb{P}(pa)$ и отображение

$$r_a:\quad \mathbb{P}(p)\to\mathbb{P}(pa)$$

индуцирует изоморфизм расслоений $f=(f_{\mathbb{P}},f_{\mathbb{G}})$, где $f_{\mathbb{P}}=r_a$, с соответствующим изоморфизмом структурных групп:

$$f_{\mathbb{G}}=\mathsf{ad}\,(a^{-1}):\quad \Phi(p)\ni b\mapsto a^{-1}ba\in\Phi(pa),$$

т.к. группы голономии в различных точках сопряжены друг другу (предложение 14.4.1). Легко видеть, что для двух произвольных точек $p,q\in\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ существует такой элемент $a\in\mathbb{G}$, что $\mathbb{P}(p)=\mathbb{P}(qa)$. Поэтому все расслоения голономий $\mathbb{P}(p)$ изоморфны друг другу.

Теорема 14.8.2 (Амброз–Зингер). *Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение со связной базой $\mathbb{M}$. Пусть Γ – связность на $\mathbb{P}$ с формой кривизны R, $\Phi(p_0)$ – группа голономии в точке $p_0\in\mathbb{P}$ и $\mathbb{P}(p_0)$ – расслоение голономии через p_0. Тогда алгебра Ли группы голономии $\Phi(p_0)$ совпадает с подпространством в алгебре Ли $\mathfrak{g}$ структурной группы $\mathbb{G}$, порожденной всеми элементами вида $R_p(\widetilde{X},\widetilde{Y})$ для всех $p\in\mathbb{P}(p_0)$ и всех горизонтальных векторных полей $\widetilde{X},\widetilde{Y}$ в точке p.*

Доказательство. Используется паракомпактность $\mathbb{M}$ [**?**]. □

В компонентах ???

Теорема 14.8.3. *Пусть $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – главное расслоение со связным пространством расслоения $\mathbb{P}$. Если $\mathsf{dim}\,\mathbb{M}\geq 2$, то существует связность на $\mathbb{P}$ такая, что расслоения голономии $\mathbb{P}(p)$ для всех $p\in\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ совпадают с главным расслоением $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$.*

Доказательство. Явное построение связности. При этом используется паракомпактность $\mathbb{M}$ [**?**]. Для линейных связностей это утверждение было доказано в [**?**]. В общем случае доказательство дано в [**?**]. □

Следствие. Любая связная группа Ли $\mathbb{G}$ может быть реализована как группа голономии некоторой связности в тривиальном главном расслоении $\mathbb{P}=\mathbb{M}\times\mathbb{G}$, где $\mathbb{M}$ – произвольное дифференцируемое многообразие размерности $\mathsf{dim}\,\mathbb{M}\geq 2$. □

Доказательство. Выберем связную окрестность $\mathbb{U}\subset\mathbb{M}$. Тогда расслоение $\mathbb{U}\times\mathbb{G}$ связно и мы попадаем в зону деятельности теоремы 14.8.3. Связность с $\mathbb{U}\times\mathbb{G}$ продолжается на связность на $\mathbb{M}\times\mathbb{G}$ согласно предложению 14.6.1. □

14.9 Плоские связности

Рассмотрим тривиальное главное расслоение $\mathbb{P} = \mathbb{M} \times \mathbb{G}$. Для каждого элемента структурной группы $a \in \mathbb{G}$ множество $\mathbb{M}\times\{a\}$ есть подмногообразие в $\mathbb{P}$. В частности, $\mathbb{M} \times \{e\}$, где e – единица группы, есть редуцированное подрасслоение в $\mathbb{P}$.

Определение. *Канонической плоской связностью* на тривиальном главном расслоении $\mathbb{P} = \mathbb{M} \times \mathbb{G}$ называется распределение горизонтальных подпространств $\mathbb{H}_p(\mathbb{P})$, образованное касательными пространствами к $\mathbb{M} \times \{a\}$ для всех точек $p = (x, a) \in \mathbb{M} \times \mathbb{G}$. □

Предложение 14.9.1. *Связность* Γ *на тривиальном главном расслоении* $\mathbb{P} = \mathbb{M}\times\mathbb{G}$ *является канонической плоской тогда и только тогда, когда она редуцируема к единственной связности на* $\mathbb{M} \times \{e\}$.

Доказательство. Выберем нулевое сечение $\sigma_0 : \ x \mapsto (x, e)$. Это сечение является главным расслоением $\mathbb{P}(\mathbb{M}, \pi, e)$, на котором существует единственная связность. Эта связность взаимно однозначно определяет каноническую плоскую связность на $\mathbb{P} = \mathbb{M} \times \mathbb{G}$. □

Пусть θ – каноническая форма на группе Ли $\mathbb{G}$, определенная в разделе 9.2. Обозначим через $\mathsf{pr} : \ \mathbb{M} \times \mathbb{G} \to \mathbb{G}$ естественную проекцию и положим

$$\omega := \mathsf{pr}^*\theta = \omega^{*\mathrm{A}} L_{\mathrm{A}}. \tag{14.63}$$

Эта 1-форма является частным случаем формы связности (14.5) и определяет каноническую плоскую связность на $\mathbb{P}$. Формула Маурера–Картана для канонической 1-формы (9.26) влечет, что каноническая плоская связность имеет нулевую кривизну, так как

$$d\omega = d(\,\mathsf{pr}^*\theta) = \,\mathsf{pr}^*(d\theta) = \,\mathsf{pr}^*\left(-\frac{1}{2}[\theta,\theta]\right) = -\frac{1}{2}[\,\mathsf{pr}^*\theta,\,\mathsf{pr}^*\theta] = -\frac{1}{2}[\omega,\omega].$$

Сравнивая полученное равенство со структурным уравнением (14.26), заключаем, что форма кривизны канонической плоской связности тождественно равна нулю, $R = 0$.

Теперь рассмотрим случай произвольного главного расслоения. Дадим общее

Определение. Связность Γ на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ называется *плоской*, если каждая точка базы $x \in \mathbb{M}$ имеет окрестность $\mathbb{U}$ такую, что индуцированная связность на $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$ изоморфна канонической плоской связности на $\mathbb{U} \times \mathbb{G}$. Другими словами, существует изоморфизм $\chi : \ \pi^{-1}(\mathbb{U}) \to \mathbb{U} \times \mathbb{G}$, отображающий горизонтальное подпространство в каждой точке $p \in \pi^{-1}(\mathbb{U})$ в горизонтальное подпространство канонической плоской связности на $\mathbb{U} \times \mathbb{G}$ в точке $\chi(p)$. □

Предложение 14.9.2. *Пусть задано главное расслоение* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. *Тогда плоская связность на* $\mathbb{P}$ *существует и единственна с точностью до изоморфизма.*

Доказательство. Если задано главное расслоение, то определен атлас на базе $\mathbb{M} = \bigcup_i \mathbb{U}_i$ и семейство функций перехода (13.7), которые по теореме 13.1.2 с точностью до изоморфизма определяют главное расслоение. Выбрав координатное покрытие базы достаточно малым, можно считать, что все координатные окрестности $\mathbb{U}_i$ соответствуют окрестностям, входящим в определение плоской связности. Согласно теореме 14.1.1 для однозначного задания связности на $\mathbb{P}$ достаточно задать семейство

локальных форм связности на каком либо атласе базы. Это означает, что плоская связность на произвольном главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ существует и единственна с точностью до изоморфизма. □

Теорема 14.9.1. *Связность* Γ *на главном расслоении* $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ *является плоской тогда и только тогда, когда ее форма кривизны равна нулю,* $R=0$.

Доказательство. Необходимость очевидна. Обратно. Допустим, что форма кривизны равна нулю. Пусть $\mathbb{U}$ – односвязная окрестность точки $x\in\mathbb{P}$ и рассмотрим индуцированную связность на $\mathbb{P}|_{\mathbb{U}}=\pi^{-1}(\mathbb{U})$. По теоремам 14.4.1 и Амброза–Зингера 14.8.2 группа голономии индуцированной связности на $\mathbb{P}|_{\mathbb{U}}$ состоит только из единицы. Применяя теорему редукции 14.8.1, мы видим, что индуцированная связность на $\mathbb{P}|_{\mathbb{U}}$ изоморфна канонической плоской связности на $\mathbb{U}\times\mathbb{G}$. □

Следствие. Любая связность Γ на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ с одномерной базой $\mathbb{M}$ является плоской. □

Доказательство. Любая 2-форма на одномерном многообразии равна нулю. Отсюда следует, что все локальные формы кривизны тоже равны нулю. Так как для формы кривизны только горизонтальные компоненты являются нетривиальными, то она также обращается в нуль. □

Следствие. Пусть Γ – связность на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ такая, что ее форма кривизны равна нулю, $R=0$. Если база $\mathbb{M}$ односвязна, то главное расслоение $\mathbb{P}$ изоморфно тривиальному расслоению $\mathbb{M}\times\mathbb{G}$ и связность Γ изоморфна канонической плоской связности на $\mathbb{M}\times\mathbb{G}$. □

Доказательство. Группа голономии в рассматриваемом случае состоит из единственного элемента – единицы. Поэтому расслоение голономии $\mathbb{P}(p)$ пересекает каждый слой ровно в одной точке. Следовательно, каждое расслоение голономии задает глобальное сечение, и поэтому главное расслоение тривиально. При этом горизонтальные подпространства касательны к расслоению голономии. Пусть $q\in\mathbb{P}(p)$ – произвольная точка расслоения голономии через p и $\sigma_0=(x,e)$ – нулевое сечение главного расслоения $\mathbb{P}=\mathbb{M}\times\mathbb{G}$. Тогда для каждой точки базы $x\in\mathbb{M}$ существует единственный элемент $a(x)\in\mathbb{G}$ такой, что $\sigma_0=qa$, где $x=\pi(q)$. При этом вертикальный автоморфизм $p\mapsto pa$ переводит связность Γ на $\mathbb{P}=\mathbb{M}\times\mathbb{G}$ в каноническую плоскую связность. □

Если форма кривизны равна нулю, то распределение горизонтальных векторных полей находится в инволюции. Это сразу следует из (14.39), т.к. векторы D_α образуют базис распределения горизонтальных векторных полей. Согласно теореме Фробениуса для плоской связности через каждую точку p главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ проходит интегральное подмногообразие. Все касательные векторы к интегральным подмногообразиям горизонтальны и интегральное подмногообразие, проходящее через точку p, совпадает с расслоением голономии $\mathbb{P}(p)$ через p.

Пример 14.9.1 (Локальное рассмотрение). Пусть $\mathbb{Q}=\mathbb{U}\times\mathbb{G}$ – тривиальное главное расслоение, база $\mathbb{U}$ которого покрыта одной картой. Общий вид связности на $\mathbb{Q}$ был найден ранее (14.5). Сравнение этого выражения с (14.63) показывает, что связность $\mathbb{Q}$ является канонической плоской тогда и только тогда, когда часть ее компонент, описывающих произвол в выборе связности на нулевом сечении, обращаются в нуль, $\overset{\circ}{A}{}_\alpha{}^{\mathrm{A}}(x)=0$. Из равенства (14.12) следует, что компоненты локальной формы

канонической плоской связности для произвольного сечения $\sigma:\ x\mapsto\Big(x,b(x)\Big)$ имеют вид

$$A_\alpha{}^{\text{A}}(x)=\partial_\alpha b^{\text{B}}L^{-1}{}_{\text{B}}{}^{\text{A}}(b).$$

После перехода к какому либо представлению структурной группы

$$\rho:\ \ \mathbb{G}\ni a\mapsto\{S_i{}^j(a)\}\in\mathsf{aut}\,\mathbb{V},$$

для компонент локальной формы плоской связности справедливо равенство (14.14)

$$A_\alpha=\partial_\alpha SS^{-1},\tag{14.64}$$

где мы, для простоты, опустили матричные индексы.

Определение. Калибровочное поле A_α, заданное равенством (14.64) на координатной окрестности U, называется *чистой калибровкой.* □

Теперь рассмотрим случай, когда база $\mathbb{M}$ главного расслоения не является односвязной. Пусть Γ – связность на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ со связной базой $\mathbb{M}$. Выберем произвольную точку $p_0\in\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ и обозначим через $\widetilde{\mathbb{M}}=\mathbb{P}(p_0)$ расслоение голономии через p_0. Тогда $\widetilde{\mathbb{M}}\Big(\mathbb{M},\pi,\Phi(p_0)\Big)$ есть главное расслоение над $\mathbb{M}$ со структурной группой $\Phi(p_0)$. Так как суженная группа голономии $\Phi_0(p_0)$ для плоской связности всегда тривиальна, то по теоремам 14.4.1 и Амброза–Зингера 14.8.2 группа голономии $\Phi(p_0)$ при неодносвязной базе дискретна. Поэтому отображение $\pi:\ \widetilde{\mathbb{M}}\to\mathbb{M}$ является накрытием со связным накрывающим пространством.

Пусть $x_0=\pi(p_0)\in\mathbb{M}$. Каждая замкнутая кривая в базе $\mathbb{M}$, исходящая из x_0, при помощи параллельного переноса слоев вдоль нее определяет некоторый элемент группы голономии $\Phi(p_0)$. Поскольку суженная группа голономии тривиальна, то любые две замкнутые и гомотопные относительно начала кривые, представляющие один и тот же элемент фундаментальной группы $\pi(\mathbb{M},x_0)$, порождают один и тот же элемент из $\Phi(p_0)$. Таким образом мы получаем сюрьективное отображение фундаментальной группы $\pi(\mathbb{M},p_0)$ на группу голономии $\Phi(p_0)$. Легко видеть, что это отображение является гомоморфизмом групп. Пусть $\mathbb{H}$ – нормальная подгруппа в $\Phi(p_0)$ и положим $\mathbb{M}'=\widetilde{\mathbb{M}}/\mathbb{H}$. Тогда $\mathbb{M}'$ – главное расслоение над $\mathbb{M}$ со структурной группой $\Phi(p_0)/\mathbb{H}$. В частности, отображение $\pi:\ \mathbb{M}'\to\mathbb{M}$ – накрытие. Пусть $\mathbb{P}'(\mathbb{M}',\pi,\mathbb{G})$ – главное расслоение, индуцированное из $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ накрывающей проекцией $f_{\mathbb{M}}=\pi:\ \mathbb{M}'\to\mathbb{M}$. Пусть $f:\ \mathbb{P}'(\mathbb{M}',\pi,\mathbb{G})\to\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ – естественный гомоморфизм главных расслоений, см. теорему 13.3.4, тогда справедливо

Предложение 14.9.3. *Существует единственная связность Γ' на главном расслоении $\mathbb{P}'(\mathbb{M}',\pi,\mathbb{G})$, которая отображается на связность Γ на $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ гомоморфизмом $f:\ \mathbb{P}'\to\mathbb{P}$. Связность Γ' плоская. Если точка $p'_0\in\mathbb{P}'$ такова, что $f_{\mathbb{M}}(p'_0)=p_0$, то группа голономии $\Phi(p'_0)$ для связности Γ' изоморфно отображается на $\mathbb{H}$ гомоморфизмом $f_{\mathbb{G}}$.*

Доказательство. См., например, [**?**]. □

В данном утверждении всегда можно выбрать в качестве нормальной подгруппы единицу группы голономии $e\in\Phi(p_0)$.

Следствие. Пусть $\widetilde{\mathbb{P}}(\widetilde{\mathbb{M}},\pi,\mathbb{G})$ – главное расслоение, индуцированное из $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ накрывающей проекцией $f_{\mathbb{M}}=\pi:\ \widetilde{\mathbb{M}}\to\mathbb{M}$. Тогда существует единственная связность $\tilde{\Gamma}$ на главном расслоении $\widetilde{\mathbb{P}}(\widetilde{\mathbb{M}},\pi,\mathbb{G})$, которая отображается на связность Γ на $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ гомоморфизмом $f:\ \widetilde{\mathbb{P}}\to\mathbb{P}$. Связность $\tilde{\Gamma}$ плоская, и соответствующая ей группа голономии тривиальна.

14.10 Локальные и инфинитезимальные группы голономии

Группа голономии $\Phi(p)$ в точке $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, которая является важнейшей глобальной характеристикой связности, была определена при помощи множества всех замкнутых путей в базе $\Omega\Big(\mathbb{M}, \pi(p)\Big)$. Это крайне неудобно для практических вычислений, так как зачастую структура многообразия $\mathbb{M}$ просто неизвестна. Возникает вопрос, можно ли каким либо образом вычислить группу голономии, исходя из локальных характеристик связности? В некоторых случаях суженная группа голономии $\Phi_0(p)$ действительно определяется локальными свойствами связности. В настоящем разделе мы опишем два подхода к этой проблеме, которые основаны на понятиях локальной и инфинитезимальной групп голономии.

Начнем рассмотрение с локальной группы голономии. Пусть на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ со связной базой $\mathbb{M}$ задана связность Γ. Для каждого связного и односвязного открытого подмножества базы $\mathbb{U} \subset \mathbb{M}$ обозначим через $\Gamma_{\mathbb{U}}$ связность на $\mathbb{P}|_{\mathbb{U}} = \pi^{-1}(\mathbb{U})$, которая индуцирована из связности Γ. В силу следствия из предложения 14.6.4 связность $\Gamma_{\mathbb{U}}$ существует и единственна – это сужение связности Γ на $\pi^{-1}(\mathbb{U})$. Для каждой точки $p \in \pi^{-1}(\mathbb{U})$ обозначим через $\Phi_0(p, \mathbb{U})$ и $\mathbb{P}(p, \mathbb{U})$ суженную группу голономии с опорной точкой p и расслоение голономии через точку p для связности $\Gamma_{\mathbb{U}}$ соответственно. Расслоение голономии $\mathbb{P}(p, \mathbb{U})$ состоит из тех точек $q \in \pi^{-1}(\mathbb{U})$, которые можно соединить с точкой p горизонтальной кривой, целиком лежащей в $\pi^{-1}(\mathbb{U})$.

Рассмотрим две окрестности $\mathbb{U}_2 \subset \mathbb{U}_1$, которые содержат точку $x = \pi(p)$. Тогда всякая замкнутая петля, целиком лежащая в $\mathbb{U}_2$, будет также петлей в $\mathbb{U}_1$. Поэтому справедливо включение

$$\Phi_0(p, \mathbb{U}_2) \subset \Phi_0(p, \mathbb{U}_1).$$

Как подгруппа суженной группы голономии $\Phi_0(p)$ группа $\Phi_0(p, \mathbb{U}_1)$ однозначно определяется своей алгеброй Ли. Поэтому из равенства размерностей $\mathsf{dim}\, \Phi_0(p, \mathbb{U}_2) = \mathsf{dim}\, \Phi(p, \mathbb{U}_1)$ следует совпадение групп голономии $\Phi_0(p, \mathbb{U}_2) = \Phi_0(p, \mathbb{U}_1)$. Это наблюдение приводит к следующему понятию.

Определение. *Локальной группой голономии* $\Phi_{\mathrm{loc}}(p)$ в точке p называется пересечение

$$\Phi_{\mathrm{loc}}(p) = \bigcap_{\mathbb{U}} \Phi_0(p, \mathbb{U}) \tag{14.65}$$

по всем связным и односвязным открытым окрестностям $\mathbb{U}$ точки $x = \pi(p)$. □

Замечание. В настоящем разделе мы будем рассматривать только связные и односвязные открытые окрестности точек. Поэтому в дальнейшем, для краткости, мы будем говорить просто окрестности. □

Пусть $\{\mathbb{U}_k\}$, $k = 1, 2, \ldots,$ – последовательность окрестностей, сходящихся к точке x, т.е. $\mathbb{U}_k \supset \overline{\mathbb{U}}_{k+1}$ и $\bigcap_{k=1}^{\infty} \mathbb{U}_k = \{x\}$. Тогда, очевидно, имеют место включения

$$\Phi_0(p, \mathbb{U}_1) \supset \Phi_0(p, \mathbb{U}_2) \supset \Phi_0(p, \mathbb{U}_3) \supset \ldots$$

Поскольку для каждой окрестности $\mathbb{U}$ точки x существует целое число $k_{\mathbb{U}}$ такое, что $\mathbb{U}_k \subset \mathbb{U}$ для всех $k > k_{\mathbb{U}}$, то локальная группа голономии представима в виде

$$\Phi_{\mathrm{loc}}(p) = \bigcap_{k=1}^{\infty} \Phi_0(p, \mathbb{U}_k).$$

Так как каждая суженная группа голономии $\Phi_0(p, \mathbb{U}_k)$ есть связная подгруппа Ли в $\mathbb{G}$ (теорема 14.4.1), то отсюда следует, что размерность суженной группы голономии $\mathsf{dim}\,\Phi_0(p, \mathbb{U}_k)$ постоянна для достаточно больших k. Поэтому для больших значений k справедливо равенство $\Phi_{\mathrm{loc}}(p) = \Phi_0(p, \mathbb{U}_k)$.

Замечание. Конечно, суженная группа голономии $\Phi_0(p, x)$ для "окрестности", состоящей из одной точки x, состоит ровно из одного элемента – единицы, и ее размерность равна нулю. Допустим, что параметр k в последовательности $\{\mathbb{U}_k\}$ непрерывен. Это может быть, например, радиус шара в $\mathbb{R}^n$, на который можно отобразить окрестность $\mathbb{U}_k$. Тогда функция $\mathsf{dim}\,\Phi_0(p, \mathbb{U}_k)$ от k принимает значения в целых числах и не может быть непрерывной, если суженная группа голономии $\Phi_0(p)$ для всего главного расслоения нетривиальна. В дальнейшем мы увидим, что при определенных условиях локальная группа голономии $\Phi_{\mathrm{loc}}(p)$ совпадает с суженной группой $\Phi_0(p)$. □

Предложение 14.10.1. *Локальные группы голономии имеют следующие свойства:*

1) *локальная группа голономии $\Phi_{\mathrm{loc}}(p)$ есть связная подгруппа Ли в $\mathbb{G}$, содержащаяся в суженной группе голономии $\Phi_0(p)$;*
2) *каждая точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что $\Phi_{\mathrm{loc}}(p) = \Phi_0(p, \mathbb{V})$ для любой окрестности $\mathbb{V} \ni x$, содержащейся в $\mathbb{U}$;*
3) *если $\mathbb{U}$ – окрестность точки $x = \pi(p)$, о которой говорится в свойстве 2), то $\Phi_{\mathrm{loc}}(p) \supset \Phi_{\mathrm{loc}}(q)$ для всех точек $q \in \mathbb{P}(p, \mathbb{U})$;*
4) *для каждого $a \in \mathbb{G}$ имеем*

$$\Phi_{\mathrm{loc}}(pa) = \mathsf{ad}\,(a^{-1})\Phi_{\mathrm{loc}}(p);$$

5) *для каждого целого m множество точек базы*

$$\{\pi(p) = x \in \mathbb{M} : \ \mathsf{dim}\,\Phi_{\mathrm{loc}}(p) \leq m\}$$

открыто в $\mathbb{M}$.

Доказательство. Свойства 1)–4) очевидны. Докажем свойство 5). Из свойства 4) следует, что функция $\mathsf{dim}\,\Phi_{\mathrm{loc}}(p)$ постоянна на каждом слое и поэтому ее можно рассматривать как функцию на базе $\mathbb{M}$, принимающую целые значения. Из свойств 3) и 4) вытекает, что

$$\mathsf{dim}\,\Phi_{\mathsf{loc}}(q) \leq \mathsf{dim}\,\Phi_{\mathsf{loc}}(p),$$

для всех точек $q \in \pi^{-1}(\mathbb{U})$. Отсюда вытекает свойство 5). □

Теорема 14.10.1. *Пусть $\mathfrak{g}_0(p)$ и $\mathfrak{g}_{\mathrm{loc}}(p)$ – алгебры Ли для групп голономий $\Phi_0(p)$ и $\Phi_{\mathrm{loc}}(p)$ соответственно. Тогда $\Phi_0(p)$ и $\mathfrak{g}_0(p)$ порождаются соответственно всеми $\Phi_{\mathrm{loc}}(q)$ и $\mathfrak{g}_{\mathrm{loc}}(q)$ для всех точек $q \in \mathbb{P}(p)$.*

Доказательство. См., например, [**?**]. □

Теорема 14.10.2. *Если $\mathsf{dim}\,\Phi_{\mathrm{loc}}(p)$ постоянна на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, то локальная и суженная группы голономии совпадают, $\Phi_{\mathrm{loc}}(p) = \Phi_0(p)$, для всех $p \in \mathbb{P}$.*

Доказательство. По свойству 3) предложения 14.10.1 каждая точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что $\Phi_{\mathrm{loc}}(p) \supset \Phi_{\mathrm{loc}}(q)$ для каждой точки q из расслоения голономии $\mathbb{P}(p, \mathbb{U})$. Так как $\mathsf{dim}\,\Phi_{\mathsf{loc}}(p) = \mathsf{dim}\,\Phi_{\mathsf{loc}}(q)$, то сами группы голономии в различных точках совпадают, $\Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{loc}}(q)$. Отсюда следует, что если $q \in \mathbb{P}(p)$, то $\Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{loc}}(q)$ для всех $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. Из теоремы 14.10.1 вытекает равенство $\Phi_0(p) = \Phi_{\mathsf{loc}}(p)$. □

Теперь перейдем к определению инфинитезимальной группы голономии и изучим ее связь с локальной группой голономии. Инфинитезимальная группа голономии может быть определена только для гладких $\mathcal{C}^\infty$ главных расслоений $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ с гладкой $\mathcal{C}^\infty$ связностью Γ. В дальнейшем мы будем считать, что условие гладкости выполнено.

Инфинитезимальная группа голономии в точке главного расслоения $p\in\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ определяется при помощи формы кривизны R связности Γ, заданной на $\mathbb{P}$. Сначала определим индуктивно семейство подпространств $\mathfrak{m}_k(p)$, $k=0,1,2,\ldots$, в алгебре Ли $\mathfrak{g}$ структурной группы $\mathbb{G}$. Пусть $\mathfrak{m}_0(p)$ подпространство в $\mathfrak{g}$, порожденное всеми элементами вида $R_p(X,Y)$, где X,Y – произвольные горизонтальные векторы в точке $p\in\mathbb{P}$. Рассмотрим $\mathfrak{g}$-значную функцию на $\mathbb{P}$ вида

$$f_k = Z_k\ldots Z_1\Big(R(X,Y)\Big), \tag{14.66}$$

где $X,Y,Z_1,\ldots,Z_k$ – произвольные горизонтальные векторные поля на $\mathbb{P}$ и векторы $Z_1,\ldots,Z_k$ действуют на $\mathfrak{g}$-значную функцию $R(X,Y)$ как дифференцирования. Пусть $\mathfrak{m}_k(p)$ – подпространство в алгебре Ли $\mathfrak{g}$, порожденное подпространством $\mathfrak{m}_{k-1}$ и значениями в точке p всех функций f_k вида (14.66). Положим

$$\mathfrak{g}_{\mathsf{inf}}(p) = \bigcup_{k=0}^{\infty}\mathfrak{m}_k(p). \tag{14.67}$$

Предложение 14.10.2. *Подпространство $\mathfrak{g}_{\mathsf{inf}}(p)$ в $\mathfrak{g}$ есть подалгебра Ли в алгебре Ли $\mathfrak{g}_{\mathsf{loc}}(p)$ локальной группы голономии.*

Доказательство. См. [**?**]. □

Это предложение позволяет ввести новое понятие.

Определение. Связная подгруппа Ли $\Phi_{\mathsf{inf}}(p)$ в структурной группе $\mathbb{G}$, порожденная подалгеброй $\mathfrak{g}_{\mathsf{inf}}(p)$ называется *инфинитезимальной группой голономии* связности Γ в точке $p\in\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$. □

Предложение 14.10.3. *Инфинитезимальная группа голономии $\Phi_{\mathsf{inf}}(p)$ имеет следующие свойства:*

1) $\Phi_{\mathsf{inf}}(p)$ является связной подгруппой Ли локальной группы голономии $\Phi_{\mathsf{loc}}(p)$;

2) $\Phi_{\mathsf{inf}}(pa) = \mathsf{ad}\,(a^{-1})\Phi_{\mathsf{inf}}(p)$ и $\mathfrak{g}_{\mathsf{inf}}(pa) = \mathsf{ad}\,(a^{-1})\mathfrak{g}_{\mathsf{inf}}(p)$;

3) для каждого целого m множество точек базы

$$\{\pi(p)=x\in\mathbb{M}:\quad \mathsf{dim}\,\Phi_{\mathsf{inf}}(p)\geq m\} \tag{14.68}$$

открыто в $\mathbb{M}$;

4) если $\Phi_{\mathsf{inf}}(p)=\Phi_{\mathsf{loc}}(p)$ в точке p, то существует окрестность $\mathbb{U}$ точки $x=\pi(p)$ такая, что

$$\Phi_{\mathsf{inf}}(q)=\Phi_{\mathsf{loc}}(q)=\Phi_{\mathsf{inf}}(p)=\Phi_{\mathsf{loc}}(p)$$

для всех $q\in\mathbb{U}$.

Доказательство. Свойство 1) следует из предложения 14.10.2.

Свойство 2) вытекает из интуитивно понятного равенства

$$\mathfrak{m}_k(pa) = \mathsf{ad}\,(a^{-1})\mathfrak{m}_k(p),\qquad \forall k.$$

Детали доказательства приведены в [?].

3). Размерность инфинитезимальной группы голономии $\dim \Phi_{\mathsf{inf}}(p)$ в силу свойства 2) можно рассматривать как функцию на $\mathbb{M}$ со значениями в целых числах. Если значения конечного числа функций f_k вида (14.66) линейно независимы в точке p, то они таковы и в любой точке из некоторой окрестности точки p. Поэтому, если свойство (14.68) выполнено в точке x, то оно имеет место и в некоторой окрестности x.

4). Допустим, что $\Phi_{\mathsf{inf}}(p) = \Phi_{\mathsf{loc}}(p)$. Из свойства 3) предложения 14.10 и свойства 5) предложения 14.10.1 следует, что точка $x = \pi(p)$ имеет окрестность $\mathbb{U}$ такую, что

$$\dim \Phi_{\mathsf{inf}}(q) \geq \dim \Phi_{\mathsf{inf}}(p) \quad \text{и} \quad \dim \Phi_{\mathsf{loc}}(q) \leq \dim \Phi_{\mathsf{loc}}(p), \quad \forall q \in \pi^{-1}(\mathbb{U}).$$

С другой стороны, $\Phi_{\mathsf{loc}}(q) \supset \Phi_{\mathsf{inf}}(q)$ для каждого $q \in \pi^{-1}(\mathbb{U})$. Отсюда

$$\dim \Phi_{\mathsf{loc}}(q) = \dim \Phi_{\mathsf{inf}}(q) = \dim \Phi_{\mathsf{loc}}(p) = \dim \Phi_{\mathsf{inf}}(p)$$

и, следовательно, $\Phi_{\mathsf{loc}}(q) = \Phi_{\mathsf{inf}}(q)$ для каждого $q \in \pi^{-1}(\mathbb{U})$. Применяя теорему 14.10.2 к индуцированному расслоению $\mathbb{P}|_{\mathbb{U}}$, видим, что $\Phi_0(p, \mathbb{U}) = \Phi_{\mathsf{loc}}(p)$ и $\Phi_0(q, \mathbb{U}) = \Phi_{\mathsf{loc}}(q)$. Если $q \in \mathbb{P}(p, \mathbb{U})$, то $\Phi_0(p, \mathbb{U}) = \Phi_0(q, \mathbb{U})$. Поэтому $\Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{loc}}(q)$. □

Теорема 14.10.3. *Если размерность инфинитезимальной группы голономии* $\dim \Phi_{\mathsf{inf}}(q)$ *постоянна в некоторой окрестности точки* $p \in \mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$*, то локальная и инфинитезимальная группы голономий в точке* p *совпадают,* $\Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{inf}}(p)$*.*

Доказательство. См. [?] □

Следствие. Если размерность инфинитезимальной группы голономии $\dim \Phi_{\mathsf{inf}}(p)$ постоянна на всем главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, то

$$\Phi_0(p) = \Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{inf}}(p) \tag{14.69}$$

для всех $p \in \mathbb{P}$. □

Доказательство. Вытекает из теорем 14.10.2 и 14.10.3. □

Теорема 14.10.4. *Для вещественно аналитической связности* Γ *на вещественно аналитическом главном расслоении* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *следующие группы голономии равны*

$$\Phi_0(p) = \Phi_{\mathsf{loc}}(p) = \Phi_{\mathsf{inf}}(p)$$

для всех $p \in \mathbb{P}$*.*

Доказательство. См. [?]. □

Равенство групп голономий (14.69) позволяет вычислить суженную группу голономии $\Phi_0(p)$. Действительно, алгебра Ли инфинитезимальной группы голономии (14.67) порождена всеми функциями вида (14.66). Пусть x^α, $\alpha = 1, \ldots, n$, – система координат в окрестности $\mathbb{U}$ точки $x = \pi(p)$ и $\sigma : \mathbb{U} \to \mathbb{P}$ – локальное сечение. Тогда алгебра инфинитезимальной группы голономии порождается значениями компонент локальной формы кривизны $F_{\alpha\beta}{}^{\mathrm{A}}$ и всех ее ковариантных производных:

$$F_{\alpha\beta}{}^{\mathrm{A}}, \quad \nabla_{\gamma_1} F_{\alpha\beta}{}^{\mathrm{A}}, \quad \nabla_{\gamma_2}\nabla_{\gamma_1} F_{\alpha\beta}{}^{\mathrm{A}}, \quad \ldots$$

Ясно, что подпространство $\mathfrak{m}_0 \subset \mathfrak{g}$ порождено всеми компонентами тензора кривизны $F_{\alpha\beta}{}^{\mathrm{A}}$, подпространство $\mathfrak{m}_1 \subset \mathfrak{g}$ – всеми компонентами тензора кривизны $F_{\alpha\beta}{}^{\mathrm{A}}$ и их первых ковариантных производных $\nabla_{\gamma_1} F_{\alpha\beta}{}^{\mathrm{A}}$ и так далее. Таким образом, при заданной связности на главном расслоении, инфинитезимальная группа голономии позволяет в принципе вычислить суженную группу голономии. Конечно, после этого необходимо проверить, что условие следствия выполнено.

14.11 Инвариантные связности

Прежде чем рассматривать инвариантные связности общего вида, мы опишем важный частный случай.

Теорема 14.11.1. *Пусть $\mathbb{G}$ – связная группа Ли, и $\mathbb{H}$ – ее замкнутая подгруппа Ли. Пусть $\mathfrak{g}$ и $\mathfrak{h}$ – алгебры Ли для $\mathbb{G}$ и $\mathbb{H}$ соответственно.*

1) *Если существует линейное подпространство $\mathfrak{m}$ в $\mathfrak{g}$ такое, что $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и $\mathsf{ad}\,(\mathbb{H})\mathfrak{m} = \mathfrak{m}$, то $\mathfrak{h}$-компонента ω канонической 1-формы θ в $\mathbb{G}$ определяет связность на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантную относительно действия левых сдвигов из $\mathbb{G}$.*
2) *Обратно, любая связность на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантная относительно действия левых сдвигов из $\mathbb{G}$ (если она существует), определяет разложение $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ и может быть получена так, как это описано в пункте 1).*
3) *Форма кривизны R инвариантной связности, определенная формой ω из пункта 1), равна*
$$R(X, Y) = -\frac{1}{2}[X, Y]_{\mathfrak{h}},$$
где $X, Y \in \mathfrak{m}$ – произвольные левоинвариантные векторные поля на $\mathbb{G}$ из $\mathfrak{m}$ и в правой части равенства взята $\mathfrak{h}$-компонента коммутатора.
4) *Пусть $\mathfrak{g}(e)$ – алгебра Ли группы голономии $\Phi(e)$ в единице e группы Ли $\mathbb{G}$ для инвариантной связности, определенной в пункте 1). Тогда $\mathfrak{g}(e)$ порождается всеми элементами вида $[X, Y]_{\mathfrak{h}}$, где $X, Y \in \mathfrak{m}$.*

Доказательство. 1). Пусть Z^* – фундаментальное векторное поле, соответствующее элементу подалгебры $Z \in \mathfrak{h}$. Из определения канонической 1-формы (см. раздел 9.2) следует, что $\omega(Z^*) = \theta(Z^*) = Z$. Пусть $\theta_{\mathfrak{m}}$ есть $\mathfrak{m}$-компонента канонической формы θ. Для любого $a \in \mathbb{H}$ и $X \in \mathbb{T}_p(\mathbb{G})$ справедливы равенства

$$\begin{aligned} \theta(r_{a*}X) &= \omega(r_{a*}X) + \theta_{\mathfrak{m}}(r_{a*}X), \\ \mathsf{ad}\,(a^{-1})\theta(X) &= \mathsf{ad}\,(a^{-1})\omega(X) + \mathsf{ad}\,(a^{-1})\theta_{\mathfrak{m}}(X). \end{aligned}$$

Левые части этих равенств совпадают. Поскольку по условию теоремы $\mathsf{ad}\,(a^{-1})\mathfrak{m} = \mathfrak{m}$, то сравнение $\mathfrak{h}$-компонент правых частей приводит к равенству

$$\omega(r_{a*}X) = \mathsf{ad}\,(a^{-1})\omega(X),$$

т.е. 1-форма ω определяет связность на $\mathbb{G}$. Эта связность инвариантна относительно действия группы слева по построению.

2). Пусть ω – форма связности на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантная относительно действия левых сдвигов из $\mathbb{G}$. Пусть $X \in \mathfrak{m}$ – множество левоинвариантных векторных полей на $\mathbb{G}$ таких, что $\omega(X) = 0$. Тогда алгебра Ли на $\mathbb{G}$ разлагается в прямую сумму, $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$.

3). Левоинвариантное векторное поле горизонтально тогда и только тогда, когда оно лежит в $\mathfrak{m}$. Поэтому утверждение 3) следует из равенства (14.27).

4). Пусть $\mathfrak{g}_1$ – подпространство в $\mathfrak{g}$, порожденное множеством элементов вида $R_e(X, Y)$, где $X, Y \in \mathfrak{m}$. Пусть $\mathfrak{g}_2$ – подпространство в $\mathfrak{g}$, порожденное множеством элементов $R_a(X, Y)$, где $X, Y \in \mathfrak{m}$ для всех $a \in \mathbb{G}$. Тогда по теореме Амброза–Зингера $\mathfrak{g}_1 \subset \mathfrak{g}(e) \subset \mathfrak{g}_2$. С другой стороны, $\mathfrak{g}_1 = \mathfrak{g}_2$, т.к. $R_a(X, Y) = R_e(X, Y)$ для любых $X, Y \in \mathfrak{m}$ и $a \in \mathbb{G}$. Теперь 4) следует из 3). $\square$

Замечание. Линейное подпространство алгебры Ли $\mathfrak{m} \subset \mathfrak{g}$ представляет собой распределение горизонтальных векторных полей на $\mathbb{G}$, т.е. связность Γ на главном расслоении $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, инвариантную относительно действия группы слева. В общем случае $\mathfrak{m} \subset \mathfrak{g}$ является только линейным подпространством, а не подалгеброй. □

Замечание. Утверждение 1) теоремы 14.11.1 можно рассматривать как частный случай предложения 14.6.3. Пусть $\mathbb{P} = (\mathbb{G}/\mathbb{H}) \times \mathbb{G}$ – тривиальное главное расслоение над базой $\mathbb{G}/\mathbb{H}$ со структурной группой $\mathbb{G}$. Вложим расслоение $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ в $\mathbb{P}$ при помощи отображения

$$f(a) = \Big(\pi(a), a\Big), \qquad a \in \mathbb{G},$$

где $\pi : \mathbb{G} \to \mathbb{G}/\mathbb{H}$ – естественная проекция. Пусть ϕ – форма связности, определяющая каноническую плоскую связность на $\mathbb{P}$. Ее $\mathfrak{h}$-компонента, суженная на подрасслоение $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$, по предложению 14.6.3 определяет связность и совпадает с формой связности ω в утверждении 1). □

Возвращаясь к общему случаю, сначала докажем предложение, которое является основой для многих приложений.

Предложение 14.11.1. *Пусть s_t – однопараметрическая группа автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и $\widetilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное s_t. Пусть Γ – связность на $\mathbb{P}$, инвариантная относительно действия s_t. Для произвольной точки главного расслоения $p_0 \in \mathbb{P}$ определим четыре кривые p_t, x_t, $\tilde{x}_t$ и a_t следующим образом:*

$$p_t = s_t(p_0) \in \mathbb{P}, \qquad x_t = \pi(p_t) \in \mathbb{M},$$

$\tilde{x}_t$ есть горизонтальный лифт x_t такой, что $\tilde{x}_0 = p_0$ и

$$p_t = \tilde{x}_t a_t, \qquad a_t \in \mathbb{G}. \tag{14.70}$$

Тогда a_t является однопараметрической подгруппой в структурной группе $\mathbb{G}$, порожденной элементом $X = \omega_{p_0}(\widetilde{X})$, где ω – форма связности для Γ.

Доказательство. Поскольку

$$\dot{p}_t = r_{a_t *}\dot{\tilde{x}}_t + \tilde{x}_{t*}\dot{a}_t,$$

то

$$\omega(\dot{p}_t) = \mathsf{ad}\,(a_t^{-1})\omega(\dot{\tilde{x}}_t) + a_{t*}^{-1}\dot{a}_t,$$

где $\tilde{x}_{t*}$ и a_{t*} – дифференциалы отображений $\tilde{x}_t : \mathbb{G} \to \pi^{-1}(x_t)$ и $a_t : \mathbb{G} \to \mathbb{G}$ соответственно. Первое слагаемое в правой части равно нулю, т.к. кривая $\tilde{x}_t$ горизонтальна. Следовательно, $\omega(\dot{p}_t) = a_{t*}^{-1}\dot{a}_t$. С другой стороны, $\dot{p}_t = s_{t*}\widetilde{X}_{p_0}$, где s_{t*} – дифференциал отображения $s_t : \mathbb{P} \to \mathbb{P}$, и поэтому $\omega(\dot{p}_t) = \omega(\widetilde{X}_{p_0}) = X$, т.к. форма связности инвариантна относительно однопараметрической группы преобразований s_t. Отсюда следует равенство $X = a_{t*}^{-1}\dot{a}_t$. □

Определение. Кривая a_t в структурной группе Ли $\mathbb{G}$ из условия предложения 14.11.1 называется *разверткой* кривой p_t в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$. □

Замечание. В этом определении кривая p_t может быть произвольной кривой в главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, не обязательно связанной с группой симметрии связности. □

Определение. Пусть $\mathbb{K}$ – группа Ли автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ с алгеброй Ли $\mathfrak{k}$. Выберем в главном расслоении опорную точку $p_0\in\mathbb{P}$. Каждый элемент из $\mathbb{K}$ индуцирует некоторое преобразование базы при помощи проекции. Множество $\mathbb{J}$ всех элементов из $\mathbb{K}$, которые оставляют неподвижной точку $x_0=\pi(p_0)$, образуют замкнутую подгруппу в $\mathbb{K}$, которая называется *подгруппой изотропии* в $\mathbb{K}$ для точки $x_0\in\mathbb{M}$. Определим гомоморфизм групп Ли

$$\lambda:\quad \mathbb{J}\ni j\mapsto\lambda(j)=a\in\mathbb{G} \tag{14.71}$$

следующим образом. Каждый автоморфизм $j\in\mathbb{J}$ переводит p_0 в точку $j(p_0)$, которая принадлежит тому же слою $\pi^{-1}\Big(\pi(p_0)\Big)$, т.к. точка x_0 неподвижна. Следовательно, $j(p_0)=p_0a$ для некоторого $a\in\mathbb{G}$. Положим $\lambda(j)=a$. Тогда для двух элементов подгруппы изотропии, $j_1,j_2\in\mathbb{J}$, справедливы равенства

$$p_0\lambda(j_1j_2)=(j_1j_2)(p_0)=j_1\Big(p_0\lambda(j_2)\Big)=\Big(j_1(p_0)\Big)\lambda(j_2)=p_0\lambda(j_1)\lambda(j_2).$$

Тем самым $\lambda(j_1j_2)=\lambda(j_1)\lambda(j_2)$ и, следовательно, построенное отображение является гомоморфизмом групп. Нетрудно проверить, что отображение $\lambda:\ \mathbb{J}\to\mathbb{G}$ дифференцируемо. Гомоморфизм групп λ индуцирует гомоморфизм алгебр Ли,

$$\lambda:\quad \mathfrak{j}\to\mathfrak{g},$$

который мы обозначили той же буквой. □

Замечание. Отображение (14.71) зависит от выбора опорной точки $p_0\in\mathbb{P}$. В настоящем разделе мы будем считать, что опорная точка p_0 выбрана и зафиксирована. □

Предложение 14.11.2. *Пусть $\mathbb{K}$ – группа Ли автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ с алгеброй Ли $\mathfrak{k}$ и Γ – связность на $\mathbb{P}$ с формой связности ω и формой кривизны R, инвариантная относительно автоморфизмов $\mathbb{K}$. Пусть $\mathbb{J}\subset\mathbb{K}$ – подгруппа изотропии в $\mathbb{K}$ для точки $x_0=\pi(p_0)$ с алгеброй Ли $\mathfrak{j}$. Определим линейное отображение*

$$\Lambda:\quad \mathfrak{k}\ni X\mapsto\Lambda(X)=\omega_{p_0}(\widetilde{X})\in\mathfrak{g},$$

где $\widetilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное полем X. Тогда:

1) $\Lambda(X)=\lambda(X)$, $\forall X\in\mathfrak{j}$;
2) $\Lambda\Big(\mathsf{ad}\,(j)X\Big)=\mathsf{ad}\,\Big(\lambda(j)\Big)\Lambda(X)$, $\forall j\in\mathbb{J}$ *и* $\forall X\in\mathfrak{k}$, *где* $\mathsf{ad}\,(j)$ *обозначает присоединенное представление подгруппы изотропии $\mathbb{J}$ в $\mathfrak{k}$ и $\mathsf{ad}\,\Big(\lambda(j)\Big)$ – присоединенное представление структурной группы $\mathbb{G}$ в $\mathfrak{g}$;*
3) $2R_{p_0}(\widetilde{X},\widetilde{Y})=[\Lambda(X),\Lambda(Y)]-\Lambda\Big([X,Y]\Big),\qquad\forall X,Y\in\mathfrak{k}.$

Доказательство. См., например, [?]. □

Замечание. Геометрический смысл отображения Λ дается предложением 14.11.1. $\Lambda(X)$ – это тот элемент алгебры Ли $\mathfrak{g}$, который порождает однопараметрическую подгруппу a_t в структурной группе $\mathbb{G}$, определенную равенством (14.70). □

Замечание. Отображение Λ в предложении 14.11.2 является только линейным. В общем случае оно не является гомоморфизмом алгебр Ли. □

Определение. Группа автоморфизмов $\mathbb{K}$ действует на главном расслоении $\mathbb{P}(\mathbb{M},\pi,\mathbb{G})$ *слой-транзитивно*, если для любых двух слоев из $\mathbb{P}$ существует элемент в $\mathbb{K}$, отображающий один слой в другой, т.е. если действие $\mathbb{K}$ на базе $\mathbb{M}$ транзитивно. □

Предложение 14.11.3. *Если $\mathbb{J}$ – подгруппа изотропии для слой-транзитивной группы автоморфизмов $\mathbb{K}$ в точке $x_0 = \pi(p_0)$, то база $\mathbb{M}$ является однородным пространством, $\mathbb{M} \approx \mathbb{K}/\mathbb{J}$.*

Доказательство. Вытекает из теоремы 10.1.2. □

Описание $\mathbb{K}$-инвариантных связностей на главном расслоении в случае слой-транзитивного действия групп автоморфизмов дает

Теорема 14.11.2. *Если связная группа Ли $\mathbb{K}$ является слой-транзитивной группой автоморфизмов главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ и если $\mathbb{J}$ – подгруппа изотропии в $\mathbb{K}$ для точки $x_0 = \pi(p_0)$, то существует взаимно однозначное соответствие между множеством $\mathbb{K}$-инвариантных связностей на $\mathbb{P}$ и множеством линейных отображений $\Lambda : \mathfrak{k} \to \mathfrak{g}$, которые удовлетворяют условиям 1) и 2) предложения 14.11.2. Соответствие задается следующим образом:*

$$\Lambda : \quad \mathfrak{k} \ni X \mapsto \Lambda(X) = \omega_{p_0}(\widetilde{X}) \in \mathfrak{g},$$

где $\widetilde{X}$ – векторное поле на $\mathbb{P}$, индуцированное полем X.

Доказательство. См. [**?**]. □

Замечание. Из теоремы 14.1.1 следует, что связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ однозначно определяется заданием семейства локальных форм связности на произвольном покрытии базы $\mathbb{M}$. Это происходит потому что при помощи действия структурной группы распределение горизонтальных подпространств с локальных сечений можно разнести на все пространство главного расслоения $\mathbb{P}$. Если известно, что связность инвариантна также относительно слой-транзитивных автоморфизмов $\mathbb{K}$, то локальную форму связности достаточно задать в одной точке базы $\mathbb{M}$. Действительно, при помощи автоморфизмов $\mathbb{K}$ она разносится по всей базе $\mathbb{M}$, а затем, действуя структурной группой, ее можно разнести по всему пространству расслоения $\mathbb{P}$.

Если группа автоморфизмов $\mathbb{K}$ не является слой-транзитивной, тогда в задании инвариантной связности появляется значительный произвол. В этом случае форму связности можно определить на каждой орбите действия группы $\mathbb{K}$ на $\mathbb{P}$. Для этого достаточно задать ее в какой либо одной точке на каждой орбите и при помощи $\mathbb{K}$ разнести ее по орбитам и, наконец, по всему пространству главного расслоения $\mathbb{P}$, действуя структурной группой $\mathbb{G}$. После этого необходимо проверить дифференцируемость полученного распределения горизонтальных подпространств. □

Если группа автоморфизмов $\mathbb{K}$ действует на $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ слой-транзитивно, то форма кривизны R, которая является тензориальной формой типа $\mathsf{ad}\,\mathbb{G}$, инвариантная относительно $\mathbb{K}$, полностью определяется своими значениями в опорной точке $R_{p_0}(\widetilde{X}, \widetilde{Y})$, где $\widetilde{X}$ и $\widetilde{Y}$ – векторные поля на $\mathbb{P}$, индуцированные элементами $X, Y \in \mathfrak{k}$ алгебры Ли группы автоморфизмов $\mathbb{K}$. В этом случае утверждение 3) предложения 14.11.2 выражает форму кривизны $R_{p_0}(\widetilde{X}, \widetilde{Y})$ в терминах Λ.

Из предложения 14.11.2 и теорем 14.9.1 и 14.11.2 получаем

Следствие. $\mathbb{K}$-инвариантная связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, определенная отображением Λ, является плоской тогда и только тогда, когда отображение $\Lambda : \mathfrak{k} \to \mathfrak{g}$ есть гомоморфизм алгебр Ли. □

Теорема 14.11.3. *Допустим, что в теореме 14.11.2 алгебра Ли* $\mathfrak{k}$ *содержит линейное подпространство* $\mathfrak{m}$ *такое, что* $\mathfrak{k} = \mathfrak{j} \oplus \mathfrak{m}$ *и* $\mathsf{ad}\,(\mathbb{J})\mathfrak{m} = \mathfrak{m}$, *где* $\mathsf{ad}\,(\mathbb{J})$ *– присоединенное представление* $\mathbb{J}$ *в* $\mathfrak{k}$. *Тогда:*

1) Существует взаимно однозначное соответствие между множеством $\mathbb{K}$*-инвариантных связностей на главном расслоении* $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ *и множеством линейных отображений* $\Lambda_{\mathfrak{m}} : \ \mathfrak{m} \to \mathfrak{g}$ *таких, что*

$$\Lambda_{\mathfrak{m}}\Big(\mathsf{ad}\,(j)X\Big) = \mathsf{ad}\,\Big(\lambda(j)\Big)\Lambda_{\mathfrak{m}}(X), \qquad \forall X \in \mathfrak{m} \ \textit{и} \ \ \forall j \in \mathbb{J};$$

соответствие задается теоремой 14.11.2 следующим образом

$$\Lambda(X) = \begin{cases} \lambda(X), & \textit{если} \ \ X \in \mathfrak{j}, \\ \Lambda_{\mathfrak{m}}(X), & \textit{если} \ \ X \in \mathfrak{m}. \end{cases}$$

2) Форма кривизны R *для* $\mathbb{K}$*-инвариантной связности, определяемой при помощи отображения* $\Lambda_{\mathfrak{m}}$, *удовлетворяет следующему равенству*

$$2R_{p_0}(\widetilde{X}, \widetilde{Y}) = [\Lambda_{\mathfrak{m}}(X), \Lambda_{\mathfrak{m}}(Y)] - \Lambda_{\mathfrak{m}}\Big([X, Y]_{\mathfrak{m}}\Big) - \lambda\Big([X, Y]_{\mathfrak{j}}\Big), \qquad \forall X, Y \in \mathfrak{m},$$

где $[X, Y]_{\mathfrak{m}}$ *и* $[X, Y]_{\mathfrak{j}}$ *обозначают соответственно* $\mathfrak{m}$*- и* $\mathfrak{j}$*-компоненту коммутатора* $[X, Y] \in \mathfrak{k}$.

Доказательство. Пусть $\Lambda : \ \mathfrak{k} \to \mathfrak{g}$ – линейное отображение, удовлетворяющее утверждениям 1) и 2) предложения 14.11.2. Пусть $\Lambda_{\mathfrak{m}}$ – сужение отображения Λ на $\mathfrak{m}$. Нетрудно проверить, что отображение $\Lambda \to \Lambda_{\mathfrak{m}}$ является взаимно однозначным и согласно теореме 14.11.2 дает желаемое соответствие. Утверждение 2) следует из 3) предложения 14.11.2. □

Определение. В теореме 14.11.3 $\mathbb{K}$-инвариантная связность на главном расслоении $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$, определяемая условием $\Lambda_{\mathfrak{m}} = 0$ называется *канонической инвариантной связностью* относительно разложения $\mathfrak{g} = \mathfrak{j} \oplus \mathfrak{m}$. □

Следующая теорема определяет алгебру Ли группы голономии $\mathbb{K}$-инвариантной связности.

Теорема 14.11.4. *В предположениях и обозначениях теоремы 14.11.2 алгебра Ли* $\mathfrak{g}(p_0)$ *группы голономии* $\Phi(p_0)$ *для* $\mathbb{K}$*-инвариантной связности, определяемой при помощи линейного отображения* $\Lambda : \ \mathfrak{k} \to \mathfrak{g}$, *задается суммой*

$$\mathfrak{m}_0 + [\Lambda(\mathfrak{k}), \mathfrak{m}_0] + \Big[\Lambda(\mathfrak{k}), [\Lambda(\mathfrak{k}), \mathfrak{m}_0]\Big] + \ldots,$$

где $\mathfrak{m}_0$ *– линейное подпространство в* $\mathfrak{g}$, *порожденное множеством*

$$\{[\Lambda(X), \Lambda(Y)] - \Lambda([X, Y])\}, \qquad \forall X, Y \in \mathfrak{k}.$$

Доказательство. См., например, [**?**]. □

Замечание. Утверждения 1) и 3) теоремы 14.11.1 следуют из теоремы 14.11.3, если в качестве главного расслоения $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ выбрать $\mathbb{G}(\mathbb{G}/\mathbb{H}, \pi, \mathbb{H})$ и положить $\mathbb{K} = \mathbb{G}$. Тогда инвариантная связность из теоремы 14.11.1 является канонической инвариантной связностью. Утверждение 4) теоремы 14.11.1 следует из теоремы 14.11.4. □

Глава 15

Приложения в квантовой механике

В настоящей главе рассмотрены некоторые приложения дифференциальной геометрии в квантовой механике. Нетривиальные геометрические структуры, а речь идет о нетривиальной связности на главном расслоении, часто возникают при решении уравнений математической физики. В настоящей главе будет показано, как возникает нетривиальная связность на главном расслоении со структурной группой $\mathbb{U}(1)$ или $\mathbb{U}(\mathrm{N})$ в нерелятивистской квантовой механике при решении уравнения Шредингера. Удивительно не столько то, что главное расслоение возникает естественным образом, а то, что предсказанные эффекты были подтверждены экспериментально.

Сначала мы дадим геометрическую интерпретацию нерелятивистской квантовой механике в конечномерном случае. Будет показано, что гамильтониан квантовой системы задает компоненты локальной формы связности на главном расслоении, а уравнение Шредингера определяет параллельный перенос слоев. При этом базой является одномерное многообразие, соответствующее времени, а структурной группой – унитарная группа $\mathbb{U}(\mathrm{N})$, где N – размерность гильбертова пространства состояний квантовомеханической системы. Решение квантовомеханической задачи не зависит от выбора базиса в гильбертовом пространстве, и его выбирают из соображений удобства. Использование базиса, состоящего из собственных векторов гамильтониана, позволяет упростить доказательство адиабатической теоремы и сделать его более прозрачным.

В качестве приложения адиабатической теоремы рассмотрена фаза Берри. В заключительном разделе настоящей главы рассмотрен эффект Ааронова–Бома, который, хотя и не имеет прямого отношения к адиабатической теореме, с геометрической точки зрения аналогичен фазе Берри.

Эффект Ааронова–Бома [**?**] и фаза Берри [**?**] привлекают большое внимание теоретиков и экспериментаторов в течении многих лет. Интерес вызван двумя обстоятельствами. Во-первых, в обоих случаях при решении уравнения Шредингера естественным образом возникает $\mathbb{U}(1)$-связность. Во-вторых, в теории калибровочных полей распространено мнение, что к наблюдаемым эффектам может приводить нетривиальная напряженность поля, а не сами потенциалы, которые не являются калибровочно инвариантными. Вопреки этому мнению Ааронов и Бом и Берри показали, что интеграл от калибровочного поля вдоль замкнутой кривой может привести к наблюдаемым эффектам. Эти выводы вскоре были подтверждены экспериментально.

Понятие фазы Берри было обобщено на неабелев случай, соответствующий вырожденным уровням энергии гамильтониана, Вилчеком и Зи [**?**]. В этом случае при решении уравнения Шредингера естественным образом возникают неабелевы калибровочные поля.

Во всех перечисленных выше случаях к наблюдаемым эффектам приводят элементы группы голономии (см. раздел 14.4) соответствующих связностей. Элементы группы голономии в общем случае являются ковариантными объектами, а для абелевой группы $\mathbb{U}(1)$ – инвариантными. Показано, что главное расслоение может быть тривиальным, но связность, которая на нем возникает, в общем случае имеет нетривиальную группу голономии и приводит к наблюдаемым эффектам. Отсюда следует, что фаза Берри и эффект Ааронова–Бома имеют геометрическую природу.

15.1 Адиабатическая теорема

Адиабатическая теорема [?] занимает одно из центральных мест в нерелятивистской квантовой механике, т.к. позволяет находить приближенное решение уравнения Шредингера при медленном изменении гамильтониана во времени. Первоначально она была доказана для дискретного (возможно, бесконечного) спектра гамильтониана при некоторых ограничениях на возможное пересечение уровней энергии [?]. Доказательство для невырожденных уровней энергии приведено, например, в [?]. Различным вариантам адиабатической теоремы посвящено большое количество статей, ссылки на которые можно найти в [?]. Доказательства адиабатической теоремы довольно сложны. Ниже будет приведено доказательство адиабатической теоремы в конечномерном случае [?].

В нерелятивистской квантовой механике состояние системы описывается вектором гильбертова пространства (*волновой функцией*) $\psi(t) \in \mathbb{H}$, зависящим от времени t и некоторого набора других переменных, который определяется рассматриваемой задачей. Эволюция квантовой системы во времени описывается *уравнением Шредингера* [?, ?]

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi, \tag{15.1}$$

где H – самосопряженный линейный оператор, действующий в гильбертовом пространстве $\mathbb{H}$, который называется *гамильтонианом* системы, и $\hbar$ – постоянная Планка.

Для уравнения Шредингера, как правило, ставится задача Коши с начальным условием

$$\psi(0) = \psi_0, \tag{15.2}$$

где $\psi_0 \in \mathbb{H}$ – некоторый фиксированный вектор гильбертова пространства.

В дальнейшем, для простоты, положим $\hbar = 1$ и обозначим частную производную по времени, точкой, $\dot{\psi} = \partial_t\psi$.

Предположим, для простоты, что гильбертово пространство представляет собой конечномерное комплексное пространство $\mathbb{H} = \mathbb{C}^{\mathrm{N}}$ комплексной размерности $\dim\mathbb{H} = \mathrm{N}$. В гильбертовом пространстве задано скалярное произведение, которое обозначим круглыми скобками,

$$\mathbb{H}\times\mathbb{H} \ni \ \psi,\phi \mapsto (\psi,\phi) \ \in \mathbb{C}.$$

По определению, скалярное произведение линейно по первому аргументу ψ и выполнено равенство: $(\psi,\phi)^\dagger = (\phi,\psi)$, где символ $\dagger$ обозначает комплексное сопряжение. Квадрат вектора гильбертова пространства (ψ,ψ) является вещественным числом, при этом мы требуем, чтобы квадратичная форма (ψ,ψ) была строго положительно определена, т.е. $(\psi,\psi) \geq 0$, причем $(\psi,\psi) = 0$ тогда и только тогда, когда $\psi = 0$.

Тогда скалярное произведение определяет норму вектора гильбертова пространства:

$$\|\psi\| = \sqrt{(\psi, \psi)}.$$

Поскольку уравнение Шредингера линейно по ψ, а гамильтониан самосопряжен, то норма произвольного решения уравнений Шредингера сохраняется во времени. Отсюда следует, что векторы состояния можно нормировать. Обычно предполагается, что векторы состояния нормированы на единицу,

$$\|\psi\| = 1. \tag{15.3}$$

Нормировка вектора состояния не устраняет полностью произвол в выборе вектора гильбертова пространства, т.к. остается произвол в выборе постоянного фазового множителя.

Замечание. Мы используем символ † для обозначения эрмитова сопряжения матриц, т.е. транспонирования матрицы и комплексного сопряжения всех элементов. В частном случае, когда матрица состоит из одного элемента, эрмитово сопряжение совпадает с комплексным. □

Пусть в гильбертовом пространстве $\mathbb{H}$ выбран некоторый базис e_k, $k = 1, \ldots, \mathrm{N}$. Тогда гамильтониан квантовомеханической системы задается эрмитовой $\mathrm{N}\times\mathrm{N}$-матрицей, а вектор состояния $\psi = \psi^k e_k$ – строкой из N компонент,

$$\psi = (\psi^1, \ldots, \psi^{\mathrm{N}}),$$

где $\psi^1, \ldots, \psi^{\mathrm{N}}$ – комплекснозначные компоненты вектора. Если базис гильбертова пространства ортонормирован,

$$(e_k, e_l) = \delta_{kl},$$

то скалярное произведение задается равенством

$$(\psi, \phi) = \psi\phi^\dagger = \psi^1\phi_1^\dagger + \ldots + \psi^{\mathrm{N}}\phi_{\mathrm{N}}^\dagger.$$

Рассмотрим задачу Коши (15.1), (15.2) в общем случае, когда гамильтониан системы зависит от времени $H = H(t)$. Для решения этой задачи необходимо выбрать базис в гильбертовом пространстве $\mathbb{H}$. Конечно, решение задачи от выбора базиса не зависит, и его выбирают из соображений удобства. Рассмотрим два случая.

Пусть базис $e_k \in \mathbb{H}$, $k = 1, \ldots, \mathrm{N}$, ортонормирован и фиксирован, $\dot{e}_k = 0$. Произвольный вектор можно разложить по этому базису $\psi = \psi^k e_k$. При этом гамильтониан задается эрмитовой $\mathrm{N} \times \mathrm{N}$-матрицей $H_l{}^k$, а задача Коши для уравнения Шредингера в компонентах примет вид системы обыкновенных дифференциальных уравнений с некоторыми начальными условиями:

$$\begin{aligned} i\dot{\psi}^k &= \psi^l H_l{}^k, \\ \psi^k(0) &= \psi_0^k. \end{aligned} \tag{15.4}$$

Замечание. Мы записываем действие гамильтониана в конечномерном случае справа, чтобы согласовать наши обозначения с обозначениями, принятыми в дифференциальной геометрии. Напомним, что в дифференциальной геометрии для векторного поля в координатном базисе принята запись $X = X^\alpha \partial_\alpha$. (Суммирование с десяти до четырех по циферблату часов.) Альтернативная запись $X = \partial_\alpha X^\alpha$ используется для

обозначения дивергенции векторного поля. Преобразование координат мы записываем в виде

$$X = X^\alpha \partial_\alpha = X^\alpha \frac{\partial x^{\alpha'}}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x^{\alpha'}} \partial_\beta.$$

То есть действие матрицы преобразования координат записывается справа. С другой стороны, действие операторов в квантовой механике общепринято записывать слева. Поэтому, если мы хотим использовать единообразные обозначения, то необходимо чем то пожертвовать. В конечномерном случае мы будем записывать матрицу, задающую линейный оператор, справа. Это вопрос соглашения, и к нему легко привыкнуть. □

Рассмотрим теперь другой ортонормированный базис b_k, который может зависеть от времени, $b_k = b_k(t)$. Такой базис может оказаться более удобным для решения некоторых задач. Вектор гильбертова пространства ψ можно разложить также по этому базису $\psi = \psi'^k b_k$. Тогда задача Коши (15.4) будет выглядеть по другому:

$$\begin{aligned} i\dot{\psi}'^k &= \psi'^l H'^k_l, \\ \psi'^k(0) &= \psi'^k_0, \end{aligned} \tag{15.5}$$

где H'^k_l – компоненты гамильтониана относительно нового базиса, которые будут определены ниже. Поскольку базисы ортонормированы, то они связаны между собой некоторым унитарным преобразованием:

$$b_k = S_k{}^l e_l, \qquad S \in \mathbb{U}(\mathrm{N}), \tag{15.6}$$

которое в общем случае зависит от времени, $S = S(t)$. При этом компоненты вектора гильбертова пространства преобразуются с помощью обратной матрицы,

$$\psi'^k = \psi^l S^{-1}{}^k_l.$$

Отсюда следует выражение для компонент начального вектора гильбертова пространства $\psi'^k_0 = \psi^l_0 S^{-1}{}^k_l(0)$. Переписав уравнение Шредингера (15.4) в базисе b_k, получим компоненты гамильтониана относительно нового базиса

$$H' = SHS^{-1} + iS\dot{S}^{-1} = SHS^{-1} - i\dot{S}S^{-1}, \tag{15.7}$$

где мы, для краткости, опустили матричные индексы. Мы видим, что компоненты гамильтониана преобразуются также, как компоненты локальной формы $\mathbb{U}(\mathrm{N})$-связности.

Теперь можно дать геометрическую интерпретацию нерелятивистской квантовой механике. Пусть время пробегает всю вещественную прямую, $t \in \mathbb{R}$. Тогда мы имеем главное расслоение $\mathbb{P}\Big(\mathbb{R}, \pi, \mathbb{U}(\mathrm{N})\Big) \approx \mathbb{R} \times \mathbb{U}(\mathrm{N})$ с базой $\mathbb{R}$, типичным слоем $\mathbb{U}(\mathrm{N})$ и проекцией $\pi : \ \mathbb{P} \to \mathbb{R}$ (см. раздел 13.1). Это расслоение тривиально, т.к. базой является вещественная прямая. Гамильтониан квантовой системы задает компоненты локальной формы $\mathbb{U}(\mathrm{N})$-связности (1-форма на $\mathbb{R}$ со значениями в алгебре Ли):

$$A_t = \{iH_l{}^k\} \in \mathfrak{u}(\mathrm{N}),$$

где t – координатный ковариантный индекс, принимающий одно значение, который раньше обозначался α. Вектор гильбертова пространства $\psi \in \mathbb{H}$ – это сечение ассоциированного расслоения $\mathbb{E}\Big(\mathbb{R}, \pi_\mathbb{E}, \mathbb{H}, \mathbb{U}(\mathrm{N}), \mathbb{P}\Big)$, типичным слоем которого является

гильбертово пространство $\mathbb{H}$. Уравнение Шредингера имеет вид равенства нулю ковариантной производной,

$$\nabla_t \psi = \dot{\psi} + \psi A_t = 0,$$

т.е. задает параллельный перенос вектора гильбертова пространства. При изменении сечения компоненты связности преобразуются по-правилу

$$A'_t = S A_t S^{-1} + \dot{S} S^{-1},$$

как и положено компонентам локальной формы связности (14.15). Кривизна этой связности равна нулю, поскольку база одномерна.

Решение задачи Коши для уравнения Шредингера (15.1), (15.2) не зависит от выбора базиса. Поэтому его выбирают из соображений удобства.

Пример 15.1.1. Пусть в фиксированном базисе e_k вектор гильбертова пространства имеет вид $\psi^k = \psi_0^l U_l{}^k$, где унитарная матрица $U_l{}^k(t)$ задает оператор эволюции квантовой системы, который, по-определению, удовлетворяет дифференциальному уравнению

$$i\dot{U} = UH,$$

с начальным условием $U_l{}^k(0) = \delta_l^k$. Тогда нетрудно проверить, что оператор эволюции задает переход к такому базису гильбертова пространства $b_k = U^{-1}{}_k^l e_l$, в котором гамильтониан равен нулю, $H' = 0$. Следовательно, вектор гильбертова пространства, описывающий эволюцию квантовой системы, в этом базисе имеет постоянные компоненты ψ_0^k, которые определяются начальным состоянием. □

Замечание. При преобразовании базиса, которое зависит от времени, эрмитова матрица, соответствующая гамильтониану, испытывает калибровочное преобразование (15.7). При этом преобразовании в общем случае собственные значения матрицы меняются. Рассмотренный выше пример показывает, что если гамильтониан, заданный в постоянном базисе, имел некоторый спектр, то после перехода к новому базису, заданному оператором эволюции, гамильтониан H' обращается в нуль, и имеет только нулевые собственные значения. В квантовой механике уравнение Шредингера обычно задают, определив гамильтониан в постоянном базисе, исходя из физических соображений. Затем, если это удобнее, можно перейти к новому базису, зависящему от времени. □

Перейдем к определению адиабатического предела и описанию базиса $b_k(t)$, который будет использован при доказательстве адиабатической теоремы. Адиабатическая теорема справедлива для гамильтонианов, которые медленно меняются со временем. А именно, мы предполагаем, что гамильтониан достаточно гладко зависит от вещественного параметра $\nu = \epsilon t$, где $\epsilon > 0$, который меняется на конечном отрезке, $\nu \in [0, \nu_0]$. Тогда медленное изменение гамильтониана означает, что параметр ν меняется на конечную величину при малых ϵ и больших временах t.

Определение. Двойной предел в решении задачи Коши для уравнения Шредингера (15.1) и (15.2) на отрезке $[0, t]$

$$\epsilon \to 0, \quad t \to \infty, \quad \text{при условии} \ \ \epsilon t = \nu = \mathsf{const}. \tag{15.8}$$

называется *адиабатическим*. □

При исследовании адиабатического предела время t в уравнении Шредингера удобно заменить на параметр ν:

$$i\epsilon\frac{\partial\psi}{\partial\nu} = \psi H(\nu). \tag{15.9}$$

Вектор состояния $\psi(\nu,\epsilon)$ в таком случае зависит также от параметра ϵ, а адиабатический предел соответствует простому пределу $\epsilon\to 0$ при каждом значении параметра $\nu\in[0,\nu_0]$.

Асимптотическое решение уравнения вида (15.9) в общем случае построено в [**?**, **?**].

Для доказательства адиабатической теоремы нам понадобится специальный базис $b_k(\nu)$, зависящий от ν. Пусть исходный гамильтониан $H(\nu)$ квантовой системы задан в некотором фиксированном базисе e_k. Тогда существует унитарная матрица $S(\nu)$, которая диагонализирует гамильтониан,

$$SH(\nu)S^{-1} = H_{\text{D}}(\nu) = \mathsf{diag}\Big(E_1(\nu),\dots,E_{\text{N}}(\nu)\Big), \tag{15.10}$$

где $E_1\le E_2\le\ldots\le E_{\text{N}}$ – уровни энергии собственных состояний гамильтониана H, которые будем считать упорядоченными. Пусть b_k – собственные векторы исходного гамильтониана,

$$b_kH = E_kb_k, \tag{15.11}$$

для всех ν. Как известно, строками матрицы преобразования ${S_k}^l$, где индекс k фиксирован и $l=1,\dots,\text{N}$, являются компоненты собственных векторов $b_k = {b_k}^le_l$ гамильтониана H: ${S_k}^l = {b_k}^l$. То есть гамильтониан $H(\nu)$ в базисе b_k диагонален. Унитарная матрица S определена неоднозначно, и произвол в ее выборе в дальнейшем рассмотрении будет использован.

Мы допускаем, что часть уровней энергии может быть вырождена. Обозначим через Υ_n множество индексов, для которых $E_j(\nu)=E_n(\nu)$ при $j\in\Upsilon_n$. Конечно, в качестве индекса n можно выбрать любой индекс, принадлежащий Υ_n. Если уровень E_n невырожден, то множество индексов состоит из одного элемента: $\Upsilon_n=\{n\}$. Мы докажем адиабатическую теорему в случае, когда множество индексов Υ_n для всех n не меняется со временем, т.е. уровни энергии не пересекаются.

Мы также предполагаем, что гамильтониан H, уровни энергии $E_1,\dots,E_{\text{N}}$ и матрица преобразования S достаточно гладко зависят от ν на конечном отрезке $[0,\nu_0]$.

Для доказательства адиабатической теоремы нам понадобится

Лемма 15.1.1. *Существует унитарная матрица S в (15.10) такая, что выполнено условие*

$$\left(\frac{dS}{d\nu}S^{-1}\right)_k^{\ j} = 0, \qquad \forall k\in\Upsilon_j. \tag{15.12}$$

Доказательство. Рассмотрим два случая. Пусть уровень энергии E_k невырожден. Матрица преобразования S определена с точностью до умножения каждой строки на фазовый множитель ${S_k}^j\mapsto {S_k}^j\,\mathsf{e}^{i\alpha_k(\nu)}$ для всех $j=1,\dots,\text{N}$. Это соответствует произволу в выборе фазового множителя у собственного вектора состояния (15.11). Пусть фазовый множитель удовлетворяет уравнению

$$\frac{d\alpha_k}{d\nu} = i\sum_{j=1}^{\text{N}}\frac{d{S_k}^j}{d\nu}{S^{-1}}_j^{\,k},$$

где суммирование по k в правой части отсутствует. Тогда нетрудно проверить, что после преобразования для любого решения этого уравнения выполнено равенство

$$\left(\frac{dS}{d\nu}S^{-1}\right)_k^k = 0. \tag{15.13}$$

Это можно проделать для всех невырожденных уровней одновременно, выбрав подходящим образом фазы $\alpha_k(\nu)$.

Теперь предположим, что все уровни энергии вырождены, $E_1 = \ldots = E_{\mathrm{N}}$. Тогда матрица S определена с точностью до унитарного преобразования

$$S \mapsto WS, \qquad W(\nu) \in \mathbb{U}(\mathrm{N}).$$

Пусть матрица W удовлетворяет уравнению

$$\frac{dW}{d\nu} + W\frac{dS}{d\nu}S^{-1} = 0,$$

которое всегда имеет решение. Тогда после преобразования для любого решения будет выполнено равенство (15.12) для всех j, k.

Если вырождена только часть уровней, то соответствующее унитарное преобразование необходимо проделать только с этими уровнями. Таким образом, равенство (15.12) будет выполнено для всех уровней с $E_j = E_k$. □

Доказательство адиабатической теоремы будет проведено в ортонормированном базисе (15.6), где матрица S выбрана таким образом, как описано в лемме 15.1.1. Этот базис состоит из собственных векторов исходного гамильтониана и гамильтониан $H(\nu)$ в нем диагонален (15.10). Компоненты вектора состояния в базисе b_k, как и ранее, пометим штрихом, $\psi = \psi'^k b_k$. Поскольку гамильтониан H в этом базисе диагонален, то квадрат модуля k-той компоненты вектора состояния

$$|(\psi, b_k)|^2 = |\psi'^k|^2,$$

где круглые скобки обозначают скалярное произведение в $\mathbb{H}$, равен вероятности обнаружить квантовую систему в состоянии E_k.

Для формулировки теоремы нам понадобится функция

$$\triangle E_n(\nu) = \min_{j,\sigma} |E_j(\sigma) - E_n(\sigma)|, \qquad \forall \sigma \in [0, \nu],$$

где минимум $|E_j - E_n|$ берется по всем j, для которых $E_j \neq E_n$, и всем $\sigma \in [0, \nu]$. Поскольку уровни энергии не пересекаются, то для каждого значения параметра ν функция $\triangle E_n(\nu)$ конечна и равна минимальному расстоянию от уровня энергии E_n до остальных уровней энергии.

Теорема 15.1.1 (Адиабатическая теорема). *Пусть гамильтониан $H = H(\nu)$, его собственные состояния $b_k(\nu)$ и уровни энергии $E_k(\nu)$ достаточно гладко зависят от ν на конечном отрезке $\nu \in [0, \nu_0]$. Предположим, что число вырожденных собственных состояний постоянно во времени. Пусть $\psi_{(n)}(\nu, \epsilon)$ – решение уравнения Шредингера, которое в начальный момент времени совпадает с собственным состоянием $b_n(0)$ гамильтониана $H(0)$, соответствующем уровню энергии $E_n(0)$. Тогда в адиабатическом пределе (15.8) справедлива следующая оценка нормы*

$$1 - \sum_{j \in \Upsilon_n} |(\psi_{(n)}, b_j)|^2 = \frac{\mathsf{O}(\epsilon^2)}{\triangle E_n^2(\nu)}, \qquad \forall \nu \in [0, \nu_0]. \tag{15.14}$$

То есть в процессе эволюции квантовая система будет оставаться в собственном состоянии гамильтониана $H(\nu)$, соответствующем уровню энергии $E_n(\nu)$, с точностью порядка ϵ^2.

Доказательство. Будем решать задачу Коши (15.5) в базисе (15.6). Гамильтониан, который входит в уравнение Шредингера, в этом базисе диагонален с точностью до линейных членов по ϵ,

$$H' = H_{\text{D}} - i\epsilon \frac{dS}{d\nu} S^{-1}.$$

Пусть матрица S выбрана таким образом, как описано в лемме 15.1.1. Предположим, что в начальный момент времени система находится в собственном состоянии гамильтониана H_{D} и, следовательно, в собственном состоянии исходного гамильтониана $H = S^{-1} H_{\text{D}} S$. Это значит, что начальное условие в базисе b_k имеет вид

$$\psi_{(n)}(0,\epsilon) = b_n(0) = (\underbrace{0,\dots,0}_{n-1}, 1, 0\dots, 0).$$

Любое решение уравнения Шредингера представимо в виде

$$\psi_{(n)}(\nu,\epsilon) = \phi_{(n)}(\nu,\epsilon)\,\mathsf{exp}\left(-\frac{i}{\epsilon}\int_0^\nu d\sigma H_{\text{D}}(\sigma)\right), \tag{15.15}$$

где $\phi_{(n)}$ – некоторый вектор гильбертова пространства $\mathbb{H}$. Тогда для вектора $\phi_{(n)}$ получаем уравнение

$$\frac{\partial \phi_{(n)}}{\partial \nu} = -\phi_{(n)}\,\mathsf{exp}\left(-\frac{i}{\epsilon}\int_0^\nu d\sigma H_{\text{D}}\right)\frac{dS}{d\nu}S^{-1}\,\mathsf{exp}\left(\frac{i}{\epsilon}\int_0^\nu d\sigma H_{\text{D}}\right),$$

Полученное уравнение вместе с начальным условием перепишем в виде интегрального уравнения

$$\phi_{(n)}(\nu,\epsilon) = \psi'_{(n)}(0) - \int_0^\nu d\sigma \phi_{(n)}\,\mathsf{exp}\left(-\frac{i}{\epsilon}\int_0^\sigma d\lambda H_{\text{D}}\right)\frac{dS}{d\sigma}S^{-1}\,\mathsf{exp}\left(\frac{i}{\epsilon}\int_0^\sigma d\lambda H_{\text{D}}\right). \tag{15.16}$$

При $\epsilon \to 0$ подынтегральное выражение содержит быстро осциллирующий множитель и его легко оценить. Рассмотрим модуль компоненты решения $\psi'^j_{(n)}$, которая соответствует собственному состоянию гамильтониана H с энергией E_j, где $E_j \neq E_n$,

$$\left|\psi'^j_{(n)}\right| = \left|\phi^j_{(n)}\right| = \left|\sum_{k=1}^{\text{N}} \int_0^\nu d\sigma\,\mathsf{exp}\left(\frac{i}{\epsilon}\int_0^\sigma d\lambda(E_j - E_k)\right)\phi^k_{(n)}\left(\frac{dS}{d\nu}S^{-1}\right)_k{}^j\right|. \tag{15.17}$$

В сумме справа слагаемые с $E_k = E_j$ вклада не дают в силу равенства (15.12). При $E_k \neq E_j$ каждое слагаемое проинтегрируем по частям:

$$\begin{aligned}\frac{\epsilon}{i(E_j - E_k)}\,\mathsf{exp}\left(\frac{i}{\epsilon}\int_0^\sigma d\lambda(E_j - E_k)\right)\phi^k_{(n)}\left(\frac{dS}{d\nu}S^{-1}\right)_k{}^j\Bigg|_0^\nu - \\ - \frac{\epsilon}{i}\int_0^\nu d\sigma\,\mathsf{exp}\left(\frac{i}{\epsilon}\int_0^\sigma d\lambda(E_j - E_k)\right)\frac{1}{E_j - E_k}\frac{d}{d\sigma}\left[\phi^k_{(n)}\left(\frac{dS}{d\nu}S^{-1}\right)_k{}^j\right].\end{aligned} \tag{15.18}$$

По предположению подынтегральное выражение во втором слагаемом является дифференцируемой функцией и его снова можно проинтегрировать по частям. В результате получим, что оно имеет порядок ϵ^2, и им можно пренебречь. Модуль первого слагаемого, очевидно, ограничен. Таким образом получаем оценку

$$\left|\psi'^j_{(n)}(\nu,\epsilon)\right| = \frac{\mathsf{O}(\epsilon)}{\mathsf{min}\left|E_j(\sigma) - E_k(\sigma)\right|}, \quad \forall j \notin \Upsilon_n, \tag{15.19}$$

где минимум берется по всем k, для которых $E_k \neq E_j$, и всем $\sigma \in [0, \nu]$.

Теперь снова вернемся к выражению (15.18). Из оценки (15.19) вытекает, что $|\phi^k_{(n)}|$ для всех k при $E_k \neq E_n$ имеет порядок не ниже ϵ. Поэтому в сумме (15.17) все слагаемые с индексом $k \notin \Upsilon_n$ дают вклад не ниже ϵ^2, и ими можно пренебречь. Поэтому оценку (15.19) можно улучшить

$$\left|\psi'^j_{(n)}(\nu, \epsilon)\right| = \frac{\mathsf{O}(\epsilon)}{\min \left|E_j(\sigma) - E_n(\sigma)\right|}, \qquad \forall j \notin \Upsilon_n,$$

где минимум берется только по $\sigma \in [0, \nu]$.

Норма произвольного решения сохраняется во времени и равна единице. Следовательно,

$$1 - \sum_{j \in \Upsilon_n} |\psi'^j_{(n)}(\nu, \epsilon)|^2 = \sum_{j \notin \Upsilon_n} |\psi'^j_{(n)}(\nu, \epsilon)|^2.$$

Поскольку число уровней конечно, то отсюда вытекает оценка (15.14). □

Замечание. В теореме функция $\triangle E_n(\nu)$ для каждого ν равна константе и ее можно включить в $\mathsf{O}(\epsilon^2)$. Тем не менее мы выделили множитель $\triangle E_n$ с тем, чтобы показать, что предположение о том, что уровни энергии не пересекаются, является существенным. При пересечении уровней энергии знаменатель в (15.14) обращается в нуль и доказательство не проходит. В этом случае требуются дополнительные предположения о степени касания уровней энергии и дополнительное исследование. В своей оригинальной статье [?] Борн и Фок рассмотрели случай, когда спектр гамильтониана дискретен, но может быть неограничен. Неявно ими было сделано предположение о невырожденности спектра почти для всех моментов времени. Кроме того, допускалась возможность определенного пересечения уровней энергии с течением времени. Мы рассмотрели более простой конечномерный случай, когда уровни энергии не пересекаются. Это позволило упростить доказательство и выявить наиболее существенные моменты. Оценка (15.14) согласуется с оценкой, приведенной в [?]. □

Адиабатическая теорема утверждает, что, если в начальный момент времени система находилась в собственном состоянии гамильтониана, соответствующем уровню энергии $E_n(0)$, и этот уровень невырожден, то в адиабатическом пределе она будет оставаться в собственном состоянии $E_n(\nu)$ с точностью порядка ϵ^2 при конечных значениях параметра ν. Если уровень энергии E_n вырожден, то система будет находиться в одном из собственных состояний E_j, где $j \in \Upsilon_n$, с той же точностью. В следующем разделе мы увидим, что оценка (15.14) неулучшаема, а в процессе эволюции система может оказаться в любом из вырожденных состояний E_j, $j \in \Upsilon_n$, с вероятностью порядка единицы. Эти утверждения, естественно, не зависят от выбора базиса, который использовался при доказательстве адиабатической теоремы.

Рассмотрим теперь, как выглядит в адиабатическом пределе решение задачи Коши (15.4) в фиксированном базисе в невырожденном случае. Пусть $\varphi(\nu)$ – собственная функция гамильтониана $H(\nu)$, отвечающая невырожденному собственному значению энергии $E(\nu)$,

$$\varphi H = E\varphi, \qquad \forall \nu \in [0, \nu_0].$$

Эти собственные функции определены с точностью до фазового множителя, который может зависеть от ν. Пусть в начальный момент времени система находилась в собственном состоянии $\psi_0 = \varphi(0)$. В адиабатическом приближении она будет находиться в собственном состоянии, соответствующем энергии $E(\nu)$. Поскольку собственное состояние невырождено, то решение задачи Коши (15.4) может отличаться от φ не

более, чем на фазовый множитель. Поэтому будем искать решение в виде $\psi = \mathsf{e}^{i\Theta}\varphi$, где $\Theta(t)$ – неизвестная функция времени. Тогда из уравнения Шредингера следует уравнение на фазу

$$\dot{\Theta} = i(\dot{\varphi}, \varphi) - E. \tag{15.20}$$

Поскольку в начальный момент времени $\Theta(0) = 0$, то фаза имеет вид

$$\Theta(t) = i\int_0^t ds(\dot{\varphi}, \varphi) - \int_0^t dsE(\epsilon s) = i\int_0^\nu d\sigma \left(\frac{d\varphi}{d\nu}, \varphi\right) - \int_0^t dsE(\epsilon s). \tag{15.21}$$

Покажем, что фазу собственной функции φ всегда можно выбрать таким образом, что будет выполнено равенство

$$\left(\frac{d\varphi}{d\nu}, \varphi\right) = 0, \tag{15.22}$$

если $\nu \in [0, \infty)$. Действительно, пусть $\varphi = \mathsf{e}^{i\beta}\chi$, где функция $\beta(\nu)$ удовлетворяет уравнению

$$i\frac{d\beta}{d\nu} = \left(\frac{d\varphi}{d\nu}, \varphi\right) \tag{15.23}$$

с некоторым начальным условием, например, $\beta(0) = 0$. Тогда нетрудно проверить, что для новых собственных функций выполнено равенство $(d\chi/d\nu, \chi) = 0$. Поскольку уравнение (15.23) всегда имеет решение на полупрямой, то собственные функции φ гамильтониана всегда можно выбрать таким образом, что будет выполнено равенство (15.22).

Однако уравнение (15.23) может не иметь решения на окружности $\mathbb{S}^1$. Будем считать, что на окружности $\nu \in [0, 2\pi]$. Тогда необходимым условием существования решения является равенство

$$i\int_0^{2\pi} d\nu \left(\frac{d\varphi}{d\nu}, \varphi\right) = 2\pi m, \qquad m = 0, \pm 1, \pm 2, \ldots .$$

Ясно, что это условие в общем случае не выполняется. Поэтому уравнение (15.23) может не иметь решения на окружности. В этом случае первое слагаемое в выражении для фазы (15.21) устранить нельзя. По сути дела это и есть фаза Берри.

Решение задачи Коши на окружности $\nu \in \mathbb{S}^1$ означает наличие машины времени. Эти решения можно отбросить как нефизические. Однако Берри предложил другую схему рассуждений, которая будет рассмотрена в разделе 15.2.1.

15.1.1 Двухуровневая система

В настоящем разделе в качестве примера мы рассмотрим двухуровневую квантово-механическую систему, для которой уравнение Шредингера решается явно. Будет показано, что оценка, данная в адиабатической теореме, является неулучшаемой.

Чтобы упростить задачу, поступим следующим образом. Вместо того, чтобы задать исходный гамильтониан в фиксированном базисе, а затем его диагонализировать, мы зададим диагональную матрицу $H_{\text{д}}$ и унитарную матрицу S, которые определяют исходный гамильтониан $H = S^{-1}H_{\text{д}}S$. Пусть диагонализированный гамильтониан имеет вид

$$H_{\text{д}} = \begin{pmatrix} E_1(\nu) & 0 \\ 0 & E_2(\nu) \end{pmatrix},$$

где $E_{1,2}(\nu)$ – некоторые заданные функции. Унитарную матрицу S в (15.10) выберем в виде

$$S = \begin{pmatrix} \cos\frac{\alpha}{2} & i\sin\frac{\alpha}{2} \\ i\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix},$$

где $\alpha(\nu) \in \mathbb{R}$ – также некоторая заданная функция. Следовательно, исходный гамильтониан задачи имеет вид

$$H = S^{-1}H_{\text{D}}S = \begin{pmatrix} E_1\cos^2\frac{\alpha}{2} + E_2\sin^2\frac{\alpha}{2} & -\frac{i}{2}(E_2 - E_1)\sin\alpha \\ \frac{i}{2}(E_2 - E_1)\sin\alpha & E_1\sin^2\frac{\alpha}{2} + E_2\cos^2\frac{\alpha}{2} \end{pmatrix}$$

и зависит от трех, пока произвольных, функций от параметра ν, которые предполагаются достаточно гладкими.

Будем решать уравнение Шредингера в базисе (15.6), в котором гамильтониан имеет вид (20.62). Простые вычисления приводят к гамильтониану

$$H' = \begin{pmatrix} E_1(\nu) & \frac{\dot\alpha}{2} \\ \frac{\dot\alpha}{2} & E_2(\nu) \end{pmatrix},$$

где точка обозначает дифференцирование по времени t. Ищем решение уравнения Шредингера (15.5) в виде строки

$$\psi' = \left(\exp\left(-i\int_0^t ds E_1\right)\phi,\ \exp\left(-i\int_0^t ds E_2\right)\chi\right),$$

где $\phi(t)$ и $\chi(t)$ – неизвестные функции. Подстановка в уравнение Шредингера приводит к системе уравнений для компонент:

$$\begin{aligned} i\dot\phi &= \frac{\dot\alpha}{2}\exp\left(-i\int_0^t ds(E_2 - E_1)\right)\chi, \\ i\dot\chi &= \frac{\dot\alpha}{2}\exp\left(i\int_0^t ds(E_2 - E_1)\right)\phi. \end{aligned} \tag{15.24}$$

При $\dot\alpha \neq 0$ из первого уравнения следует равенство

$$\chi = \frac{2i}{\dot\alpha}\exp\left(i\int_0^t ds(E_2 - E_1)\right)\dot\phi. \tag{15.25}$$

Продифференцируем это равенство по времени и подставим во второе уравнение. В результате получим уравнение второго порядка для ϕ:

$$\ddot\phi + \left(i(E_2 - E_1) - \frac{\ddot\alpha}{\dot\alpha}\right)\dot\phi + \left(\frac{\dot\alpha}{2}\right)^2\phi = 0. \tag{15.26}$$

Для того, чтобы решить это уравнение в явном виде зафиксируем произвольные функции, которые входят в задачу:

$$\begin{aligned} E_1 &= E_1^{(0)} + \epsilon t, \quad & E_{1(0)} &= \text{const}, \\ E_2 &= E_2^{(0)} + \epsilon t, \quad & E_{2(0)} &= \text{const}, \\ \alpha &= 2\epsilon t. \end{aligned} \tag{15.27}$$

Тогда уравнение (15.26) примет простой вид

$$\ddot\phi + 2i\triangle E\dot\phi + \epsilon^2\phi = 0, \tag{15.28}$$

где $\triangle E = E_2^{(0)} - E_1^{(0)}$ – расстояние между уровнями энергии. Общее решение этого уравнения зависит от двух постоянных интегрирования $C_{1,2}$ и имеет вид

$$\phi = \mathsf{e}^{-i\triangle E t}\left(C_1\,\mathsf{e}^{i\omega_\epsilon t} + C_2\,\mathsf{e}^{-i\omega_\epsilon t}\right),$$

где

$$\omega_\epsilon := \sqrt{\triangle E^2 + \epsilon^2}.$$

$C_{1,2}$ – постоянные интегрирования. Компонента χ определяется по формуле (15.25). Предположим, что в начальный момент времени система находилась в состоянии E_1, т.е.

$$\phi(0) = 1, \qquad \chi(0) = 0. \tag{15.29}$$

Простые вычисления дают решение задачи Коши для системы уравнений (15.24):

$$\begin{aligned} \phi &= \mathsf{e}^{-i\triangle E t}\left[\cos\,(\omega_\epsilon t) + \frac{i\triangle E}{\omega_\epsilon}\sin\,(\omega_\epsilon t)\right], \\ \chi &= \mathsf{e}^{i\triangle E t}\left[-\frac{i\epsilon}{\omega_\epsilon}\sin\,(\omega_\epsilon t)\right]. \end{aligned} \tag{15.30}$$

Выпишем также компоненты соответствующего вектора состояния

$$\begin{aligned} \psi'^1 &= \mathsf{e}^{-i\left(\frac{\nu^2}{2\epsilon}+E_1^{(0)}\frac{\nu}{\epsilon}-\triangle E\frac{\nu}{\epsilon}\right)}\left[\cos\frac{\omega_\epsilon\nu}{\epsilon} + \frac{i\triangle E}{\omega_\epsilon}\sin\frac{\omega_\epsilon\nu}{\epsilon}\right], \\ \psi'^2 &= \mathsf{e}^{-i\left(\frac{\nu^2}{2\epsilon}+E_2^{(0)}\frac{\nu}{\epsilon}+\triangle E\frac{\nu}{\epsilon}\right)}\left[-\frac{i\epsilon}{\omega_\epsilon}\sin\frac{\omega_\epsilon\nu}{\epsilon}\right], \end{aligned} \tag{15.31}$$

где мы перешли от t, ϵ к переменным ν, ϵ. Отсюда следует, что адиабатический предел для самого вектора состояний неопределен, т.к. его фаза стремится к бесконечности. Однако оценку квадрата модуля компонент можно дать. Для решения (15.31) следует оценка

$$1 - |\psi'^1(\nu,\epsilon)|^2 = \frac{\mathsf{O}(\epsilon^2)}{(\triangle E)^2}, \qquad |\psi'^2(\nu,\epsilon)|^2 = \frac{\mathsf{O}(\epsilon^2)}{(\triangle E)^2},$$

которая совпадает с оценкой в адиабатической теореме. Отсюда следует, что данная оценка неулучшаема.

Теперь рассмотрим случай вырожденных состояний, $E_1 = E_2$, при заданных ранее функциях (15.27). В этом случае уравнение (15.28) сводится к уравнению для свободного осциллятора:

$$\ddot{\phi} + \epsilon^2\phi = 0,$$

и легко интегрируется. Выпишем соответствующее решение задачи Коши (15.29) для компонент вектора состояния:

$$\begin{aligned} \psi'^1 &= \quad \mathsf{e}^{-i\left(\frac{\nu^2}{2\epsilon}-E_1^{(0)}\frac{\nu}{\epsilon}\right)}\cos\nu, \\ \psi'^2 &= -i\,\mathsf{e}^{-i\left(\frac{\nu^2}{2\epsilon}-E_1^{(0)}\frac{\nu}{\epsilon}\right)}\sin\nu. \end{aligned}$$

Мы снова видим, что адиабатический предел у вектора состояния отсутствует. Однако квадраты модулей компонент хорошо определены:

$$|\psi'^1|^2 = \cos^2\nu, \qquad |\psi'^2|^2 = \sin^2\nu.$$

Отсюда следует, что по мере увеличения параметра ν вектор состояния ψ' осциллирует между вырожденными состояниями. Это показывает, что, если в начальный момент времени система находится в одном из вырожденных состояний, то в процессе эволюции она может оказаться в любом из вырожденных состояний с вероятностью порядка единицы.

В настоящее время адиабатическая теорема доказана для гамильтонианов с бесконечным дискретным спектром. Кроме того адиабатическая теорема справедлива для изолированных уровней, энергия которых отделена от остального спектра, который может быть как дискретным, так и непрерывным [?].

15.2 Фаза Берри

Перейдем к задаче, которую рассмотрел М. Берри [?], в ее простейшем варианте.

Пусть гильбертово пространство конечномерно и гамильтониан $H = H(\lambda)$ достаточно гладко зависит от точки некоторого многообразия $\lambda \in \mathbb{M}$ размерности $\dim \mathbb{M} = n$. Если на $\mathbb{M}$ выбрать координатную окрестность $\mathbb{U} \subset \mathbb{M}$, то гамильтониан будет зависеть от n параметров λ^k, $k = 1, \dots, n$, (координат точки λ). Будем считать, что положение точки λ на $\mathbb{M}$ зависит от времени t некоторым наперед заданным образом, т.е. гамильтониан зависит от точки некоторой кривой $\lambda(t)$, $t \in [0, t_0]$. Мы будем решать уравнение Шредингера в адиабатическом приближении, т.е. величина t_0 должна быть достаточно велика. Предположим также, что гамильтониан зависит от времени только через точку $\lambda(t) \in \mathbb{M}$.

15.2.1 Абелев случай: невырожденное состояние

Рассмотрим задачу на собственные значения

$$\phi H = E\phi, \qquad E = \mathsf{const},$$

где $\phi \in \mathbb{H}$ при всех $\lambda \in \mathbb{M}$. Предположим, что существует невырожденное собственное значение энергии E и соответствующее собственное состояние ϕ, которые достаточно гладко зависят от $\lambda \in \mathbb{M}$. Не ограничивая общности, предположим, что собственная функция ϕ нормирована на единицу, $(\phi, \phi) = 1$. Тогда она единственна с точностью до умножения на фазовый множитель, который может зависеть от λ. Зафиксируем этот фазовый множитель каким либо образом.

Теперь будем решать задачу Коши для уравнения Шредингера (15.1) с начальным условием

$$\psi|_{t=0} = \phi_0, \tag{15.32}$$

где $\phi_0 := \phi\big(\lambda(0)\big)$. В адиабатическом приближении квантовая система в процессе эволюции будет оставаться в собственном состоянии, соответствующем уровню энергии $E(\lambda)$. Поэтому ищем решение в виде

$$\psi = \mathsf{e}^{i\Theta}\phi,$$

где $\Theta = \Theta(t)$ – некоторая неизвестная функция от времени. Тогда из уравнения Шредингера следует уравнение на фазу (15.20) с начальным условием $\Theta|_{t=0} = 0$. Поскольку $\dot\phi = \dot\lambda^k \partial_k \phi$, то решение задачи Коши для уравнения (15.20) имеет вид

$$\Theta = \int_0^t dt \dot\lambda^k A_k - \int_0^t ds E\big(\lambda(s)\big) = \int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k - \int_0^t ds E\big(\lambda(s)\big), \tag{15.33}$$

где введено обозначение

$$A_k(\lambda) := i(\phi, \partial_k \phi) \tag{15.34}$$

и интеграл по λ берется вдоль кривой $\lambda(t)$.

Таким образом, интеграл (15.33) в адиабатическом приближении дает решение задачи Коши для уравнения Шредингера (15.1) с начальным условием (15.32). Первое слагаемое в (15.33) называется *геометрической фазой* или *фазой Берри*, а второе – *динамической фазой*.

Заметим, что компоненты (15.34) вещественны вследствие нормировки волновой функции. Действительно, дифференцируя условие нормировки $(\phi, \phi) = 1$, получаем равенство

$$(\partial_k \phi, \phi) + (\phi, \partial_k \phi) = (\phi, \partial_k \phi)^\dagger + (\phi, \partial_k \phi) = 0.$$

Отсюда вытекает вещественность компонент (15.34) и, следовательно, фазы Берри.

Теперь рассмотрим множество замкнутых кривых $\gamma \in \Omega(\mathbb{M}, \lambda_0)$ на многообразии параметров $\mathbb{M}$ с началом и концом в точке $\lambda_0 \in \mathbb{M}$. Тогда для полного изменения фазы волновой функции получаем ответ

$$\Theta = \Theta_{\text{в}} - \int_0^{t_0} dt E\Big(\lambda(t)\Big),$$

где

$$\Theta_{\text{в}} = \oint_\gamma d\lambda^k A_k. \tag{15.35}$$

Динамическая часть фазы волновой функции расходится. Однако в экспериментах наблюдается разность фаз двух векторов состояний с одинаковой динамической фазой, которая определяется фазой Берри. Поэтому рассмотрим фазу Берри подробнее.

Заметим, что выражение для фазы Берри не зависит от параметризации кривой $\gamma(t)$. Это значит, что переход к адиабатическому пределу в уравнении Шредингера влияет на фазу Берри только через компоненты локальной формы связности (15.34).

В таком виде можно дать геометрическую интерпретацию фазе Берри $\Theta_{\text{в}}$, которая определяется первым слагаемым в полученном выражении. А именно, мы имеем главное расслоение $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(1)\Big)$, базой которого является многообразие параметров $\lambda \in \mathbb{M}$, а структурной группой – группа $\mathbb{U}(1)$ (фаза вектора состояния $\mathsf{e}^{i\Theta}$). Вектор гильбертова пространства $\phi \in \mathbb{H}$ представляет собой локальное сечение ассоциированного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}, \mathbb{U}(1), \mathbb{P}\Big)$, типичным слоем которого является гильбертово пространство $\mathbb{H}$.

Рассмотрим изменение локального сечения ассоциированного расслоения, которое вызвано умножением вектора гильбертова пространства на фазовый множитель (вертикальный автоморфизм),

$$\phi' = \mathsf{e}^{ia}\phi,$$

где $a = a(\lambda) \in \mathcal{C}^2(\mathbb{M})$ – произвольная дважды дифференцируемая функция. Тогда компоненты (15.34) преобразуются по-правилу

$$A'_k = A_k - \partial_k a.$$

Сравнивая это правило с преобразованием компонент локальной формы связности, заключаем, что поля $A_k(\lambda)$ можно интерпретировать, как компоненты локальной формы связности для группы $\mathbb{U}(1)$. Другими словами, $A_k(\lambda)$ – это калибровочное поле для одномерной унитарной группы $\mathbb{U}(1)$. Если база ассоциированного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}, \mathbb{U}(1), \mathbb{P}\Big)$ покрыта некоторым семейством карт, $\mathbb{M} = \cup_j \mathbb{U}_j$, то множество

сечений, заданных в каждой координатной окрестности $\mathbb{U}_j$, определяет семейство локальных форм связности на главном расслоении $\mathbb{P}\big(\mathbb{M},\pi,\mathbb{U}(1)\big)$. Семейство локальных форм связности $d\lambda^k A_k$ определяет единственную с точностью до изоморфизма связность на $\mathbb{P}$ (теорема 14.1.1).

Вспомним выражение для элемента группы голономии в виде упорядоченной P-экспоненты (14.56). В рассматриваемом случае группа $\mathbb{U}(1)$ абелева и P-экспонента совпадает с обычной экспонентой. Поэтому фаза Берри (15.35) определяет элемент $\mathsf{e}^{i\Theta_{\text{в}}}$ группы голономии $\Phi(\lambda_0,e)\subset\mathbb{U}(1)$ главного расслоения в точке $(\lambda_0,e)\in\mathbb{P}$, соответствующей нулевому сечению $\mathbb{M}\ni\lambda\mapsto(\lambda,e)\in\mathbb{P}$, где $\lambda_0:=\lambda(0)$ и e – единица структурной группы $\mathbb{U}(1)$. Сечение является нулевым, поскольку в начальный момент времени фаза Берри равна нулю, $\Theta_{\text{в}}|_{t=0}=0$. Локальная форма связности $d\lambda^k A_k$ также соответствует нулевому сечению.

Если база $\mathbb{M}$ односвязна, то выражение для фазы Берри (15.35) можно переписать в виде поверхностного интеграла от компонент локальной формы кривизны. Используя формулу Стокса, получаем следующее выражение

$$\Theta_{\text{в}}=\frac{1}{2}\iint_S d\lambda^k\wedge d\lambda^l F_{kl}, \tag{15.36}$$

где S – поверхность в $\mathbb{M}$ с границей $\gamma\in\Omega(\mathbb{M},\lambda_0)$ и $F_{kl}=\partial_k A_l-\partial_l A_k$ – компоненты локальной формы кривизны (напряженности калибровочного поля). Если база $\mathbb{M}$ не является односвязной, то выражение для фазы Берри в виде поверхностного интеграла (15.36) имеет место только для тех замкнутых путей, которые стягиваются в точку.

15.2.2 Частица со спином $1/2$ в магнитном поле

В качестве примера вычислим фазу Берри для частицы со спином $1/2$, находящейся во внешнем однородном магнитном поле. В нерелятивистской квантовой механике частица со спином $1/2$ описывается двухкомпонентной волновой функцией

$$\psi=(\psi_+,\psi_-).$$

Будем считать, что она находится в евклидовом пространстве $\mathbb{R}^3$ с заданным однородным магнитным полем. Пусть напряженность магнитного поля $H^k(t)$, $k=1,2,3$, не зависит от точки пространства, но меняется со временем t некоторым заданным образом. Кроме того, для простоты, пренебрежем кинетической энергией частицы и будем считать, что другие поля отсутствуют. В этом случае гильбертово пространство $\mathbb{H}$ двумерно, и гамильтониан частицы состоит из одного слагаемого – взаимодействия магнитного момента частицы с внешним магнитным полем (см., например, [**?**, **?**]),

$$H=-\mu H^k\sigma_k^{\text{т}},$$

где $\sigma_k^{\text{т}}$ – транспонированные матрицы Паули (поскольку в наших обозначениях они действуют справа) и μ – магнетон (размерная постоянная, равная отношению магнитного момента частицы к ее спину). Чтобы привести гамильтониан к виду, который был рассмотрен ранее, ведем новые переменные $\lambda^k=-\mu H^k$. Тогда гамильтониан примет вид

$$H=\lambda^k\sigma_k^{\text{т}}=\begin{pmatrix}\lambda^3 & \lambda^+\\ \lambda^- & -\lambda^3\end{pmatrix}, \tag{15.37}$$

где $\lambda^\pm=\lambda^1\pm i\lambda^2$.

Собственные значения гамильтониана (15.37) находятся из уравнения

$$\det(H - E\mathbb{1}) = 0,$$

которое имеет два вещественных решения

$$E_\pm = \pm|\lambda|, \tag{15.38}$$

где

$$|\lambda| = \sqrt{(\lambda^1)^2 + (\lambda^2)^2 + (\lambda^3)^2}$$

– длина вектора $\lambda = \{\lambda^k\} \in \mathbb{R}^3$. Нетрудно проверить, что уравнение на собственные функции,

$$\phi_\pm H = E_\pm \phi_\pm,$$

имеет два решения

$$\phi_\pm = \frac{1}{\sqrt{2|\lambda|}} \left(\pm \frac{\lambda^-}{\sqrt{|\lambda| \mp \lambda^3}}, \sqrt{|\lambda| \mp \lambda^3} \right). \tag{15.39}$$

В полученном выражении множитель выбран таким образом, что собственные функции нормированы на единицу,

$$(\phi_\pm, \phi_\pm) = 1.$$

Таким образом, гамильтониан (15.37) частицы со спином $1/2$, находящейся во внешнем однородном магнитном поле, имеет два невырожденных собственных состояния (15.39), соответствующих уровням энергии (15.38).

Для дальнейших вычислений в пространстве параметров $\lambda \in \mathbb{R}^3$ удобно ввести сферические координаты $|\lambda|, \theta, \varphi$:

$$\begin{aligned} \lambda^1 &= |\lambda| \sin\theta \cos\varphi, \\ \lambda^2 &= |\lambda| \sin\theta \sin\varphi, \\ \lambda^3 &= |\lambda| \cos\theta. \end{aligned}$$

Тогда собственные функции примут вид

$$\phi_+ = \left(\cos\frac{\theta}{2}\, e^{-i\varphi}, \sin\frac{\theta}{2} \right), \qquad \phi_- = \left(-\sin\frac{\theta}{2}\, e^{-i\varphi}, \cos\frac{\theta}{2} \right).$$

Замечание. Собственные функции гамильтониана в сферической системе координат не определены при $\theta = 0, \pi$. Это – координатная сингулярность, присущая сферической системе координат. В декартовой системе координат собственные функции (15.39) никакой сингулярности не имеют. □

Допустим, что экспериментатор, наблюдающий за частицей, достаточно гладко меняет однородное магнитное со временем. То есть параметры $\lambda^k(t)$, от которых зависит гамильтониан, достаточно гладко зависят от времени. Предположим также, что в начальный момент времени $t = 0$ частица находилась в состоянии ϕ_+. Соответствующее решение уравнения Шредингера (15.1) в адиабатическом приближении имеет вид

$$\psi = e^{i\Theta} \phi_+,$$

где фаза Θ удовлетворяет уравнению (15.20). Нетрудно вычислить компоненты локальной формы связности $A_k = i(\phi_+, \partial_k \phi_+)$ для собственного состояния ϕ_+,

$$A_{|\lambda|} = 0, \quad A_\theta = 0, \quad A_\varphi = \cos^2\frac{\theta}{2}.$$

Соответствующая локальная форма кривизны имеет только две отличные от нуля компоненты:

$$F_{\theta\varphi} = -F_{\varphi\theta} = -\frac{1}{2}\sin\theta.$$

Теперь вычислим фазу Берри для замкнутой кривой в пространстве параметров $\gamma = \{\lambda^k(t)\} \in \mathbb{M}$,

$$\begin{aligned}\Theta_{\text{в}} &= \oint_\gamma d\lambda^k A_k = \frac{1}{2}\iint_S d\lambda^k \wedge d\lambda^l F_{kl} = \\ &= \iint_S d\theta \wedge d\varphi F_{\theta\varphi} = -\frac{1}{2}\iint_S d\theta \wedge d\varphi \sin\theta = -\frac{1}{2}\Omega(\gamma),\end{aligned} \tag{15.40}$$

где S – поверхность в $\mathbb{R}^3$ с границей γ и $\Omega(\gamma)$ – телесный угол, который занимает поверхность S, если смотреть из начала координат.

Если в начальный момент времени частица находилась в состоянии ϕ_-, то вычисления проводятся аналогично. В этом случае

$$A_{|\lambda|} = 0, \quad A_\theta = 0, \quad A_\varphi = \sin^2\frac{\theta}{2},$$

и компоненты локальной формы кривизны отличаются знаком:

$$F_{\theta\varphi} = -F_{\varphi\theta} = \frac{1}{2}\sin\theta.$$

Следовательно, фаза Берри также отличается только знаком.

Таким образом, если в начальный момент времени частица находилась в одном из состояний $\phi_\pm$, то после изменения однородного магнитного поля вдоль замкнутой кривой γ, ее волновая функция изменится на фазовый множитель, геометрическая часть которого равна

$$\Theta_{\text{в}\pm} = \mp\frac{1}{2}\Omega(\gamma), \tag{15.41}$$

где $\Omega(\gamma)$ – телесный угол, под которым виден замкнутый контур γ из начала координат. Этот результат имеет место в адиабатическом приближении, когда параметры $\gamma(t)$ медленно меняются со временем.

Выражение для фазы Берри (15.41) было подтверждено экспериментально [**?**] для рассеяния поляризованных нейтронов в спиральном магнитном поле.

В рассмотренном примере однородное магнитное поле может иметь произвольное направление и величину. Следовательно, база $\mathbb{M}$ главного расслоения $\mathbb{P}\big(\mathbb{M}, \pi, \mathbb{U}(1)\big)$ совпадает с евклидовым пространством, $\mathbb{M} = \mathbb{R}^3$. Поэтому, согласно теореме 13.1.1, главное расслоение $\mathbb{P}$ тривиально, $\mathbb{P} \simeq \mathbb{R}^3 \times \mathbb{U}(1)$. В случае фазы Берри связность на этом расслоении определяется сечением ассоциированного расслоения, например, ϕ_+, которое находится путем решения уравнения Шредингера. Это сечение (15.39), как нетрудно проверить, имеет особенность на положительной полуоси $\lambda^3 \geq 0$. Компоненты локальной формы связности (**??**) относительно декартовой системы координат

имеют вид

$$A_1 = \frac{\partial\varphi}{\partial\lambda^1}A_\varphi = -\frac{\sin\varphi\cos\frac{\theta}{2}}{2|\lambda|\sin\frac{\theta}{2}},$$
$$A_2 = \frac{\partial\varphi}{\partial\lambda^2}A_\varphi = \frac{\cos\varphi\cos\frac{\theta}{2}}{2|\lambda|\sin\frac{\theta}{2}}, \tag{15.42}$$
$$A_3 = \frac{\partial\varphi}{\partial\lambda_3}A_\varphi = 0.$$

Здесь мы вынуждены перейти в декартову систему координат, поскольку сферическая система координат сингулярна на оси λ^3 и непригодна для исследования особенностей, которые здесь расположены. Как видим, компоненты локальной формы связности имеют особенность на положительной полуоси $\lambda^3 \geq 0$ как и вектор ϕ_+. Теперь вычислим компоненты локальной формы тензора кривизны. У нее отличны от нуля все компоненты:

$$F_{12} = -F_{21} = -\frac{\cos\theta}{2|\lambda|^2},$$
$$F_{13} = -F_{31} = \frac{\sin\theta\sin\varphi}{2|\lambda|^2},$$
$$F_{23} = -F_{32} = -\frac{\sin\theta\cos\varphi}{2|\lambda|^2}.$$

Наконец, вычислим квадрат тензора кривизны, который является инвариантом,

$$F^2 = 2\Big(F_{12}^2 + F_{13}^2 + F_{23}^2\Big) = \frac{1}{2|\lambda|^4}.$$

Таким образом, форма кривизны сингулярна только в начале координат.

Вернемся к нашему главному расслоению $\mathbb{R}^3 \times \mathbb{U}(1)$. Локальная форма связности (15.42) на нем неопределена, т.к. имеет особенность на полуоси $\lambda^3 \geq 0$, которую мы обозначим $\{\lambda^3_+\}$. Поэтому, чтобы построить главное расслоение с заданной связностью, мы вынуждены удалить прообраз $\pi^{-1}\Big(\{\lambda^3_+\}\Big)$, где $\pi : \ \mathbb{R}^3 \times \mathbb{U}(1) \to \mathbb{R}^3$ – естественная проекция. В результате получаем тривиальное главное расслоение $\Big(\mathbb{R}^3 \setminus \{\lambda^3_+\}\Big) \times \mathbb{U}(1)$, которое является подрасслоением исходного расслоения. На этом главном расслоении локальная форма (15.42) бесконечно дифференцируема.

Можно рассуждать по-другому. Поскольку магнитное поле является внешним, то мы вправе предположить, что оно меняется, например, в полупространстве $\mathbb{R}^3_+$, определяемом условием $\lambda_1 > 0$. Поскольку полупространство $\mathbb{R}^3_+$ диффеоморфно всему евклидову пространству $\mathbb{R}^3$, то соответствующее главное расслоение тривиально: $\mathbb{P} \approx \mathbb{R}^3_+ \times \mathbb{U}(1)$. В этом случае никаких вопросов в определением связности вообще не возникает, т.к. локальная форма связности (15.42) гладкая. При этом выражение для фазы Берри (15.41) останется прежним.

Таким образом, фаза Берри является не топологическим понятием, а геометрическим, т.к. топология главного расслоения тривиальна. Она обязана своим происхождением нетривиальной связности, которая определяется сечениями ассоциированного расслоения.

15.2.3 Неабелев случай: вырожденное состояние

Понятие фазы Берри было обобщено на случай, когда уровни энергии гамильтониана вырождены [?]. В этом случае при решении уравнения Шредингера возникает

главное расслоение $\mathbb{P}\Big(\mathbb{M},\pi,\mathbb{U}(r)\Big)$ со структурной группой $\mathbb{U}(r)$, где r – количество независимых собственных функций, соответствующих вырожденному уровню энергии E. Опишем эту конструкцию подробно.

Предположим, что гамильтониан квантовой системы зависит от точки некоторого многообразия $\lambda(t)\in\mathbb{M}$, как и ранее. Пусть E – вырожденное собственное значение гамильтониана H, которому соответствуют r независимых собственных функций ϕ^a, $a=1,\ldots,r$,

$$\phi^a H = E\phi^a$$

для всех моментов времени. Мы предполагаем, что $E(\lambda)$ и $\phi^a(\lambda)$ являются достаточно гладкими функциями от точки многообразия λ, и число собственных функций r не меняется со временем.

Собственные функции можно выбрать ортонормированными,

$$(\phi^a,\phi_b)=\delta^a_b,$$

где δ^a_b – символ Кронекера. Будем искать решение задачи Коши ψ^a для уравнения Шредингера (15.1) с начальным условием

$$\psi^a(0)=\psi^a_0=\phi^a\Big(\lambda(0)\Big).$$

То есть в начальный момент времени система находится в одном из собственных состояний ϕ^a. В адиабатическом приближении решение ψ^a для всех моментов времени является собственной функцией гамильтониана $H(\lambda)$, соответствующей значению энергии $E(\lambda)$. Поэтому его можно разложить по собственным функциям вырожденного состояния

$$\psi^a=\phi^b U^{-1}{}^a_b, \tag{15.43}$$

где $U(\lambda)\in\mathbb{U}(r)$ – некоторая унитарная матрица, которая достаточно гладко зависит от точки $\lambda\in\mathbb{M}$.

Унитарность матрицы U обусловлена следующим обстоятельством. Рассмотрим решения ψ^a для всех значений индекса $a=1,\ldots,r$. Дифференцируя скалярное произведение (ψ^a,ψ_b) по времени и используя уравнение Шредингера, получаем уравнение

$$\frac{\partial}{\partial t}(\psi^a,\psi_b)=-i(\psi^a H,\psi_b)+i(\psi^a,H\psi_b)=0.$$

Последнее равенство следует из самосопряженности гамильтониана. Отсюда следует, что, если в начальный момент времени векторы $\psi^a_0=\phi^a\Big(\lambda(0)\Big)$ ортонормированы, то соответствующие решения уравнения Шредингера останутся таковыми и во все последующие моменты времени. Поэтому матрица U в разложении (15.43) унитарна.

Для искомого решения (15.43) уравнение Шредингера сводится к уравнению

$$i\phi^c\dot{U}^{-1}{}^b_c+i\dot{\phi}^c U^{-1}{}^b_c=\phi^c H U^{-1}{}^b_c.$$

Возьмем скалярное произведение левой и правой части с ϕ_a. В результате получаем уравнение на унитарную матрицу

$$\dot{U}^{-1}{}^b_a=\dot{\lambda}^k A_{ka}{}^c U^{-1}{}^b_c-iEU^{-1}{}^b_a, \tag{15.44}$$

где введено обозначение

$$A_{ka}{}^c:=-(\partial_k\phi^c,\phi_a). \tag{15.45}$$

Из условия ортонормированности собственных функций ϕ^a следует антиунитарность компонент ${A_{ka}}^b$ для всех $k = 1, \ldots, n$, если индексы a, b рассматриваются, как матричные. Действительно, дифференцируя условие ортонормированности $(\phi^b, \phi_a) = \delta^b_a$, получаем равенство

$$(\partial_k \phi^b, \phi_a) + (\phi^b, \partial_k \phi_a) = (\phi^a, \partial_k \phi_b)^\dagger + (\phi^b, \partial_k \phi_a) = 0.$$

То есть матрицы A_k антиунитарны и поэтому принадлежат алгебре Ли $\mathfrak{u}(r)$. Следовательно, матрицы A_k определяют 1-формы в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ со значениями в алгебре Ли, как и компоненты локальной формы связности.

Начальное условие для унитарной матрицы имеет вид

$$U^{-1}{}^b_a|_{t=0} = \delta^b_a.$$

Решение задачи Коши для уравнения (15.44) можно записать в виде упорядоченного P-произведения

$$\begin{aligned} U^{-1}(t) &= \mathsf{P}\,\mathsf{exp}\,\left(\int_0^t ds \dot{\lambda}^k(s) A_k(s) - i\int_0^t ds E\Big(\lambda(s)\Big)\right) = \\ &= \mathsf{P}\,\mathsf{exp}\,\left(\int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k\right) \times\, \mathsf{exp}\,\left(-i\int_0^t ds E\Big(\lambda(s)\Big)\right), \end{aligned} \tag{15.46}$$

где мы, для краткости, опустили матричные индексы.

Первый сомножитель является обобщением фазы Берри на случай вырожденных состояний, а второй сомножитель – это динамическая фаза. Динамическая фаза имеет тот же вид, что и в случае невырожденного состояния.

Первый сомножитель в решении (15.46) представляет собой унитарную матрицу Вилчека–Зи

$$U^{-1}_{\mathrm{wz}} = \mathsf{P}\,\mathsf{exp}\,\left(\int_{\lambda(0)}^{\lambda(t)} d\lambda^k A_k\right), \tag{15.47}$$

которой можно дать следующую геометрическую интерпретацию. Мы имеем главное расслоение $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(r)\Big)$ со структурной группой $\mathbb{U}(r)$ (преобразование (15.43)). Набор собственных функций ϕ^a представляет собой сечение ассоциированного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}^r, \mathbb{U}(r), \mathbb{P}\Big)$, типичным слоем которого является тензорное произведение гильбертовых пространств

$$\mathbb{H}^r := \underbrace{\mathbb{H} \otimes \ldots \otimes \mathbb{H}}_{r}.$$

При вертикальном автоморфизме, который задан унитарной матрицей $U(\lambda) \in \mathbb{U}(r)$,

$$\phi'^a = \phi^b {U^{-1}}^a_b, \qquad \phi'_a = {U_a}^b \phi_b,$$

поля (15.45) преобразуются по правилу

$$A'_k = U A_k U^{-1} + \partial_k U U^{-1}, \tag{15.48}$$

где мы опустили матричные индексы. Отсюда следует, что поля A_k можно интерпретировать, как компоненты локальной формы связности или поля Янга–Миллса. Совокупность этих компонент, заданная на координатном покрытии базы $\mathbb{M}$, однозначно задает связность на главном расслоении $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(r)\Big)$.

Если путь замкнут, $\gamma \in \Omega(\mathbb{M}, \lambda_0)$, то унитарная матрица Вилчека–Зи (15.47) представляет собой элемент группы голономии $U_{\rm wz}^{-1} \in \Phi(\lambda_0, e)$, в точке $(\lambda_0, e) \in \mathbb{P}$, соответствующей нулевому сечению $\mathbb{M} \ni \lambda \mapsto (\lambda, e) \in \mathbb{P}$, где $\lambda_0 := \lambda(0)$ и e – единица структурной группы $\mathbb{U}(r)$.

Таким образом, в случае вырожденного уровня энергии гамильтониана возникает главное расслоение $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(r)\Big)$. В рассматриваемом случае базой $\mathbb{M}$ является многообразие параметров $\lambda \in \mathbb{M}$, от точки которого зависит гамильтониан. Мы предполагаем, что это многообразие конечномерно. Структурной группой является унитарная группа $\mathbb{U}(r)$, которая также конечномерна. Связность на главном расслоении определяется сечениями ассоциированного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{H}^r, \mathbb{U}(r), \mathbb{P}\Big)$. В общем случае типичным слоем ассоциированного расслоения может быть бесконечномерное гильбертово пространство $\mathbb{H}^r$. В настоящей монографии мы не рассматриваем бесконечномерных многообразий и расслоений, чтобы избежать возникающих при этом трудностей [**?**]. Однако в данном случае все, что нужно, это формула преобразования компонент локальной формы связности (15.48), которую легко проверить в каждом конкретном случае. Если база $\mathbb{M}$ не покрывается одной картой, то состояние квантовой системы задается семейством локальных сечений на координатном покрытии базы. Оно определяет семейство локальных форм связности (15.45). В свою очередь семейство локальных форм связности однозначно с точностью до изоморфизма определяет связность на главном расслоении $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(r)\Big)$.

Опять мы видим, что главные и ассоциированные расслоения могут быть тривиальными или нет, это зависит от рассматриваемой задачи. Связность на главном расслоении $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(r)\Big)$ может быть нетривиальна и приводить к нетривиальной матрице Вилчека–Зи (15.47), которая описывает параллельный перенос слоев вдоль пути на базе $\lambda(t) \in \mathbb{M}$, даже для тривиальных расслоений. Это говорит о ее геометрическом, а не топологическом происхождении. Для замкнутых путей $\gamma \in \Omega(\mathbb{M}, \lambda_0)$ с началом и концом в точке $\lambda_0 \in \mathbb{M}$ матрица Вилчека–Зи определяет элемент группы голономии $U_{\rm wz} \in \Phi(\lambda_0, e) \subset \mathbb{U}(r)$.

При рассмотрении фазы Берри и матрицы Вилчека–Зи мы, для простоты, предположили, что гильбертово пространство квантовой системы конечномерно. Это предположение можно существенно ослабить. Полученные формулы справедливы для всех уровней, для которых справедлива адиабатическая теорема. То есть это должен быть изолированный уровень, энергия которого отделена от остального спектра.

15.3 Эффект Ааронова–Бома

Другой пример возникновения нетривиальной связности на тривиальном главном расслоении $\mathbb{P}\Big(\mathbb{R}^4, \pi, \mathbb{U}(1)\Big) \simeq \mathbb{R}^4 \times \mathbb{U}(1)$ в нерелятивистской квантовой механике дает эффект Ааронова–Бома [**?**]. В этом случае в отличии от фазы Берри в качестве базы $\mathbb{M}$ главного расслоения выступает не пространство параметров, а пространство-время $\mathbb{R}^4$, в котором частица движется. Эффект Ааронова–Бома не связан с адиабатической теоремой, и обсуждается только с геометрической точки зрения.

Рассмотрим уравнение Шредингера (15.1), в котором гамильтониан описывает движение точечной частицы массы m в трехмерном евклидовом пространстве $\mathbb{R}^3$ с декартовыми координатами x^μ, $\mu = 1, 2, 3$,

$$H_0 = -\frac{\eta^{\mu\nu} p_\mu p_\nu}{2m} + U = -\frac{\hbar^2}{2m}\triangle + U, \tag{15.49}$$

где $p_\mu = i\hbar\partial_\mu$ – оператор импульса частицы, $\eta_{\mu\nu} = \mathsf{diag}\,(- - -)$ – отрицательно определенная пространственная метрика, $\triangle = \partial_1^2 + \partial_2^2 + \partial_3^2$ – оператор Лапласа и $U(x)$ – потенциальная энергия частицы.

Четырехмерный оператор импульса имеет вид $p_\alpha = i\hbar\partial_\alpha$, $\alpha = 0, 1, 2, 3$. При этом нулевая компонента 4-импульса $p_0 = i\hbar\partial_0 = i\hbar\partial_t$ имеет физический смысл оператора энергии частицы.

Если частица взаимодействует с электромагнитным полем, то это взаимодействие описывается с помощью минимальной подстановки для всех четырех компонент импульса

$$p_\alpha \quad \mapsto \quad i\hbar\partial_\alpha - \frac{e}{c}A_\alpha, \tag{15.50}$$

где e – заряд частицы, c – скорость света и A_α – потенциал электромагнитного поля (компоненты локальной формы $\mathbb{U}(1)$-связности). При этом нулевая компонента, разделенная на скорость света, A_0/c, имеет физический смысл потенциала электрического поля, а пространственные компоненты A_μ – ковекторного потенциала магнитного поля. Таким образом, точечная частица, находящаяся в электромагнитном поле, описывается уравнением Шредингера

$$i\hbar\frac{\partial\psi}{\partial t} = \left[\frac{\hbar^2}{2m}\eta^{\mu\nu}\left(\partial_\mu + i\frac{e}{\hbar c}A_\mu\right)\left(\partial_\nu + i\frac{e}{\hbar c}A_\nu\right) + \frac{e}{c}A_0\right]\psi + U\psi. \tag{15.51}$$

Здесь мы вернулись к стандартным обозначениям квантовой механики, когда гамильтониан действует на вектор гильбертова пространства слева, поскольку он содержит производные.

В дальнейшем, для простоты, положим $\hbar = 1$ и $c = 1$.

С геометрической точки зрения минимальная подстановка (15.50) с точностью до постоянных совпадает с заменой частной производной на ковариантную:

$$\partial_\alpha \quad \mapsto \quad \partial_\alpha + ieA_\alpha.$$

Посмотрим на уравнение Шредингера (15.51) с геометрической точки зрения. Оно решается во всем пространстве, $\psi = \psi(t, x)$, поэтому базой расслоения является четырехмерное евклидово пространство, $(t, x) \in \mathbb{R}\times\mathbb{R}^3 = \mathbb{R}^4$. При этом $\mathbb{R}^4$ рассматривается просто как четырехмерное многообразие без какой либо четырехмерной метрики. При желании метрику можно ввести, однако ее наличие никак не влияет на структуру главного расслоения и связности. Метрика $\eta_{\mu\nu}$ определена только на пространственных сечениях $t = \mathsf{const}$, поскольку она входит в уравнение Шредингера. Волновая функция $\psi(t, x)$ представляет собой сечение ассоциированного расслоения $\mathbb{E}\Big(\mathbb{R}^4, \pi_{\mathbb{E}}, \mathbb{C}, \mathbb{U}(1), \mathbb{P}\Big)$, типичным слоем которого является комплексная плоскость $\mathbb{C}$ и которое ассоциировано с некоторым главным расслоением $\mathbb{P}\Big(\mathbb{R}^4, \pi, \mathbb{U}(1)\Big)$. Это главное расслоение всегда тривиально, $\mathbb{P} \simeq \mathbb{R}^4 \times \mathbb{U}(1)$, т.к. базой является четырехмерное евклидово пространство. На этом главном расслоении задана локальная форма $\mathbb{U}(1)$-связности, которая определяется электромагнитным потенциалом $A_\alpha(t, x)$. В нерелятивистской квантовой механике рассматривается не все множество сечений ассоциированного расслоения, а лишь подмножество, состоящее из тех дифференцируемых функций $\psi(t, x)$, которые в каждый момент времени t принадлежат гильбертову пространству квадратично интегрируемых функций $\mathbb{H} = \mathcal{L}_2(\mathbb{R}^3)$ на пространственных сечениях $\mathbb{R}^3$.

Рассмотрим два случая.

15.3.1 Электрический потенциал

Предположим, что магнитный потенциал равен нулю, $A_\mu = 0$, $\mu = 1, 2, 3$. Запишем уравнение Шредингера в виде

$$i\dot{\psi} = (H_0 + eA_0)\psi, \tag{15.52}$$

где H_0 – гамильтониан системы в отсутствии электромагнитного поля (15.49). Предположим также, что электрический потенциал зависит только от времени, $A_0 = A_0(t)$. Будем искать решение уравнения Шредингера (15.52) в виде $\psi = \mathsf{e}^{-i\Theta}\phi$, где ϕ – решение свободного уравнения Шредингера,

$$i\dot{\phi} = H_0\phi, \tag{15.53}$$

и $\Theta = \Theta(t)$ – некоторая фаза, не зависящая от точки пространства. Подстановка $\psi = \mathsf{e}^{-i\Theta}\phi$ в исходное уравнение Шредингера (15.52) приводит к уравнению на фазу

$$\dot{\Theta} = eA_0,$$

где мы сократили общий фазовый множитель $\mathsf{e}^{-i\Theta}$ и ϕ. Это можно сделать, т.к. уравнение Шредингера должно выполняться для всех t и x. Решение полученного уравнения имеет вид

$$\Theta(t) = \Theta_0 + e\int_0^t ds A_0(s), \tag{15.54}$$

где Θ_0 – значение фазы волновой функции в начальный момент времени.

Аронов и Бом предложили следующий эксперимент, схема которого показана на рис.15.1. Пучок электронов делится на два пучка, которые пропускаются через две металлические трубки, на которые подается различный потенциал. Затем пучки собираются и на экране наблюдается интерференционная картина. Электрический потенциал, который подается на трубки, зависит от времени. Предполагается, что он равен нулю, пока оба пучка не окажутся в своих трубках. Затем он возрастает до некоторых значений, которые отличаются внутри каждой трубки, и снова падает до нуля перед выходом пучков из трубок. Таким образом, пучки находятся в поле A_0 только внутри трубок. Интерференционная картина зависит от разности фаз электронов в пучках, которую можно приближенно оценить следующим образом.

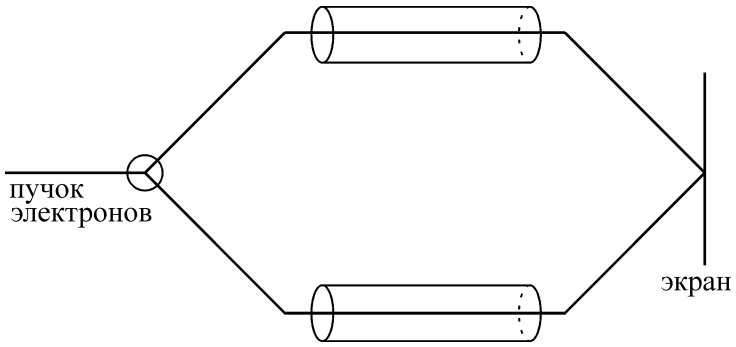


Рис. 15.1: Пучок электронов делится на два пучка, которые пропускаются через две металлические трубки, имеющие разные потенциалы. Затем пучки вновь собираются и наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

Предположим, что электрон описывается волновой функцией $\psi(t,x)$, которая удовлетворяет уравнению Шредингера (15.52) во всем пространстве-времени $\mathbb{R}^4$. Мы считаем, что в каждый момент времени носитель волновой функции отличен от нуля в небольшой окрестности пространства вблизи траектории частицы. Только в этом случае вообще можно говорить о траектории частицы. В частности, при прохождении электрона сквозь металлическую трубку предполагается, что носитель волновой функции целиком лежит внутри трубки. Математически это можно описать, введя в уравнение Шредингера (15.52) соответствующий потенциал. Этот гипотетический потенциал не меняет пространство-время, т.е. базу главного расслоения, а только обеспечивает движение электронов по заданной траектории. Изменение фазы электрона в верхнем пучке оценим следующим образом. Поскольку потенциал электрического поля однороден внутри трубки и носитель волновой функции целиком содержится внутри трубки, то можно считать, что фаза волновой функции определяется интегралом (15.54). Обозначим моменты времени, соответствующие расщеплению пучка и достижению экрана, соответственно t_1 и t_2. Тогда фаза волновой функции электрона в верхнем пучке при достижении экрана изменится на величину, задаваемую интегралом

$$\Theta_1 = e\int_{t_1}^{t_2} dt A_0^{(1)}(t),$$

где $A_0^{(1)}(t)$ – потенциал электрического поля в момент времени t, т.е. в той точке пространства, где в момент времени t находится электрон из верхнего пучка. Аналогично, изменение фазы волновой функции электрона из нижнего пучка равно

$$\Theta_2 = e\int_{t_1}^{t_2} dt A_0^{(2)}(t),$$

где $A_0^{(2)}$ – потенциал электрического поля вдоль нижней траектории. Ясно, что разность фаз электронов в верхнем и нижнем пучке, $\Theta_{\text{АВ}} = \Theta_2 - \Theta_1$, можно записать в виде интеграла

$$\Theta_{\text{АВ}} = e\oint_{\gamma} dt A_0(t). \tag{15.55}$$

вдоль замкнутого контура γ в пространстве-времени, когда сначала проходится нижняя половина контура, изображенного на рис.15.1, а затем – верхняя половина в обратную сторону. На рис.15.1 показана проекция контура γ на пространственную плоскость.

Вернемся к геометрической интерпретации. Разность фаз электронов дается интегралом (15.55), который однозначно определяется контуром γ и заданным на нем потенциалом A_0. Электрический потенциал A_0 представляет собой временну́ю компоненту локальной формы $\mathbb{U}(1)$-связности на главном расслоении $\mathbb{P}\Big(\mathbb{R}^4,\pi,\mathbb{U}(1)\Big)$. Поэтому разность фаз (15.55) определяет элемент группы голономии $\mathsf{e}^{i\Theta_{\text{АВ}}} \in \Phi\Big((t_0,x_0),e\Big) \subset \mathbb{U}(1)$ в точке $(t_0=0,x_0)$, где x_0 – точка пространства, в которой пучок расщепляется, а e – единица группы. База главного расслоения $\mathbb{P}$ имеет тривиальную топологию. Поэтому контур γ всегда можно стянуть в точку. При этом металлические трубки остаются на месте, а контур их непрерывно пересекает. При стягивании контура фаза Аронова–Бома $\Theta_{\text{АВ}}$ стремится к нулю, а соответствующий элемент группы голономии – к единице $e \in \mathbb{U}(1)$. Это хорошо известно, т.к. группа голономии является подгруппой Ли в структурной группе. Заметим, что стягивание контура никакого отношения к эксперименту не имеет, где наблюдается определенный элемент группы голономии, соответствующий начальному расположению контура.

Предлагаемая геометрическая конструкция не является единственной. Поскольку мы предполагаем, что носитель волновой функции сосредоточен в небольшой окрестности траектории, то можно построить много главных расслоений $\mathbb{P}'\Big(\mathbb{M},\pi,\mathbb{U}(1)\Big)$ с различными базами $\mathbb{M}$. Для этого достаточно взять расслоение $\mathbb{P}\Big(\mathbb{R}^4,\pi,\mathbb{U}(1)\Big)$ с определенной на нем связностью и сузить базу путем выкалывания произвольного числа и произвольно расположенных прямых в пространстве $\mathbb{R}^3$, которые перпендикулярны плоскости рисунка 15.1 и не лежат на контуре γ. В результате база $\mathbb{M}$ перестанет быть односвязным многообразием. Данная процедура представляется неестественной. Более того, она приводит к существенному усложнению задачи, т.к. исходное уравнение Шредингера (15.51) в таком случае необходимо решать на топологически нетривиальном многообразии.

В заключение данного раздела рассмотрим преобразование компоненты локальной формы $\mathbb{U}(1)$-связности при изменении сечения. Из уравнения Шредингера (15.52) следует, что при вертикальном автоморфизме

$$\psi' = \mathsf{e}^{ia}\psi,$$

где $a = a(t)$ – дифференцируемая функция времени, компонента локальной формы $\mathbb{U}(1)$ связности преобразуется по правилу

$$eA'_0 = eA_0 + \dot{a},$$

как и подобает компонентам локальной формы $\mathbb{U}(1)$-связности.

Таким образом, в основе эффекта Ааронова–Бома, так же как и для фазы Берри, лежит не топология, а нетривиальная геометрия, т.е. связность с нетривиальной группой голономии. При этом топология пространства может быть как тривиальной, так и нетривиальной.

15.3.2 Магнитный потенциал

Рассмотрим теперь случай, когда потенциал электрического поля равен нулю, $A_0 = 0$. Предположим, что ковекторный потенциал магнитного поля зависит только от пространственных координат x^μ и не завит от времени t (статическое поле). Тогда уравнение Шредингера примет вид

$$\begin{aligned} i\dot{\psi} &= \frac{1}{2m}\eta^{\mu\nu}(\partial_\mu + ieA_\mu)(\partial_\nu + ieA_\nu)\psi + U\psi \\ &= \frac{1}{2m}\eta^{\mu\nu}(\partial^2_{\mu\nu}\psi + 2ieA_\mu\partial_\nu\psi + ie\partial_\mu A_\nu\psi - e^2A_\mu A_\nu\psi) + U\psi. \end{aligned} \tag{15.56}$$

Пусть ϕ – решение уравнения Шредингера в отсутствии потенциала магнитного поля (15.53). Тогда нетрудно проверить, что функция

$$\psi = \mathsf{e}^{-i\Theta}\phi,$$

где фаза Θ удовлетворяет уравнению

$$\partial_\mu\Theta = eA_\mu \tag{15.57}$$

является решением исходного уравнения Шредингера (15.56).

Ааронов и Бом предложили эксперимент для определения фазы Θ, схема которого показана на рис.15.2. В этом эксперименте пучок электронов делится на два пучка,

которые огибают бесконечно длинный соленоид с постоянным магнитным потоком Ф, который перпендикулярен плоскости рисунка, с разных сторон. Затем пучки собираются вместе и на экране наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

Для оценки разности фаз электронов сделаем те же предположения, что и в случае электрического поля. А именно, будем считать, что уравнение Шредингера без магнитного потенциала имеет решение с носителем, который сосредоточен в малой окрестности траектории электрона. Мы предполагаем, что это можно осуществить путем введения в уравнение (15.53) соответствующего потенциала. Этот потенциал не меняет топологию пространства-времени, а только обеспечивает движение электронов вдоль заданной траектории. Тогда для фазы решения уравнения Шредингера с магнитным потенциалом справедливы уравнения (15.57). Поскольку магнитное поле вне соленоида равно нулю, $\partial_\mu A_\nu - \partial_\nu A_\mu = 0$, то условия интегрируемости для системы уравнений (15.57) выполнены. Поэтому разность фаз можно представить в виде контурного интеграла

$$\Theta_{\text{АВ}} = e \oint_\gamma dx^\mu A_\mu, \tag{15.58}$$

где γ – замкнутый контур в четырехмерном пространстве времени, который охватывает соленоид. Отметим, что слагаемое $dx^0 A_0$ в подынтегральном выражении равно нулю, т.к. $A_0 = 0$ по-предположению. Этот интеграл не зависит от выбора контура, охватывающего соленоид, поскольку магнитное поле вне соленоида равно нулю, $\partial_\mu A_\nu - \partial_\nu A_\mu = 0$.

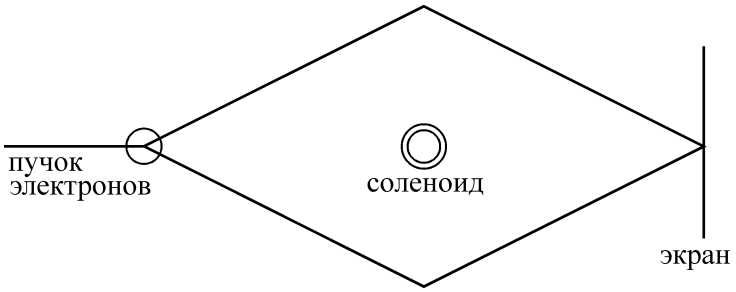


Рис. 15.2: Пучок электронов делится на два пучка, которые огибают тонкий соленоид с разных сторон. Затем пучки вновь собираются и наблюдается интерференционная картина, которая зависит от разности фаз электронов в разных пучках.

Фазу Аронова–Бома, используя формулу Стокса, можно записать в виде поверхностного интеграла

$$\Theta_{\text{АВ}} = \frac{1}{2} e \iint_S dx^\mu \wedge dx^\nu F_{\mu\nu} = e\Phi, \tag{15.59}$$

где $F_{\mu\nu}$ – напряженность магнитного поля (компоненты локальной 2-формы кривизны) и Ф – полный поток магнитного поля через соленоид. Заметим, что для применения формулы Стокса, необходимо считать, что магнитное поле определено в пространстве $\mathbb{R}^3$ всюду, включая сам соленоид.

Геометрическая интерпретация рассмотренного эффекта Аронова–Бома состоит в следующем. Мы имеем то же самое главное расслоение, что и в случае электрического потенциала, $\mathbb{P}\big(\mathbb{R}^4, \pi, \mathbb{U}(1)\big)$, базой которого является четырехмерное евклидово пространство, $(t, x) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$, в котором движутся пучки электронов, а структурной группой – унитарная группа $\mathbb{U}(1)$ (фазовый множитель $\mathsf{e}^{i\Theta}$ волновой функ-

ции). Однако связность на нем другая: отличны от нуля только пространственные компоненты локальной формы связности A_μ, $\mu = 1, 2, 3$. Разность фаз Ааронова–Бома (15.58) однозначно определяется контуром γ (тем же, что и в случае электрического потенциала) и значениями компонент связности A_μ на нем. При записи контурного интеграла в виде поверхностного (15.59) предполагается, что связность задана на всем пространстве-времени $\mathbb{R}^4$. Тем самым, мы рассматриваем соленоид конечного радиуса, чтобы избежать сингулярностей.

Таким образом, главное расслоение тривиально, а фаза Ааронова–Бома $\Theta_{\text{АВ}}$, зависящая от связности и контура, однозначно определяет элемент группы голономии. При стягивании контура в точку соответствующий элемент группы голономии стремится к единице $e \in \mathbb{U}(1)$, как и должно быть. При этом мы предполагаем, что контур свободно проходит через соленоид.

Волновая функция электрона, как и в случае электрического потенциала, является сечением ассоциированного расслоения $\mathbb{E}\Big(\mathbb{R}^4, \pi_{\mathbb{E}}, \mathbb{C}, \mathbb{U}(1), \mathbb{P}\Big)$, базой которого является евклидово пространство $\mathbb{R}^4$, а структурной группой – унитарная группа $\mathbb{U}(1)$. При вертикальном автоморфизме

$$\psi' = \mathsf{e}^{ia}\psi,$$

где $a = a(x)$ – дифференцируемая функция пространственных координат x^μ, $\mu = 1, 2, 3$, потенциал магнитного поля преобразуется по-правилу

$$eA'_\mu = eA_\mu - \partial_\mu a.$$

Это следует из уравнения Шредингера (15.56). Таким образом, компоненты потенциала магнитного поля действительно ведут себя, как компоненты локальной формы связности.

Поскольку разность фаз электронов (15.58) определяется значениями компонент локальной формы связности только вблизи контура интегрирования, то базу тривиального главного расслоения $\mathbb{P}\Big(\mathbb{R}^4, \pi, \mathbb{U}(1)\Big)$ можно сузить, не меняя ответа. Например, можно удалить область пространства-времени, лежащую внутри контура γ и содержащую соленоид. Тогда база расслоения перестанет быть односвязной. По этой причине эффект Ааронова–Бома часто называют топологическим. Как было показано выше, это совершенно необязательно. Достаточно считать, что магнитное поле отлично от нуля в ограниченной области на плоскости рис.15.2 внутри контура интегрирования. Если считать, что базой расслоения является евклидово пространство $\mathbb{R}^4$ с выколотым соленоидом, то это создает дополнительные трудности, т.к. уравнение Шредингера надо решать на соответствующем неодносвязном многообразии. Кроме того, формулу Стокса в этом случае применить нельзя. Таким образом, эффект Ааронова–Бома, вызванный магнитным потенциалом, также следует рассматривать как геометрический, а не топологический.

Эффект Ааронова–Бома как с электрическим, так и с магнитным потенциалом привлек большое внимание физиков по следующей причине. Согласно современным представлениям в калибровочных моделях наблюдаемыми величинами являются только калибровочно инвариантные функции. С этой точки зрения потенциал электромагнитного поля A_α, $\alpha = 0, 1, 2, 3$, сам по себе ненаблюдаем, т.к. не является калибровочно инвариантным. В рассмотренных примерах пучки электронов не подвергаются действию электромагнитных сил, поскольку напряженности электрического и магнитного поля в областях, через которые пролетают электроны, равны нулю. Поэтому, казалось бы, разность фаз в пучках электронов должна быть равна

нулю. Однако из уравнения Шредингера следует, что это не так. Следует отметить, что наблюдаемым является не сам потенциал электромагнитного поля, а интеграл от него по замкнутому контуру, который определяет элемент группы голономии $\mathbb{U}(1)$-связности, который является инвариантным объектом.

Вскоре после публикации статьи, эффект Ааронова–Бома был подтвержден экспериментально. Эффект, вызванный магнитным потенциалом наблюдался в экспериментах [**?**, **?**, **?**].

Глава 16

Векторные поля Киллинга

В разделе 10.3 были рассмотрены геометрические структуры на многообразии, которые инвариантны относительно действия некоторой группы преобразований. Вопрос ставился так. Пусть задана группа преобразований, и требуется найти такие структуры, которые инвариантны относительно этих преобразований. Обратная задача нахождения группы преобразований, которую допускает заданная геометрическая структура также очень важна.

Изучение преобразований, которые сохраняют метрику пространства-времени играет исключительно важную роль в математической физике. Достаточно сказать, что с такими преобразованиями связаны наиболее важные законы сохранения. В настоящей главе мы дадим определение векторных полей Киллинга, которые являются генераторами локальных симметрий метрики, а также изучим некоторые из их свойств.

16.1 Изометрии и инфинитезимальные изометрии

Рассмотрим n-мерное (псевдо-)риманово многообразие $(\mathbb{M}, g)$ с метрикой $g(x) = g_{\alpha\beta}(x)dx^\alpha \otimes dx^\beta$, $\alpha, \beta = 0, 1, \ldots, n-1$ и соответствующей связностью Леви-Чивиты Γ.

Определение. Диффеоморфизм

$$\imath : \quad \mathbb{M} \ni x \mapsto x' = \imath(x) \in \mathbb{M}$$

называется *изометрией* или *движением* многообразия $\mathbb{M}$, если он сохраняет метрику,

$$g(x) = \imath^* g(x'), \tag{16.1}$$

где $\imath^*$ – возврат отображения $\imath$. □

Замечание. В настоящей главе, для простоты, мы не будем использовать знак тильды для обозначения компонент связности Леви–Чивиты, т.к. аффинная связность общего вида с кручением и неметричностью использоваться не будет. □

Из условия инвариантности метрики (16.1) следует инвариантность скалярного произведения векторов. Пусть $X, Y \in \mathbb{T}_x(\mathbb{M})$ – два произвольных вектора из касательного пространства в точке $x \in \mathbb{M}$. Тогда справедливо равенство

$$g(X,Y)|_x = \imath^* g(\imath_* X, \imath_* Y)|_{\imath(x)} = g(\imath_* X, \imath_* Y)|_{\imath(x)}.$$

Поскольку изометрия сохраняет метрику, то она сохраняет также связность Леви-Чивиты, соответствующий тензор кривизны, экстремали и, вообще, все геометрические объекты, которые определяются только метрикой.

Запишем отображение (16.1) в координатах. Пусть обе точки x и x' принадлежат одной координатной окрестности и имеют, соответственно, координаты x^α и x'^α. Тогда изометрия $\imath$ в координатах запишется в виде условия

$$g_{\alpha\beta}(x) = \frac{\partial x'^\gamma}{\partial x^\alpha}\frac{\partial x'^\delta}{\partial x^\beta} g_{\gamma\delta}(x'), \tag{16.2}$$

связывающего компоненты метрики в различных точках многообразия. Это условие по виду совпадает с правилом преобразования компонент метрики при преобразовании координат (5.4). Разница заключается в следующем. При преобразовании координат мы считаем, что одной и той же точке $x \in \mathbb{M}$ соответствует два набора координат $\{x^\alpha\}$ и $\{x^{\alpha'} := x'^\alpha\}$ в двух различных системах координат. При рассмотрении изометрий x и x' – это две различные точки одного и того же многообразия $\mathbb{M}$.

Предложение 16.1.1. *Множество всех изометрий данного (псевдо-)риманова многообразия* $(\mathbb{M}, g)$ *является группой, которую обозначим* $\mathbb{I}(\mathbb{M})$.

Доказательство. Две последовательных изометрии также являются изометрией. Тождественное отображение многообразия $\mathbb{M}$ является изометрией и представляет собой единицу группы. У каждого диффеоморфизма $\imath$ есть обратной диффеоморфизм $\imath^{-1}$, который является обратной изометрией. □

Если метрика на многообразии задана, т.е. определены значения ее компонент во всех точках x, то соотношение (16.2) представляет собой уравнение на функции $x'(x)$, которые определяют изометрию. В общем случае это уравнение не имеет решений и у соответствующего (псевдо-)риманова многообразия нет никаких нетривиальных изометрий. В этом случае группа изометрий состоит из одного единичного элемента. Чем шире группа изометрий, тем уже класс соответствующих (псевдо-)римановых многообразий.

Пример 16.1.1. Евклидово пространство $\mathbb{R}^n$ с евклидовой метрикой $\delta_{\alpha\beta}$ допускает группу изометрий, которая состоит из преобразований неоднородной группы вращений $\mathbb{IO}(n,\mathbb{R})$, $\mathsf{dim}\,\mathbb{IO}(n,\mathbb{R}) = \frac{1}{2}n(n+1)$, состоящей из вращений, сдвигов и отражений. □

Группа изометрий $\mathbb{I}(\mathbb{M})$ может быть дискретной или группой Ли.

Определение. Если группа изометрий $\mathbb{I}(\mathbb{M})$ является группой Ли, то имеет смысл говорить об инфинитезимальных преобразованиях (см. раздел 10.2). В этом случае мы говорим об *инфинитезимальных изометриях*. Каждая инфинитезимальная изометрия генерируется некоторым достаточно гладким векторным полем, которое называется векторным *полем Киллинга*. □

Замечание. Дискретные изометрии (псевдо-)риманова многообразия, например, отражения, не генерируются никакими векторными полями. □

Запишем условие инвариантности метрики относительно инфинитезимальных преобразований из группы изометрий в координатах. В разделе 3.6.5 было показано,

что каждое векторное поле генерирует однопараметрическую группу преобразований, которая называется экспоненциальным отображением. Формально условие инвариантности метрики записывается в виде равенства нулю производной Ли вдоль векторного поля Киллинга $K = K^\alpha \partial_\alpha$ от метрики [**?**]

$$\mathsf{L}_K g = 0. \tag{16.3}$$

Используя явное выражение для производной Ли (3.120), это уравнение в локальной системе координат принимает вид

$$\nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0, \tag{16.4}$$

где $K_\alpha = K^\beta g_{\beta\alpha}$, а ковариантная производная

$$\nabla_\alpha K_\beta = \partial_\alpha K_\beta - \Gamma_{\alpha\beta}{}^\gamma K_\gamma$$

строится по символам Кристоффеля $\Gamma_{\alpha\beta}{}^\gamma$.

Определение. Уравнение (16.4) называется *уравнением Киллинга*, а интегральные кривые полей Киллинга называются *траекториями Киллинга*. Если $K = K^\alpha \partial_\alpha$ – векторное поле Киллинга, то ему соответствует 1-форма $K = dx^\alpha K_\alpha$, где $K_\alpha = K^\beta g_{\beta\alpha}$, которая называется *формой Киллинга*, и для которой мы сохранили то же обозначение. □

На любом (псевдо-)римановом многообразии $(\mathbb{M}, g)$ уравнения Киллинга (16.3) всегда имеют тривиальное решение $K = 0$. Если уравнения Киллинга имеют только тривиальное решение, то в этом случае непрерывные изометрии отсутствуют.

Траектории Киллинга $\{x^\alpha(t)\} \in \mathbb{M}$, где $t \in \mathbb{R}$, определяются системой обыкновенных дифференциальных уравнений

$$\dot{x}^\alpha = K^\alpha.$$

Если траектория Киллинга при $t = 0$ проходит через точку $p = \{p^\alpha\} \in \mathbb{M}$, то при малых t она имеет вид

$$x^\alpha(t) = p^\alpha + tK^\alpha(p) + \mathsf{o}(t). \tag{16.5}$$

Если в некоторой точке векторное поле Киллинга равно нулю, то эта точка остается неподвижной, т.е. является стационарной точкой группы изометрий. Поскольку изометрии определены для всего многообразия $\mathbb{M}$ и образуют группу, то векторные поля Киллинга обязаны быть полными, т.е. параметр t должен меняться на всей вещественной прямой $\mathbb{R}$.

Если для (псевдо-)риманова многообразия $(\mathbb{M}, g)$ известно векторное поле Киллинга, то оно определяет не только инфинитезимальные изометрии, но и всю однопараметрическую группу диффеоморфизмов. Для этого нужно найти интегральные кривые x(t), проходящие, через все точки многообразия $p \in \mathbb{M}$. Если $x(0) = p$, то каждому значению $t \in \mathbb{R}$ соответствует диффеоморфизм

$$\imath : \quad \mathbb{M} \ni p \mapsto x(t) \in \mathbb{M}.$$

Уравнения для векторных полей Киллинга в ковариантной форме (16.4) можно переписать в частных производных,

$$\partial_\alpha K_\beta + \partial_\beta K_\alpha - 2\Gamma_{\alpha\beta}{}^\gamma K_\gamma = 0.$$

В моделях математической физики часто ставится задача нахождения векторов Киллинга для заданной метрики на многообразии. Для решения этой задачи бывает удобнее использовать контравариантные компоненты векторов Киллинга, для которых уравнение Киллинга принимает вид

$$g_{\alpha\gamma}\partial_\beta K^\gamma + g_{\beta\gamma}\partial_\alpha K^\gamma + K^\gamma \partial_\gamma g_{\alpha\beta} = 0. \tag{16.6}$$

Полученное уравнение линейно по компонентам метрики и компонентам векторов Киллинга. Отсюда сразу следует, что две метрики, которые отличаются постоянным множителем, имеют один и тот же набор векторов Киллинга. Кроме того, векторное поле Киллинга определено с точностью до умножения на произвольную постоянную, отличную от нуля. В частности, если K^α – векторное поле Киллинга, то и $-K^\alpha$ также является полем Киллинга. Также любая линейная комбинация полей Киллинга является полем Киллинга.

Замечание. Векторные поля Киллинга не выдерживает умножения на функцию. Поэтому они не образуют $\mathcal{C}^\infty(\mathbb{M})$-модуль в отличие от множества всех векторных полей $\mathcal{X}(\mathbb{M}$. □

Предложение 16.1.2. *Пусть (псевдо-)риманово многообразие $(\mathbb{M}, g)$ имеет $k < \dim \mathbb{M}$ отличных от нуля коммутирующих между собой и линейно независимых векторных полей Киллинга K_i, $i = 1, \ldots, k$. Тогда существует такая система координат, в которой все компоненты метрики не зависят от k координат, соответствующих траекториям Киллинга.*

Доказательство. В разделе 3.6.5 была построена специальная система координат, связанная с произвольным векторным полем, отличным от нуля. Применительно к векторным полям Киллинга K_i это означает, что существует такая система координат $(x^1, \ldots, x^n)$, в которой каждое поле Киллинга имеет только одну постоянную компоненту, $K_i = \partial_i$. В этой системе координат уравнение (16.6) для каждого поля Киллинга принимает особенно простой вид

$$\partial_i g_{\alpha\beta} = 0, \qquad i = 1, \ldots, k.$$

Это значит, что все компоненты метрики не зависят от координат x^i. В этой системе координат траектории Киллинга определяются уравнениями

$$\dot{x}^i = 1, \quad \dot{x}^\mu = 0, \qquad \mu = 1, \ldots, i-1, i+1, \ldots, n.$$

Отсюда следует, что координатными линиями x^i являются траектории Киллинга. □

Если риманово многообразие $(\mathbb{M}, g)$ имеет два или более некоммутирующих векторных полей Киллинга, то это отнюдь не означает, что существует такая система координат, в которой компоненты метрики не зависят от двух или более координат.

Пример 16.1.2. Рассмотрим двумерную сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Пусть метрика g на сфере индуцирована вложением. Риманово пространство $(\mathbb{S}^2, g)$ имеет три векторных поля Киллинга, соответствующих $\mathbb{SO}(3)$ вращениям евклидова пространства $\mathbb{R}^3$. Легко понять, что на сфере не существует локальной системы координат, в которой компоненты метрики не зависели бы от двух координат. Действительно, это означает, что в данной системе координат компоненты метрики постоянны, и, следовательно, кривизна равна нулю. Но это невозможно, поскольку кривизна сферы постоянна. □

В общей теории относительности мы предполагаем, что пространство-время является псевдоримановым многообразием с метрикой лоренцевой сигнатуры. Используя понятие векторного поля Киллинга, можно дать инвариантное

Определение. Пространство-время или его область называются *статическими*, если на них существует времениподобное векторное поле Киллинга. □

Векторные поля Киллинга определены глобально и удовлетворяют уравнениям Киллинга на всем $\mathbb{M}$. В то же время уравнения Киллинга – это локальный объект, в том смысле, что они определены в каждой окрестности и могут иметь нетривиальные решения только на некотором подмногообразии $\mathbb{U} \subset \mathbb{M}$.

Пример 16.1.3. Рассмотрим гладкую замкнутую двумерную поверхность $\mathbb{M}$, вложенную в трехмерное евклидово пространство $\mathbb{R}^3$, как показано на рис.16.1. Отличительной особенностью этой поверхности является то, что ее нижняя часть является плоской. Пусть метрика на $\mathbb{M}$ индуцирована вложением $\mathbb{M} \hookrightarrow \mathbb{R}^3$. Тогда уравнения Киллинга в нижней части поверхности легко интегрируются, как и на евклидовой плоскости. Однако найденные нетривиальные решения не будут в общем случае нетривиальными на всем $\mathbb{M}$. Действительно, верхняя часть поверхности может быть искривлена так, что уравнения Киллинга на ней имеют только тривиальное решение. Следовательно, векторные поля Киллинга могут быть нетривиальны только на части многообразия $\mathbb{M}$. □

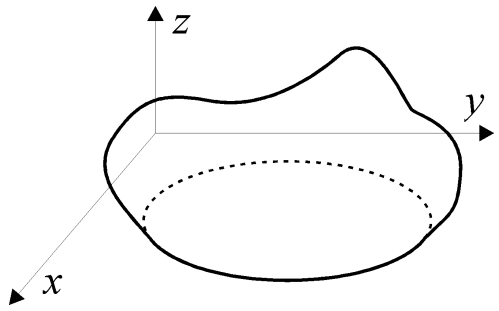


Рис. 16.1: Двумерная поверхность, вложенная в трехмерное евклидово пространство. Нижняя часть поверхности является плоской.

Группа изометрий $\mathbb{I}(\mathbb{M})$ является группой преобразований многообразия $\mathbb{M}$, которая действует слева или справа, в зависимости от нашего соглашения. Эта группа преобразований может иметь подгруппу $\mathbb{H} \subset \mathbb{I}(\mathbb{M})$, действующую на $\mathbb{M}$ свободно и собственно разрывно. Тогда фактор пространство $\mathbb{M}/\mathbb{H}$ является многообразием. Более того, на нем можно определить метрику, поскольку $\mathbb{H}$ – изометрия. Таким образом, фактор пространство $\mathbb{M}/\mathbb{H}$ превращается в (псевдо-)риманово многообразие. Поскольку каждое векторное поля Киллинга на $\mathbb{M}$ при факторизации переходит в некоторое поле Киллинга на $\mathbb{M}/\mathbb{H}$ и линейная независимость полей Киллинга при этом сохраняется, то группа изометрий фактор пространства $\mathbb{M}/\mathbb{H}$ совпадает с $\mathbb{I}(\mathbb{M})$.

16.2 Однородные и изотропные многообразия

С каждым полем Киллинга как и с произвольным векторным полем связана однопараметрическая группа преобразований, которая в данном случае сохраняет метрику.

Произвольная линейная комбинация векторов Киллинга ввиду линейности уравнения Киллинга (16.4) снова дает вектор Киллинга. То есть поля Киллинга образуют векторное пространство над полем вещественных чисел. В этом векторном пространстве можно ввести билинейную операцию. Простые вычисления показывают, что коммутатор двух векторных полей Киллинга K_1 и K_2 снова дает поле Киллинга:

$$\mathsf{L}_{[K_1,K_2]}g = \mathsf{L}_{K_1} \circ \mathsf{L}_{K_2}g - \mathsf{L}_{K_2} \circ \mathsf{L}_{K_1}g = 0,$$

Отсюда следует, что векторные поля Киллинга образуют алгебру Ли $\mathfrak{i}(\mathbb{M})$ над полем вещественных чисел, которая является подалгеброй алгебры Ли множества всех векторных полей, $\mathfrak{i}(\mathbb{M}) \subset \mathcal{X}(\mathbb{M})$. Эта алгебра является алгеброй Ли группы изометрий $\mathbb{I}(\mathbb{M})$.

Уравнения Киллинга (16.4) накладывают сильные ограничения на векторные поля Киллинга, которые мы сейчас обсудим. Воспользовавшись тождеством для коммутатора ковариантных производных (7.82), получаем

$$\nabla_\alpha\nabla_\beta K_\gamma - \nabla_\beta\nabla_\alpha K_\gamma = -R_{\alpha\beta\gamma}{}^{\delta}K_\delta. \tag{16.7}$$

Теперь воспользуемся тождеством (7.75) для тензора кривизны и уравнениями Киллинга (16.4). В результате получим тождество для векторных полей Киллинга:

$$\nabla_\alpha\nabla_\beta K_\gamma + \nabla_\beta\nabla_\gamma K_\alpha + \nabla_\gamma\nabla_\alpha K_\beta = 0,$$

где слагаемые отличаются циклической перестановкой индексов. Это равенство позволяет переписать уравнение (16.7) в виде

$$\nabla_\gamma\nabla_\alpha K_\beta = R_{\alpha\beta\gamma}{}^{\delta}K_\delta. \tag{16.8}$$

Полученное равенство является следствием уравнений Киллинга, но не эквивалентно им. Оно позволяет сделать важные выводы. Предположим, что компоненты вектора Киллинга бесконечно дифференцируемы в некоторой точке многообразия $p \in \mathbb{M}$ и разлагаются в ряд Тейлора, который сходится в некоторой окрестности $\mathbb{U}_p$. Допустим, что в точке $p \in \mathbb{M}$ нам известны компоненты формы Киллинга $K_\alpha(p)$ и их первых производных $\partial_\beta K_\alpha(p)$. Тогда соотношения (16.8) позволяют вычислить все вторые производные от компонент формы Киллинга $\partial^2_{\beta\gamma}K_\alpha$. Теперь возьмем ковариантную производную от равенства (16.8) и получим некоторое соотношение, линейное по третьим производным. Из него можно найти все третьи производные от вектора Киллинга и т.д. до бесконечности. Важно отметить, что все соотношения линейны по компонентам формы Киллинга и их производным. Это значит, что в окрестности $\mathbb{U}_p$ компоненты векторного поля Киллинга имеют вид

$$K_\alpha(x) = A_\alpha{}^{\beta}(x,p)K_\beta(p) + B_\alpha{}^{\beta\gamma}(x,p)\left[\partial_\beta K_\gamma(p) - \partial_\beta K_\alpha(p)\right], \tag{16.9}$$

где $A_\alpha{}^{\beta}(x,p)$ и $B_\alpha{}^{\beta\gamma}(x,p)$ – некоторые функции. Антисимметрия последнего слагаемого по индексам β,γ связана с тем, что симметризованная частная производная выражается через компоненты формы Киллинга в силу уравнения Киллинга (16.4). Таким образом, компоненты формы Киллинга в окрестности $\mathbb{U}_p$ являются линейными функциями от компонент формы Киллинга в точке p и ее внешней производной в той же точке.

Пусть векторные поля Киллинга гладкие на $\mathbb{M}$ и их компоненты разлагаются в ряды Тейлора в окрестности каждой точки $p \in \mathbb{M}$. Обозначим через $\mathbb{U}_p$ окрестность

точки p, в которой разложение (16.9) справедливо и обратимо, т.е. аргументы x и p можно поменять местами для некоторых новых матриц A и B. Рассмотрим точку q, которая лежит вне $\mathbb{U}_p$. Для этой точки также справедливо обратимое разложение вида (16.9) в некоторой окрестности $\mathbb{U}_q$. Предположим, что точка q лежит достаточно близко к $\mathbb{U}_p$ так, что окрестности пересекаются, $\mathbb{U}_p \cap \mathbb{U}_q \neq \emptyset$. Тогда для всех точек из пересечения $x \in \mathbb{U}_p \cap \mathbb{U}_q$ справедливо разложение (16.9) по компонентам форм Киллинга $K(p)$ и $K(q)$ и их внешним производным. Отсюда следует, что компоненты формы Киллинга и ее внешней производной в точке q линейно выражаются через компоненты формы Киллинга и ее внешней производной в точке p. Таким образом, разложение (16.9) справедливо также в объединении $\mathbb{U}_p \cup \mathbb{U}_q$. Это построение можно продолжить на все многообразие $\mathbb{M}$. Поэтому разложение (16.9) справедливо для всех точек $x \in \mathbb{M}$.

Теперь предположим, что (псевдо-)риманово многообразие $(\mathbb{M}, g)$ имеет несколько векторных полей Киллинга K_i, $i = 1, \ldots, \mathrm{N}$. Тогда для каждого векторного поля Киллинга справедливо разложение (16.9)

$$K_{i\alpha}(x) = A_\alpha{}^\beta(x,p) K_{i\beta}(p) + B_\alpha{}^{\beta\gamma}(x,p)\left[\partial_\beta K_{i\gamma}(p) - \partial_\gamma K_{i\beta}(p)\right], \qquad i = 1, \ldots, \mathrm{N}, \quad (16.10)$$

с теми же функциями $A_\alpha{}^\beta(x,p)$ и $B_\alpha{}^{\beta\gamma}(x,p)$. Эти функции одинаковы для всех полей Киллинга, потому что соотношения (16.8) линейны по компонентам векторов Киллинга и их производным. Они полностью определяются метрикой, тензором кривизны и их ковариантными производными. В полученном разложении точка $p \in \mathbb{M}$ произвольна, но фиксирована, а точка $x \in \mathbb{M}$ пробегает все многообразие.

Соотношение (16.8) представляет собой систему уравнений в частных производных на компоненты формы Киллинга, у которого есть нетривиальные условия разрешимости. Одно из этих условий в ковариантной форме имеет вид

$$[\nabla_\gamma \nabla_\delta]\, \nabla_\alpha K_\beta = -R_{\gamma\delta\alpha}{}^\epsilon \nabla_\epsilon K_\beta - R_{\gamma\delta\beta}{}^\epsilon \nabla_\alpha K_\epsilon,$$

где квадратные скобки обозначают коммутатор ковариантных производных. Подстановка в левую часть этого уравнения исходного выражения для вторых производных от формы Киллинга (16.8) после несложных алгебраических преобразований дает

$$\left(R_{\alpha\beta\gamma}{}^\epsilon \delta^\zeta_\delta - R_{\alpha\beta\delta}{}^\epsilon \delta^\zeta_\gamma + R_{\gamma\delta\alpha}{}^\epsilon \delta^\zeta_\beta - R_{\gamma\delta\beta}{}^\epsilon \delta^\zeta_\alpha\right) \nabla_\zeta K_\epsilon = \left(\nabla_\gamma R_{\alpha\beta\delta}{}^\epsilon - \nabla_\delta R_{\alpha\beta\gamma}{}^\epsilon\right) K_\epsilon. \quad (16.11)$$

Если кривизна нетривиальна, то это уравнение дает некоторые линейные соотношения между компонентами формы Киллинга K_α и их ковариантными производными $\nabla_\beta K_\alpha$. Наоборот, если существует некоторая информация в формах Киллинга, то полученное уравнение может определить структуру тензора кривизны. В теореме 16.2.1, которая сформулирована ниже, соотношение (16.11) будет использовано для доказательства того, что однородное и изотропное многообразие является пространством постоянной кривизны.

Перейдем к определениям.

Определение. (Псевдо-)риманово многообразие $(\mathbb{M}, g)$ размерности $\dim \mathbb{M} = n$ называется *однородным в точке* $p \in \mathbb{M}$, если существуют инфинитезимальные изометрии, которые переводят эту точку в любую другую точку из некоторой окрестности $\mathbb{U}_p$. Другими словами, метрика должна допускать такие векторные поля Киллинга, которые в точке p имеют все возможные направления. Поскольку векторы Киллинга образуют линейное пространство, то в сопряженном пространстве достаточно существования такого набора из n форм Киллинга $K^{(\gamma)} = dx^\alpha K_\alpha{}^{(\gamma)}(x,p)$, где индекс γ в

скобках нумерует формы Киллинга, что выполнены условия:

$$K_\alpha{}^{(\gamma)}(p,p) = \delta^\gamma_\alpha. \tag{16.12}$$

Если (псевдо-)риманово многообразие $(\mathbb{M}, g)$ однородно в каждой своей точке, то оно называется *однородным*.

(Псевдо-)риманово многообразие $(\mathbb{M}, g)$ называется *изотропным в точке* $p \in \mathbb{M}$, если существуют такие инфинитезимальные изометрии с формами Киллинга $K(x)$, которые оставляют эту точку на месте, т.е. $K(p) = 0$, и для которых внешняя производная $dK(p)$ в точке p принимает любое значение в пространстве 2-форм $\Lambda_2(\mathbb{M})|_p$ в точке p. Для этого достаточно существования такого набора из $\frac{1}{2}n(n-1)$ форм Киллинга $K^{[\gamma\delta]} = -K^{[\delta\gamma]} = dx^\alpha K_\alpha{}^{[\gamma\delta]}(x,p)$, где индексы γ, δ нумеруют формы Киллинга, что выполнены условия:

$$\begin{gathered} K_\alpha{}^{[\gamma\delta]}(p,p) = 0, \\ \left.\frac{\partial K_\beta{}^{[\gamma\delta]}(x,p)}{\partial x^\alpha}\right|_{x=p} = \delta^{\gamma\delta}_{\alpha\beta} - \delta^{\delta\gamma}_{\alpha\beta}. \end{gathered} \tag{16.13}$$

Если (псевдо-)риманово многообразие $(\mathbb{M}, g)$ изотропно в каждой своей точке, то оно называется *изотропным*. □

В силу непрерывности, наборы форм $K^{(\gamma)}$ и $K^{[\gamma\delta]}$ линейно независимы в некоторой окрестности точки p.

Предложение 16.2.1. *Любое изотропное (псевдо-)риманово многообразие* $(\mathbb{M}, g)$ *является также однородным.*

Доказательство. Если многообразие изотропно, то формы Киллинга $K^{[\gamma,\delta]}(x,p)$ и $K^{[\gamma,\delta]}(x,p+dp)$ удовлетворяют условиям (16.13) в близких точках p и $p+dp$ соответственно. Любая их линейная комбинация будет формой Киллинга и, следовательно, производная

$$\frac{\partial K^{[\gamma\delta]}(x,p)}{\partial p^\alpha}$$

также будет формой Киллинга. Вычислим производную по x этой формы Киллинга в точке p. Из первого условия в (16.13) следует равенство

$$\frac{\partial}{\partial p^\alpha} K_\beta{}^{[\gamma\delta]}(p,p) = \left.\frac{\partial K_\beta{}^{[\gamma\delta]}(x,p)}{\partial x^\alpha}\right|_{x=p} + \left.\frac{\partial K_\beta{}^{[\gamma\delta]}(x,p)}{\partial p^\alpha}\right|_{x=p} = 0.$$

Откуда, с учетом второго условия в (16.13), получаем

$$\left.\frac{\partial K_\beta{}^{[\gamma\delta]}(x,p)}{\partial p^\alpha}\right|_{x=p} = -\delta^{\gamma\delta}_{\alpha\beta} + \delta^{\delta\gamma}_{\alpha\beta}.$$

Отсюда следует, что из форм Киллинга $K^{[\gamma\delta]}$ можно построить форму Киллинга, которая в точке p принимает любое заданное значение $dx^\alpha a_\alpha$, где $a_\alpha \in \mathbb{R}$. Для этого достаточно положить

$$K_\alpha = \frac{a_\gamma}{n-1} \frac{\partial K_\alpha{}^{[\gamma\delta]}(x,p)}{\partial p^\delta}.$$

Выбрав соответствующим образом постоянные a_γ, получим набор форм Киллинга, который удовлетворяет условиям (16.12). □

Теорема 16.2.1. *Алгебра Ли* $\mathfrak{i}(\mathbb{M})$ *инфинитезимальных изометрий связного (псевдо-)риманова многообразия* $\mathbb{M}$ *имеет размерность не более, чем* $\frac{1}{2}n(n+1)$, *где* $n = \dim \mathbb{M}$. *Если размерность максимальна,* $\dim \mathfrak{i}(\mathbb{M}) = \frac{1}{2}n(n+1)$, *то* $\mathbb{M}$ *есть пространство постоянной кривизны.*

Доказательство. Размерность алгебры Ли $\mathfrak{i}(\mathbb{M})$ равна максимальному числу линейно независимых векторных полей Киллинга на многообразии $\mathbb{M}$. Из равенства (16.10) следует, что число независимых векторных полей Киллинга N не может превышать числа независимых компонент формы $\{K_\alpha(p)\}$ и ее внешней производной $\{\partial_\beta K_\alpha(p) - \partial_\alpha K_\beta(p)\}$ в фиксированной точке $p \in \mathbb{M}$. Число независимых компонент любой 1-формы в фиксированной точке не превосходит n, а число независимых компонент внешней производной не может превышать $\frac{1}{2}n(n-1)$. Поэтому справедливо следующее ограничение на размерность алгебры Ли векторных полей Киллинга:

$$\dim \mathfrak{i}(\mathbb{M}) \leq n + \frac{1}{2}n(n-1) = \frac{1}{2}n(n+1).$$

Это доказывает первое утверждение теоремы.

Однородные и изотропные многообразия имеют максимальное число $\frac{1}{2}n(n+1)$ векторных полей Киллинга и, в силу разложения (16.10), определяют все возможные векторы Киллинга на многообразии $\mathbb{M}$. Следовательно, если некоторое многообразие имеет максимальное число независимых полей Киллинга, то оно с необходимостью должно быть однородным и изотропным.

Теперь докажем, что любое однородное и изотропное пространство является пространством постоянной кривизны. Если пространство однородно и изотропно, то для каждой точки $x \in \mathbb{M}$ найдутся такие формы Киллинга, для которых $K_\alpha(x) = 0$, а $\nabla_\beta K_\alpha(x)$ является произвольной антисимметричной матрицей. Отсюда следует, что антисимметризованный коэффициент при $\nabla_\zeta K_\epsilon$ в уравнении (16.11) должен быть равен нулю, что приводит к равенству

$$R_{\alpha\beta\gamma}{}^{\epsilon}\delta^{\zeta}_{\delta} - R_{\alpha\beta\delta}{}^{\epsilon}\delta^{\zeta}_{\gamma} + R_{\gamma\delta\alpha}{}^{\epsilon}\delta^{\zeta}_{\beta} - R_{\gamma\delta\beta}{}^{\epsilon}\delta^{\zeta}_{\alpha} = R_{\alpha\beta\gamma}{}^{\zeta}\delta^{\epsilon}_{\delta} - R_{\alpha\beta\delta}{}^{\zeta}\delta^{\epsilon}_{\gamma} + R_{\gamma\delta\alpha}{}^{\zeta}\delta^{\epsilon}_{\beta} - R_{\gamma\delta\beta}{}^{\zeta}\delta^{\epsilon}_{\alpha}. \quad (16.14)$$

Ранее было доказано (доказательство предложения 16.2.1), что для произвольной точки $x \in \mathbb{M}$ существуют также такие формы Киллинга, которые принимают в этой точке произвольные значения. Следовательно, из уравнений (16.11) и (16.14) вытекает равенство

$$\nabla_\gamma R_{\alpha\beta\delta}{}^{\epsilon} = \nabla_\delta R_{\alpha\beta\gamma}{}^{\epsilon}. \quad (16.15)$$

Теперь свернем уравнение (16.14) по индексам δ, ζ и опустим верхний индекс. В результате получим выражение тензора кривизны через тензор Риччи и метрику:

$$(n-1)R_{\alpha\beta\gamma\delta} = R_{\beta\delta}g_{\alpha\gamma} - R_{\alpha\delta}g_{\beta\gamma}. \quad (16.16)$$

Правая часть этой формулы должна быть антисимметрична по индексам δ, γ. Поэтому возникает дополнительное ограничение

$$R_{\beta\delta}g_{\alpha\gamma} - R_{\alpha\delta}g_{\beta\gamma} = -R_{\beta\gamma}g_{\alpha\delta} + R_{\alpha\gamma}g_{\beta\delta}.$$

Свертка полученного равенства по индексам β, γ дает связь между тензором Риччи и скалярной кривизной:

$$R_{\alpha\delta} = \frac{1}{n}Rg_{\alpha\delta}. \quad (16.17)$$

Подстановка этого выражения в (16.16) приводит к следующему выражению для тензора кривизны

$$R_{\alpha\beta\gamma\delta} = \frac{R}{n(n-1)} \left(g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma} \right). \tag{16.18}$$

Теперь осталось доказать, что скалярная кривизна R однородного и изотропного пространства постоянна. Для этой цели используем свернутые тождества Бианки

$$2\nabla_\beta R_\alpha{}^\beta - \nabla_\alpha R = 0.$$

Подставляя в это тождество выражение для тензора Риччи (16.17), получаем условие

$$\left(\frac{2}{n} - 1 \right) \partial_\alpha R = 0.$$

При $n \geq 3$ отсюда следует $R = \mathsf{const}$.

Случай $n = 2$ требует особого рассмотрения. Свертка равенства (16.15) по индексам β, ϵ дает

$$\nabla_\gamma R_{\alpha\delta} - \nabla_\delta R_{\alpha\gamma} = 0.$$

дальнейшая свертка с $g^{\alpha\delta}$ с учетом уравнения (16.17) приводит к условию $\partial_\gamma R = 0$, т.е. $R = \mathsf{const}$ и при $n = 2$.

Таким образом, скалярная кривизна в выражении для полного тензора кривизны (16.18) равна константе, $R = \mathsf{const}$, и максимально симметричное (псевдо-)риманово многообразие является пространством постоянной кривизны. □

Замечание. Если тензор кривизны имеет вид (16.18), где $R = \mathsf{const}$, то соответствующее многообразие является пространством постоянной кривизны, т.к. ковариантная производная от метрики в римановой геометрии равна нулю, $\nabla_\epsilon R_{\alpha\beta\gamma\delta} = 0$. Обратное, вообще говоря, неверно. У пространства постоянной кривизны тензор кривизны не обязательно имеет вид (16.18). Примером является полупростая группа Ли (см. раздел 9.5). Отсюда следует, что не всякое пространство постоянной кривизны является максимально симметричным. □

Пример 16.2.1. Рассмотрим евклидово пространство $\mathbb{R}^n$, на котором задана метрика нулевой кривизны, т.е. $R_{\alpha\beta\gamma\delta} = 0$. Ясно, что это пространство постоянной нулевой кривизны. Тогда в $\mathbb{R}^n$ существует такая система координат x^α, $\alpha = 1, \ldots, n$ в которой все компоненты метрики постоянны. В этой системе координат символы Кристоффеля равны нулю и уравнение для векторов Киллинга (16.8) принимает простой вид:

$$\partial^2_{\beta\gamma} K_\alpha = 0.$$

Общее решение этого уравнения имеет вид

$$K_\alpha(x) = a_\alpha + b_{\alpha\beta} x^\beta,$$

где a_α и $b_{\alpha\beta}$ – некоторые постоянные. Из уравнения Киллинга (16.4) следует, что это выражение задает форму Киллинга тогда и только тогда, когда матрица $b_{\alpha\beta}$ антисимметрична, т.е. $b_{\alpha\beta} = -b_{\beta\alpha}$. Следовательно, можно задать $\frac{1}{2} n(n+1)$ линейно независимых форм Киллинга:

$$K_\alpha{}^{(\gamma)}(x) = \delta^\gamma_\alpha,$$
$$K_\alpha{}^{[\gamma\delta]}(x) = \delta^\delta_\alpha x^\gamma - \delta^\gamma_\alpha x^\delta.$$

Тогда произвольная форма Киллинга выражается в виде линейной комбинации

$$K_\alpha = a_\gamma K_\alpha{}^{(g)} + \frac{1}{2} b_{\delta\gamma} K_\alpha{}^{[\gamma\delta]}.$$

В полученном выражении n векторов Киллинга $K^{(\gamma)}$ генерируют трансляции в $\mathbb{R}^n$ вдоль осей координат, а $\frac{1}{2}n(n-1)$ векторов $K^{[\gamma\delta]}$ – вращения вокруг начала координат. Таким образом, метрика пространства нулевой кривизны допускает максимальное число $\frac{1}{2}n(n+1)$ векторов Киллинга и поэтому является однородным и изотропным пространством.

Известно, что линейным преобразованием координат x^α метрику можно преобразовать к диагональному виду, когда на диагонали будут стоять ± 1, в зависимости от сигнатуры исходной метрики. Если метрика риманова, то после преобразования координат, она примет стандартный вид $g_{\alpha\beta} = \delta_{\alpha\beta}$. Эта метрика инвариантна относительно неоднородной группы вращений $\mathbb{IO}(n)$. □

16.3 Свойства векторных полей Киллинга

Векторные поля Киллинга обладают рядом замечательных свойств. Начнем с простейших.

Предложение 16.3.1. *Длина вектора Киллинга остается постоянной вдоль траектории Киллинга:*

$$K^\alpha \partial_\alpha K^2 = 0. \tag{16.19}$$

Доказательство. Свернем уравнения Киллинга (16.4) с $K^\alpha K^\beta$:

$$2K^\alpha K^\beta \nabla_\alpha K_\beta = K^\alpha \nabla_\alpha K^2 = 0.$$ □

Следствие. Если векторные поля Киллинга существуют на лоренцевом многообразии, то они имеют определенную ориентацию: времениподобную, светоподобную или пространственноподобную. □

Сравним траектории Киллинга с экстремалями [**?**].

Предложение 16.3.2. *Пусть $(\mathbb{M}, g)$ – (псевдо-)риманово многообразие с векторным полем Киллинга K. Траектории Киллинга являются экстремалями тогда и только тогда, когда их длина постоянна на $\mathbb{M}$, $K^2 = \mathsf{const}$.*

Доказательство. Рассмотрим траектории Киллинга $x^\alpha(t)$, которые определяются системой уравнений

$$\dot{x}^\alpha = K^\alpha. \tag{16.20}$$

Дифференцируя эти уравнения по параметру t, получим

$$\ddot{x}^\alpha = \partial_\beta K^\alpha \dot{x}^\beta = (\nabla_\beta K^\alpha - \Gamma_{\beta\gamma}{}^\alpha K^\gamma)\dot{x}^\beta.$$

С учетом уравнения (16.20) получим

$$\ddot{x}^\alpha = K^\beta \nabla_\beta K^\alpha - \Gamma_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma. \tag{16.21}$$

Уравнения Киллинга позволяют переписать первое слагаемое в виде

$$K^\beta \nabla_\beta K^\alpha = -\frac{1}{2} g^{\alpha\beta} \partial_\beta K^2.$$

Тогда уравнения (16.21) примут вид

$$\ddot{x}^\alpha = -\frac{1}{2} g^{\alpha\beta}\partial_\beta K^2 - \Gamma_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma.$$

□

Доказанное утверждение показывает, что далеко не каждая траектория Киллинга является экстремалью.

Пример 16.3.1. Рассмотрим евклидову плоскость $\mathbb{R}^2$ с евклидовой метрикой. Эта метрика инвариантна относительно трехпараметрической неоднородной группы вращений $\mathbb{IO}(2)$. Обозначим декартовы и полярные координаты на плоскости соответственно x, y и r, φ. Тогда векторные поля Киллинга имеют вид $K_1 = \partial_\varphi$ для вращений и $K_2 = \partial_x$, $K_3 = \partial_y$ для сдвигов. Квадраты длин векторов Киллинга равны:

$$K_1^2 = r^2, \qquad K_2^2 = K_3^2 = 1.$$

Векторы Киллинга K_2 и K_3 имеют постоянную длину, их траекториями Киллинга являются прямые линии, которые являются экстремалями. Это согласуется с предложением 16.3.2. Траекториями Киллинга для вращений K_1 являются концентрические окружности с центром в начале координат. Длина вектора Киллинга K_1 постоянна на траекториях в соответствии с предложением 16.3.1, однако непостоянна на всей плоскости $\mathbb{R}^2$. Соответствующие траектории Киллинга – окружности – не являются экстремалями. □

Пример 16.3.2. Рассмотри полупростую группу Ли $\mathbb{G}$, как (псевдо-)риманово пространство с формой Киллинга–Картана в качестве метрики (см. раздел 9.5). Это пространство постоянной кривизны. Левоинвариантные векторные поля генерируют групповые преобразования справа, а правоинвариантные – слева. Групповые преобразования слева и справа сохраняют метрику, и, следовательно, лево- и правоинвариантные векторные поля являются полями Киллинга. Длина этих полей Киллинга равна ± 1. Поэтому соответствующие траектории Киллинга являются экстремалями. □

Свертывая уравнения Киллинга (16.4) с метрикой, получаем, что дивергенция поля Киллинга равна нулю

$$\nabla_\alpha K^\alpha = 0. \tag{16.22}$$

Ковариантная производная ∇^β со связностью Леви–Чивита от уравнения Киллинга (16.4) с учетом уравнения (7.82) для перестановки ковариантных производных и уравнения (16.22) приводит к уравнению

$$\triangle K_\alpha = R_{\alpha\beta} K^\beta, \tag{16.23}$$

$\triangle = \nabla^\beta \nabla_\beta$ – оператор Лапласа–Бельтрами на многообразии $\mathbb{M}$. Для пространства постоянной кривизны тензор Риччи выражается через скалярную кривизну (16.17), и уравнение (16.23) принимает вид

$$\triangle K_\alpha = \frac{R}{n} K_\alpha, \qquad R = \mathsf{const}.$$

То есть каждая компонента формы Киллинга является собственным вектором оператора Лапласа–Бельтрами.

Предложение 16.3.3. *Пусть $X, Y \in \mathcal{X}(\mathbb{M})$ – два произвольных векторных векторных поля на (псевдо-)римановом многообразии $(\mathbb{M}, g)$ и K – векторное поле Киллинга. Тогда справедливо равенство*

$$g\Big((\mathsf{L}_K - \nabla_K)X, Y\Big) + g\Big(X, (\mathsf{L}_K - \nabla_K)Y\Big) = 0,$$

где $\mathsf{L}_K X = [K, X]$ – производная Ли и $\nabla_K X = K^\alpha(\partial_\alpha X^\beta + \Gamma_{\alpha\gamma}{}^\beta X^\gamma)\partial_\beta$ – ковариантная производная векторного поля X вдоль поля Киллинга K.

Доказательство. Прямая проверка с учетом явного выражения для символов Кристоффеля (7.22) и уравнения Киллинга (16.4). □

16.4 Лоренц-инвариантные метрики

Рассмотрим (псевдо-)риманово многообразие $(\mathbb{M}, g)$, которое допускает векторное поле Киллинга K. Пусть на этом многообразии заданы дополнительно какие либо тензорные поля. Тогда представляет интерес задача отыскания таких полей, которые также инвариантны относительно преобразований изометрии. Допустим, например, что на $\mathbb{M}$ задана функция $f \in \mathcal{C}^\infty(\mathbb{M})$. Эта функция инвариантна относительно инфинитезимальных преобразований, генерируемых векторным полем Киллинга K, если производная Ли равна нулю:

$$\mathsf{L}_K f = K^\alpha \partial_\alpha f = 0, \tag{16.24}$$

т.е. если производная f вдоль векторного поля K равна нулю.

Для произвольного тензорного поля T условие инвариантности имеет имеет вид

$$\mathsf{L}_K T = 0 \tag{16.25}$$

В качестве примера рассмотрим пространство Минковского $\mathbb{R}^{1,n-1}$ произвольной размерности n. В нем задана метрика Лоренца $\eta_{\alpha\beta}$ в декартовой системе координат x^α. Эта метрика инвариантна относительно преобразований Лоренца (см. раздел 2.9). Соответствующие векторные поля Киллинга имеют вид

$$K_{\epsilon\delta} = \frac{1}{2}\left(x_\delta \partial_\epsilon - x_\epsilon \partial_\delta\right) = \frac{1}{2}\left(x_\delta \delta^\gamma_\epsilon - x_\epsilon \delta^\gamma_\delta\right)\partial_\gamma, \tag{16.26}$$

где индексы ϵ, δ нумеруют $n(n-1)/2$ векторов Киллинга и $x_\alpha = x^\beta \eta_{\beta\alpha}$. Пусть в пространстве Минковского $\mathbb{R}^{1,n-1}$ задана вторая метрика $g_{\alpha\beta}$. Поставим следующую задачу: найти все метрики $g_{\alpha\beta}$, инвариантные относительно действия группы Лоренца $\mathbb{SO}(1, n-1)$.

Уравнения (16.6) для векторных полей Киллинга (16.26) принимают вид

$$g_{\alpha\epsilon}\eta_{\beta\delta} - g_{\alpha\delta}\eta_{\beta\epsilon} + g_{\beta\epsilon}\eta_{\alpha\delta} - g_{\beta\delta}\eta_{\alpha\epsilon} + x_\delta \partial_\epsilon g_{\alpha\beta} - x_\epsilon \partial_\delta g_{\alpha\beta} = 0. \tag{16.27}$$

При преобразованиях Лоренца компоненты метрики $g_{\alpha\beta}$ ведут себя как компоненты ковариантного тензора второго ранга. Их необходимо построить из метрики Лоренца $\eta_{\alpha\beta}$ и векторов $\{y^\alpha\}$. Единственная возможность – это метрика вида

$$g_{\alpha\beta} = A\eta_{\alpha\beta} + Bx^\alpha x^\beta,$$

где A и B – некоторые функции на $\mathbb{R}^{1,n-1}$. Подстановка этой метрики в уравнение Киллинга (16.27) ограничивает вид функций A и B. Можно доказать, что они могут быть произвольными функциями от одной переменной

$$s = x^\alpha x^\beta \eta_{\alpha\beta},$$

которая инвариантна относительно преобразований Лоренца. Таким образом, метрика, инвариантная относительно преобразований Лоренца, параметризуется двумя произвольными функциями $A(s)$ и $B(s)$. Ее удобно записать в несколько другом виде

$$g_{\alpha\beta} = f(s)\Pi^{\text{T}}_{\alpha\beta} + g(s)\Pi^{\text{L}}_{\alpha\beta} = f\eta_{\alpha\beta} + (g-f)\frac{x_\alpha x_\beta}{s}, \tag{16.28}$$

где Π^{T} и Π^{L} – проекционные операторы:

$$\Pi^{\text{T}}_{\alpha\beta} = \eta_{\alpha\beta} - \frac{x_\alpha x_\beta}{s},$$
$$\Pi^{\text{L}}_{\alpha\beta} = \frac{x_\alpha x_\beta}{s}.$$

На произвольные функции $f(s)$ и $g(s)$ наложено только одно условие $f(0) - g(0) = 0$, которое необходимо, чтобы метрика была определена при $s = 0$.

Здесь и до конца настоящего раздела, если не оговорено противное, для подъема и опускания индексов будет использоваться метрика Минковского.

Лоренц инвариантная метрика вида (16.28) при $f = g$ рассматривалась В. А. Фоком [**?**]. Метрике (16.28) соответствует инвариантный интервал

$$ds^2 = f dx_\alpha dx^\alpha + (g-f)\frac{(x_\alpha dx^\alpha)^2}{s}.$$

Метрический тензор (16.28) имеет одинаковый вид во всех системах координат, связанных между собой преобразованиями Лоренца. Однако его вид меняется при сдвигах $x^\alpha \to x^\alpha + a^\alpha$, поскольку метрика явно зависит от координат, и начало системы отсчета выделено.

Используя представление (28.10) для определителя суммы двух матриц и полагая

$$A_{\alpha\beta} = f\eta_{\alpha\beta}, \qquad B_{\alpha\beta} = (g-f)\frac{x_\alpha x_\beta}{s},$$

можно вычислить определитель метрики (16.28). В результате получим

$$\det g_{\alpha\beta} = (-f)^{n-1} g. \tag{16.29}$$

Таким образом лоренц инвариантная метрика вырождена тогда и только тогда, когда $fg = 0$. Мы будем предполагать, что функции f и g являются достаточно гладкими, и $f > 0$ и $g \neq 0$.

При $g > 0$ метрике (16.28) можно поставить в соответствие репер

$$e_\alpha{}^a = \sqrt{f}\delta^a_\alpha + (\sqrt{g} - \sqrt{f})\frac{x_\alpha x^a}{s}. \tag{16.30}$$

Он инвариантен относительно одновременного действия группы Лоренца на греческие и латинские индексы.

Нетрудно проверить следующие свойства проекционных операторов

$$\Pi^{\mathrm{T}\alpha\beta}x_\beta = 0, \qquad \Pi^{\mathrm{T}\,\alpha}_{\alpha} = n-1, \qquad \partial_\alpha \Pi^{\mathrm{T}}_{\beta\gamma} = -\frac{\Pi^{\mathrm{T}}_{\alpha\beta}x_\gamma + \Pi^{\mathrm{T}}_{\alpha\gamma}x_\beta}{s},$$
$$\Pi^{\mathrm{L}\alpha\beta}x_\beta = x^\alpha, \qquad \Pi^{\mathrm{L}\,\alpha}_{\alpha} = 1, \qquad \partial_\alpha \Pi^{\mathrm{L}}_{\beta\gamma} = \frac{\Pi^{\mathrm{T}}_{\alpha\beta}x_\gamma + \Pi^{\mathrm{T}}_{\alpha\gamma}x_\beta}{s},$$

которые будут использоваться при проведении вычислений.

Несложные вычисления приводят к следующему выражению для символов Кристоффеля, соответствующих метрике (16.28),

$$\Gamma_{\alpha\beta\gamma} = f'(x_\alpha\Pi^{\mathrm{T}}_{\beta\gamma} + x_\beta\Pi^{\mathrm{T}}_{\alpha\gamma} - x_\gamma\Pi^{\mathrm{T}}_{\alpha\beta}) + g'(x_\alpha\Pi^{\mathrm{L}}_{\beta\gamma} + x_\beta\Pi^{\mathrm{L}}_{\alpha\gamma} - x_\gamma\Pi^{\mathrm{L}}_{\alpha\beta}) + \frac{g-f}{s}x_\gamma\Pi^{\mathrm{T}}_{\alpha\beta},$$
$$\Gamma_{\alpha\beta}{}^{\gamma} = \frac{f'}{f}(x_\alpha\Pi^{\mathrm{T}\gamma}_{\beta} + x_\beta\Pi^{\mathrm{T}\gamma}_{\alpha}) + \frac{g'}{g}(x_\alpha\Pi^{\mathrm{L}\gamma}_{\beta} + x_\beta\Pi^{\mathrm{L}\gamma}_{\alpha} - x^\gamma\Pi^{\mathrm{L}}_{\alpha\beta}) + \frac{g-f-f's}{sg}x^\gamma\Pi^{\mathrm{T}}_{\alpha\beta}, \tag{16.31}$$

где штрих обозначает дифференцирование по аргументу s. Для подъема индекса в последнем выражении использовалась обратная метрика

$$g^{\alpha\beta} = \frac{1}{f}\Pi^{\mathrm{T}\alpha\beta} + \frac{1}{g}\Pi^{\mathrm{L}\alpha\beta}. \tag{16.32}$$

Тензор кривизны для метрики (16.28) имеет вид

$$\begin{aligned} R_{\alpha\beta\gamma}{}^{\delta} &= \Pi^{\mathrm{T}}_{\alpha\gamma}\Pi^{\mathrm{T}\delta}_{\beta}\left[\frac{(f+f's)^2}{sfg} - \frac{1}{s}\right] \\ &+ \Pi^{\mathrm{L}}_{\alpha\gamma}\Pi^{\mathrm{T}\delta}_{\beta}\left[2\left(\frac{f+f's}{f}\right)' + \left(\frac{f'}{f} - \frac{g'}{g}\right)\frac{f+f's}{f}\right] \\ &+ \Pi^{\mathrm{L}\,\delta}_{\alpha}\Pi^{\mathrm{T}}_{\beta\gamma}\left[-2\left(\frac{f+f's}{g}\right)' + \left(\frac{f'}{f} - \frac{g'}{g}\right)\frac{f+f's}{g}\right] - (\alpha \leftrightarrow \beta). \end{aligned} \tag{16.33}$$

Свернув это выражение по индексам β и δ, получим тензор Риччи

$$\begin{aligned} R_{\alpha\beta} &= \Pi^{\mathrm{T}}_{\alpha\beta}\left[\frac{n-2}{s}\left(\frac{(f+f's)^2}{fg} - 1\right) + 2\frac{(f+f's)'}{g} - \left(\frac{f'}{f} + \frac{g'}{g}\right)\frac{f+f's}{g}\right] \\ &+ \Pi^{\mathrm{L}}_{\alpha\beta}(n-1)\left[2\frac{(f+f's)'}{f} - \left(\frac{f'}{f} + \frac{g'}{g}\right)\frac{f+f's}{f}\right]. \end{aligned} \tag{16.34}$$

Дальнейшая свертка с обратной метрикой (16.32) дает скалярную кривизну

$$R = (n-1)\left[\frac{n-2}{fs}\left(\frac{(f+f's)^2}{fg} - 1\right) + 4\frac{(f+f's)'}{fg} - 2\left(\frac{f'}{f} + \frac{g'}{g}\right)\frac{f+f's}{fg}\right] \tag{16.35}$$

Как было отмечено ранее, пространства постоянной кривизны, определяемые уравнением (**??**), автоматически удовлетворяют вакуумным уравнениям Эйнштейна с космологической постоянной. Решим эти уравнения для лоренц инвариантной метрики (16.28). Для этого опустим последний индекс у тензора кривизны (16.33) с помощью метрики (16.28)

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= \Pi^{\mathrm{T}}_{\alpha\gamma}\Pi^{\mathrm{T}}_{\beta\delta}\frac{1}{s}\left[\frac{(f+f's)^2}{g} - f\right] \\ &+ (\Pi^{\mathrm{L}}_{\alpha\gamma}\Pi^{\mathrm{T}}_{\beta\delta} - \Pi^{\mathrm{L}}_{\alpha\delta}\Pi^{\mathrm{T}}_{\beta\gamma})\left[2(f+f's)' - \left(\frac{f'}{f} + \frac{g'}{g}\right)(f+f's)\right] - (\alpha \leftrightarrow \beta) \end{aligned} \tag{16.36}$$

и подставим в уравнение (**??**). В результате получим систему дифференциальных уравнений на функции f и g

$$\frac{(f+f's)^2}{sg} - \frac{f}{s} = -\frac{2K}{n(n-1)}f^2, \tag{16.37}$$

$$2(f+f's)' - \left(\frac{f'}{f} + \frac{g'}{g}\right)(f+f's) = -\frac{2K}{n(n-1)}fg. \tag{16.38}$$

Из первого уравнения получаем решение для g

$$g = \frac{(f+f's)^2}{f\left(1 - \frac{2K}{n(n-1)}fs\right)}. \tag{16.39}$$

Поскольку для невырожденности метрики необходимо, чтобы $g \neq 0$, то при $s \neq 0$ функция f должна удовлетворять неравенству

$$f \neq \frac{n(n-1)}{2Ks}, \qquad s \neq 0. \tag{16.40}$$

Подстановка этого выражения во второе уравнение (16.38) приводит к тождеству. Таким образом мы доказали следующее утверждение.

Теорема 16.4.1. *Лоренц инвариантная метрика*

$$g_{\alpha\beta} = f\Pi^{\mathrm{T}}_{\alpha\beta} + \frac{(f+f's)^2}{f\left(1 - \frac{2K}{n(n-1)}fs\right)}\Pi^{\mathrm{L}}_{\alpha\beta}, \tag{16.41}$$

где $f(s)$ – произвольная положительная функция, удовлетворяющая условию (16.40), является метрикой пространства постоянной кривизны. Обратно. Метрику пространства постоянной кривизны можно записать в лоренц инвариантном виде (16.41) для некоторой функции $f(s)$.

Доказательство. Нам осталось доказать, что произвольную метрику пространства постоянной кривизны можно привести к лоренц инвариантному виду (16.28). Чтобы ответить на этот вопрос, запишем метрику (16.41) в более известной форме. С этой целью зафиксируем функцию f, положив $f = g$. Это равенство с учетом (16.39) сводится к уравнению

$$f'^2 s + 2f'f + \frac{2K}{n(n-1)}f^3 = 0,$$

общее решение которого имеет вид

$$f = \frac{C}{(C + \frac{K}{2n(n-1)}s)^2}, \qquad C = \mathsf{const}.$$

Постоянная интегрирования убирается растяжкой координат. Поэтому без ограничения общности положим $C = 1$. В результате получим метрику постоянной кривизны

$$g_{\alpha\beta} = \frac{\eta_{\alpha\beta}}{(1 + \frac{K}{2n(n-1)}s)^2}. \tag{16.42}$$

То, что метрику пространства постоянной кривизны можно привести к такому виду – хорошо известный факт. Доказательство этого утверждения нетривиально (см., например, [**?**]). □

Замечание. Проведенные вычисления просто переносятся на случай метрики в евклидовом пространстве, которая инвариантна относительно $\mathbb{SO}(n)$-вращений. Для этого во всех формулах метрику Лоренца $\eta_{\alpha\beta}$ нужно заменить на евклидову метрику $\delta_{\alpha\beta}$. □

Поскольку предел функции $(g-f)/s$ равен

$$\lim_{s\to 0}\frac{g-f}{s} = 2f' + \frac{2K}{n(n-1)}f^2,$$

то выражение для метрики (16.41) определено и при $s=0$.

Решим вакуумные уравнения Эйнштейна с космологической постоянной

$$R_{\alpha\beta} = \Lambda g_{\alpha\beta}$$

для лоренц инвариантной метрики. Поскольку число этих уравнений меньше, чем число уравнений в условии постоянства кривизны (**??**), то можно было бы ожидать, что они допускают решения не только с постоянной кривизной. Однако для лоренц инвариантных метрик классы решений совпадают. Действительно, подстановка тензора Риччи (16.34) в уравнения Эйнштейна дает

$$\frac{n-2}{s}\left[\frac{(f+f's)^2}{fg}-1\right] + 2\frac{(f+f's)'}{g} - \left(\frac{f'}{f}+\frac{g'}{g}\right)\frac{f+f's}{g} = \Lambda f, \tag{16.43}$$

$$(n-1)\left[2\frac{(f+f's)'}{f} - \left(\frac{f'}{f}+\frac{g'}{g}\right)\frac{f+f's}{f}\right] = \Lambda g. \tag{16.44}$$

Второе уравнение при

$$\Lambda = -\frac{2K}{n}$$

совпадает с уравнением (16.38). Линейная комбинация (16.43)$/f-$(16.44)$/g$ эквивалентна уравнению (16.37).

Таким образом мы доказали, что все лоренц инвариантные решения вакуумных уравнений Эйнштейна с космологической постоянной исчерпываются пространствами постоянной кривизны.

В последнее время многие авторы понятие векторного поля Киллинга понимают в более широком смысле: как симметрию не метрики, а других тензорных полей.

Пример 16.4.1. Рассмотрим "векторное поле Киллинга" на плоскости, которое в декартовой системе координат u, v имеет вид

$$K = u\partial_u + av\partial_v, \qquad a = \mathsf{const} \neq 0.$$

Тогда уравнение Киллинга (16.24) для функции $f(u,v)$ принимает вид

$$Kf = 0. \tag{16.45}$$

Любое решение этого уравнения имеет вид $f = f(\xi)$, где

$$\xi = \frac{(av)^{1/a}}{u}.$$

Это значит, что решения уравнения (16.45) по существу являются функциями одной переменной ξ. Если функция f одновременно удовлетворяет некоторому уравнению в частных производных на поверхности, то его решения, зависящие от переменной ξ называются *автомодельными*. Для таких решений уравнение в частных производных сводится к обыкновенному дифференциальному уравнению относительно ξ, что существенно упрощает задачу. □

Глава 17

Геодезические и экстремали

Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия, т.е. метрика g и аффинная связность Γ. Тогда можно построить два типа выделенных кривых: геодезические и экстремали. Геодезические линии определяются только связностью как линии, касательный вектор к которым остается касательным при параллельном переносе. Экстремали, напротив, определяются только метрикой как линии экстремальной длины. Поскольку метрика и связность являются независимыми геометрическими объектами, то в общем случае геодезические линии и экстремали различны. В частном случае (псевдо-)римановой геометрии, когда связностью является связность Леви–Чивиты, геодезические и экстремали совпадают.

В настоящей главе мы рассмотрим оба типа кривых, т.к. они играют важную роль в моделях математической физики. Достаточно сказать, что одним из постулатов общей теории относительности является предположение о том, что свободные точечные частицы, подверженные действию только гравитационных сил, движутся по экстремалям. Кроме того, понятие полноты многообразий связано также с экстремалями.

17.1 Геодезические

В аффинной геометрии $(\mathbb{M}, g, \Gamma)$ существует выделенное семейство линий, которые называются геодезическими. Рассмотрим произвольную кривую $\gamma = x(t) = \{x^\alpha(t)\}$, где $-\infty \le t_1 < t < t_2 \le \infty$, на многообразии $\mathbb{M}$. Вектор скорости кривой, $u(t) = \{u^\alpha(t) := \dot{x}^\alpha(t)\}$, как всегда, предполагается отличным от нуля.

Определение. *Геодезической* линией на многообразии $\mathbb{M}$ называется кривая $x(t)$ класса $\mathcal{C}^2([t_1, t_2])$, касательный вектор к которой остается касательным при параллельном переносе вдоль нее. □

Замечание. В определении геодезической линии присутствует только аффинная связность. Поэтому понятие геодезической линии никакого отношения к метрике не имеет, которой может вообще не быть на многообразии. □

Получим уравнения, которым должны удовлетворять координатные функции $x^\alpha(t)$ для того, чтобы кривая γ была геодезической. Выберем произвольный отличный от нуля вектор X_0, который касается кривой γ в некоторой точке $x(t_0)$, и разнесем его вдоль всей кривой с помощью параллельного переноса. В результате получим векторное поле $X\Big(x(t)\Big)$, определенное на кривой γ. Множество векторных полей,

определенных на кривой γ, будем обозначать $\mathcal{X}(\mathbb{M},\gamma)$. Из определения геодезической следует, что это векторное поле всюду касается γ и, следовательно, пропорционально векторному полю скорости: $X^\alpha(t) = f(t)u^\alpha(t)$, где f – некоторая отличная от нуля функция на γ. Изменим параметризацию кривой $t \mapsto s(t)$. Тогда условие параллельности примет вид

$$X^\alpha = f\frac{ds}{dt}\frac{dx^\alpha}{ds}. \tag{17.1}$$

Выберем новый параметр вдоль геодезической таким образом, чтобы было выполнено уравнение

$$\frac{ds}{dt} = \frac{1}{f},$$

которое всегда имеет решение, поскольку $f \neq 0$. Таким образом, на геодезической линии существует такая параметризация, что вектор скорости u при параллельном переносе остается вектором скорости. В дальнейшем мы предполагаем, что параметр t вдоль геодезической выбран таким образом, что $f = 1$. Если вектор скорости геодезической при параллельном переносе остается касательным, то ковариантная производная от него вдоль геодезической равна нулю (7.28):

$$\nabla_u u = u^\alpha(\partial_\alpha u^\beta + \Gamma_{\alpha\gamma}{}^\beta u^\gamma)\partial_\beta = 0. \tag{17.2}$$

Поскольку $d/dt = u^\alpha\partial_\alpha$, то это уравнение в компонентах принимает вид

$$\ddot{x}^\alpha = -\Gamma_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma. \tag{17.3}$$

Таким образом, мы получили критерий того, что кривая $x(t)$ является геодезической.

Предложение 17.1.1. *Кривая $x(t)$ на многообразии $\mathbb{M}$ с заданной аффинной связностью Γ является геодезической тогда и только тогда, когда ее координатные функции $x^\alpha(t)$ удовлетворяют системе уравнений (17.3).*

Замечание. В уравнение (17.2) входит частная производная $\partial_\alpha u^\beta$ от вектора скорости. Эта производная неопределена, т.к. векторное поле скорости $u^\beta\Big(x(t)\Big)$ задано только вдоль кривой γ. Тем не менее в уравнение входит производная по направлению $u^\alpha\partial_\alpha u^\beta = \ddot{x}^\alpha$, которая имеет смысл. Это значит, что, для определения частной производной $\partial_\alpha u^\beta$, векторное поле скорости можно продолжить в некоторую окрестность кривой γ произвольным дифференцируемым образом, а конечный ответ от такого продолжения не зависит. Подобный трюк будет часто встречаться в дальнейшем. □

Определение. Векторное поле $a(t)$, определенное вдоль произвольной кривой $x(t)$ на многообразии $\mathbb{M}$ с заданной связностью Γ,

$$a := \nabla_u u = \Big(\ddot{x}^\alpha + \Gamma_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma\Big)\,\partial_\alpha, \tag{17.4}$$

называется *ускорением* кривой. □

Сравнивая ускорение кривой с уравнением для геодезической, получаем, что кривая будет геодезической тогда и только тогда, когда ее ускорение равно нулю.

Ускорение кривой, так же как и скорость, является векторным полем вдоль кривой, $u, a \in \mathcal{X}(\mathbb{M},\gamma)$, и зависит от ее параметризации. При выводе уравнений геодезической линии (17.3) была выбрана специальная параметризация кривой γ.

Определение. Параметр t, по которому проводится дифференцирование в системе обыкновенных дифференциальных уравнений (17.3), определяющих геодезическую линию, называется *каноническим* или *аффинным*. □

Поскольку уравнение (17.1) имеет тензорную форму, то канонический параметр не зависит от выбора системы координат на $\mathbb{M}$.

Предложение 17.1.2. *Канонический параметр вдоль геодезической определен с точностью до аффинного преобразования*

$$t \mapsto s = at + b, \qquad a = \mathsf{const} \neq 0, \qquad b = \mathsf{const}.$$

Доказательство. Рассмотрим вопрос как меняется уравнение для геодезических при произвольной замене канонического параметра. При другой параметризации геодезической $x^\alpha(s)$, где $s = s(t)$, $ds/dt \neq 0$, уравнение (17.3) изменится:

$$\left(\frac{ds}{dt}\right)^2 \frac{d^2x^\alpha}{ds^2} + \frac{d^2s}{dt^2}\frac{dx^\alpha}{ds} = -\left(\frac{ds}{dt}\right)^2 \Gamma_{\beta\gamma}{}^\alpha \frac{dx^\beta}{ds}\frac{dx^\gamma}{ds}. \tag{17.5}$$

Отсюда следует, что форма уравнений (17.3) не изменится тогда и только тогда, когда $d^2s/dt^2 = 0$, т.е. замена параметра является аффинной. □

Пример 17.1.1. В пространстве Минковского $\mathbb{R}^{1,3}$ точечная частица движется по некоторой мировой линии $x(t)$. Если в качестве параметра вдоль траектории выбрано время $t = x^0$, то скорость и ускорение кривой совпадают со скоростью и ускорением частицы. Если частица свободна, т.е. на нее не действуют никакие силы и, следовательно, ее ускорение равно нулю, то траекторией частицы будет одна из геодезических линий. Поскольку связность в $\mathbb{R}^{1,3}$ в декартовой системе координат имеет равные нулю компоненты, то уравнения (17.3) сводятся к уравнениям

$$\ddot{x}^\alpha = 0,$$

которые определяют прямые линии. Таким образом, в пространстве Минковского $\mathbb{R}^{1,3}$ прямые и только они являются геодезическими. Это значит, что свободная частица в пространстве Минковского равномерно движется вдоль прямой линии. □

Геодезическая линия в аффинной геометрии обобщает понятие прямой в (псевдо-)евклидовом пространстве, сохраняя то свойство, что касательный вектор остается касательным при параллельном переносе.

Пример 17.1.2. В пространствах абсолютного параллелизма, для которых тензор кривизны равен нулю $R_{\alpha\beta\gamma}{}^\delta = 0$, локально существует система координат, в которой симметричная часть аффинной связности равна нулю $\Gamma_{\{\alpha\beta\}}{}^\gamma = 0$ (см., раздел 7.7). Поскольку геодезические линии (17.3) определяются только симметричной частью аффинной связности, то в этой системе координат геодезические являются прямыми линиями. В частности, если на группе Ли задана каноническая связность, т.е. параллельный перенос отождествлен с групповым действием справа, то такая система координат локально существует. Заметим, что в правоинвариантном базисе на группе Ли компоненты связности равны нулю, однако он не является голономным и не определяет ту систему координат, о которой идет речь. □

Для двух параметризаций $x^\alpha(t)$ и $x^\alpha(s)$ одной геодезической справедливо равенство

$$\frac{d^2s}{dt^2} = \frac{ds}{dt}\frac{dx^\beta}{ds}\partial_\beta\left(\frac{ds}{dt}\right),$$

и уравнение (17.5) переписывается в эквивалентной форме

$$\begin{aligned}\frac{d^2x^\alpha}{ds^2} &= -\Gamma_{\beta\gamma}{}^\alpha\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} - \frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}\partial_\beta \ln\left|\frac{ds}{dt}\right| \\ &= -\Gamma_{\beta\gamma}{}^\alpha\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} - \frac{1}{2}\left(\delta^\alpha_\beta\partial_\gamma \ln\left|\frac{ds}{dt}\right| + \delta^\alpha_\gamma\partial_\beta \ln\left|\frac{ds}{dt}\right|\right)\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds}. \qquad (17.6)\end{aligned}$$

Будем говорить, что две геодезические линии $x^\alpha(t)$ и $x^\alpha(s)$ для различных связностей $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$ на одном и том же многообразии $\mathbb{M}$ имеют одинаковую форму, если множества точек, через которые проходят геодезические линии, совпадают. Тогда из уравнения (17.6) следует критерий совпадения формы геодезических для двух аффинных связностей.

Теорема 17.1.1. *Для того, чтобы форма геодезических $x^\alpha(t)$ и $x^\alpha(s)$ для двух аффинных связностей $\Gamma_{1\alpha\beta}{}^\gamma$ и $\Gamma_{2\alpha\beta}{}^\gamma$ на многообразии $\mathbb{M}$ совпадала, необходимо и достаточно, чтобы симметричные части этих связностей были связаны соотношением*

$$\Gamma_{2\{\alpha\beta\}}{}^\gamma = \Gamma_{1\{\alpha\beta\}}{}^\gamma + \frac{1}{2}\left(\delta^\gamma_\beta\partial_\alpha\varphi + \delta^\gamma_\alpha\partial_\beta\varphi\right), \qquad (17.7)$$

где $\varphi = \varphi(x)$ – некоторая функция, определяющая связь канонических параметров для каждой геодезической линии

$$\frac{ds}{dt} = e^\varphi. \qquad (17.8)$$

Пусть задан диффеоморфизм многообразий

$$f: \quad \mathbb{M} \to \mathbb{N}$$

такой, что аффинная связность на $\mathbb{M}$ отображается в аффинную связность на $\mathbb{N}$. Тогда диффеоморфизм f отображает геодезическую линию в геодезическую, поскольку определение геодезической линии инвариантно (не зависит от системы координат).

Уравнения для геодезических (17.3) – это система нелинейных обыкновенных дифференциальных уравнений второго порядка. Поэтому при достаточно гладких компонентах связности для однозначного решения задачи Коши необходимо задать начальную точку $\{x^\alpha(0)\}$ и вектор скорости $\{\dot{x}^\alpha(0)\}$. Геометрически это означает, что через данную точку многообразия в данном направлении проходит одна и только одна геодезическая.

Предложение 17.1.3. *Если аффинная связность Γ на многообразии $\mathbb{M}$ гладкая, то геодезические линии также являются гладкими $\mathcal{C}^\infty$ кривыми.*

Доказательство. Продифференцируем уравнение геодезических (17.3) по каноническому параметру. Правая часть полученного равенства зависит от компонент связности и их первых производных, а также от компонент вектора скорости и их первой производной. Поэтому правая часть определена и непрерывна. Следовательно третья производная от координатных функций $x^\alpha(t)$ существует и непрерывна. Дальнейшее дифференцирование приводит к гладкости геодезических. □

Для геодезической линии можно также поставить краевую задачу: найти геодезическую, соединяющую две фиксированные точки многообразия, которое предполагается линейно связным. Эта задача разрешима в малом, т.е. любые две достаточно близкие точки можно соединить геодезической и притом только одной. Для удаленных точек эта задача может не иметь решения или иметь несколько решений.

Определение. (Псевдо-)риманово многообразие $\mathbb{M}$ называется *геодезически выпуклым*, если любые две его точки могут быть соединены геодезической.

Пример 17.1.3. Рассмотрим плоскость с конической сингулярностью и положительным углом дефицита $\mathbb{M} = \mathbb{U} \cup \mathbb{V}$, где $\mathbb{U}$ – область, изометричная евклидовой плоскости $\mathbb{R}^2$ с выколотой полупрямой, и $\mathbb{V}$ – клин евклидовой плоскости, который вставлен, рис.17.1. Тогда на ней существуют точки, которые нельзя соединить геодезической. Действительно, все геодезические, проходящие через точку p, соединяют ее со всеми точками евклидовой плоскости $\mathbb{R}^2$, до того, как клин был вставлен. Поскольку при диффеоморфизме ($\mathbb{R}^2 \setminus \overline{\mathbb{R}}_+$ на $\mathbb{U}$) геодезическая переходит в геодезическую и их число сохраняется, то точку p нельзя соединить геодезической ни с какой точкой q, лежащей на клине $\mathbb{V}$, который вставлен со стороны, противоположной конической сингулярности. Это многообразие не является геодезически выпуклым. □

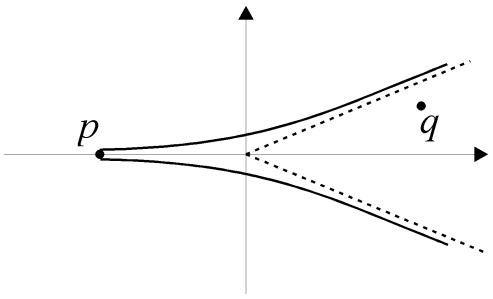


Рис. 17.1: Точки p и q на плоскости с конической сингулярностью и положительным углом дефицита нельзя соединить геодезической.

Пример 17.1.4. Рассмотрим сферу, вложенную в евклидово пространство, $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Пусть метрика на сфере индуцирована вложением и на ней задана связность Леви–Чивиты. Тогда любые две точки на сфере $\mathbb{S}^2$, не являющиеся диаметрально противоположными, можно соединить двумя разными геодезическими (две дуги большой окружности, проходящей через эти точки). Диаметрально противоположные точки соединяются бесконечным числом геодезических. Сфера $\mathbb{S}^2$ является геодезически выпуклым многообразием. □

Введем важное понятие полноты геодезической по каноническому параметру t, В силу однозначного разрешения задачи Коши, через каждую точку многообразия в заданном направлении проходит одна геодезическая. Это значит, что геодезическую линию, если она заканчивается в некоторой точке $q \in \mathbb{M}$, всегда можно продолжить. Действительно, если при конечном значении канонического параметра геодезическая попадает в точку q, то продолжим ее, склеив с геодезической, выходящей из точки q в том же направлении.

Определение. Геодезическая в $\mathbb{M}$ называется *полной*, если ее можно продолжить в обе стороны до бесконечного значения канонического параметра, $t \in (-\infty, \infty)$. □

Замечание. Поскольку канонический параметр определен с точностью до аффинного преобразования и не зависит от выбора системы координат, то данное определение корректно. □

Геодезическую линию можно рассматривать как интегральную кривую векторного поля скорости $u := \{\dot{x}^\alpha\}$. Тогда полнота геодезической означает полноту векторного поля u.

Пусть на многообразии $\mathbb{M}$ задана не только аффинная связность Γ, но и метрика g. Тогда компоненты аффинной связности можно выразить через метрику, кручение и неметричность (7.15). Хотя в уравнение для геодезических входит только симметричная часть связности, тем не менее оно нетривиально зависит от кручения и неметричности. Действительно, из формулы (7.15) следует, что симметричная часть аффинной связности имеет вид

$$\begin{aligned}\Gamma_{\{\beta\gamma\}}{}^\alpha &= \frac{1}{2}(\Gamma_{\beta\gamma}{}^\alpha + \Gamma_{\gamma\beta}{}^\alpha) \\ &= \widetilde{\Gamma}_{\beta\gamma}{}^\alpha + \frac{1}{2}(T^\alpha{}_{\beta\gamma} + T^\alpha{}_{\gamma\beta}) + \frac{1}{2}(Q_{\beta\gamma}{}^\alpha + Q_{\gamma\beta}{}^\alpha - Q^\alpha{}_{\beta\gamma}),\end{aligned} \tag{17.9}$$

где $\widetilde{\Gamma}_{\beta\gamma}{}^\alpha$ – символы Кристоффеля. Ясно, что две связности, имеющие одинаковую симметричную часть определяют одно и то же семейство геодезических.

Рассмотрим два вектора X и Y, которые параллельно переносятся вдоль геодезической γ.

Предложение 17.1.4. *Зависимость скалярного произведения (X, Y) двух векторов, которые параллельно переносятся вдоль γ, от точки геодезической определяется только тензором неметричности:*

$$\partial_u(X, Y) = \nabla_u(X^\alpha Y^\beta g_{\alpha\beta}) = -u^\gamma X^\alpha Y^\beta Q_{\gamma\alpha\beta}. \tag{17.10}$$

Отсюда следует, что в римановой геометрии и геометрии Римана–Картана, где $Q = 0$, скалярное произведение двух векторов при параллельном переносе вдоль геодезической сохраняется. В частности, квадрат вектора скорости геодезической постоянен вдоль нее.

Доказательство. Следует из определения тензора неметричности (7.14). □

Замечание. Это утверждение верно и для произвольной кривой (предложение 7.4.1). Скалярное произведение двух векторных полей, полученных в результате параллельного переноса двух векторов вдоль произвольной кривой γ в римановой геометрии и геометрии Римана–Картана, не зависит от точки кривой. □

Следствие. Если неметричность лоренцева многообразия $\mathbb{M}$ равна нулю, то квадрат вектора скорости постоянен вдоль геодезических, и их можно разделить на три класса: времениподобные, пространственноподобные и светоподобные (изотропные). При этом тип геодезической не может меняться от точки к точке. □

В заключение настоящего раздела получим уравнение девиации геодезических. Пусть на многообразии $\mathbb{M}$ задана произвольная кривая $\sigma = y(s)$, $s \in [s_1, s_2]$ и отличное от нуля векторное поле $X(s) \in \mathcal{X}(\mathbb{M}, \sigma)$, определенное вдоль этой кривой и которое нигде не касается кривой σ. Рассмотрим семейство геодезических $\gamma_s = x(t, s)$, проходящих через каждую точку кривой σ в направлении $X(s)$:

$$x(0, s) = y(s), \qquad \dot{x}(0, s) = X(s),$$

где точка, как и раньше, обозначает дифференцирование по каноническому параметру t. Если кривая $y(s)$ и векторное поле $X(s)$ достаточно гладкие, то множество точек $x(t,s)$ образует двумерное подмногообразие (поверхность) Σ в $\mathbb{M}$ при достаточно малых t. При этом параметры t, s можно выбрать в качестве координат на Σ. Этим координатам соответствуют векторные поля

$$u := \partial_t = \frac{\partial x^\alpha}{\partial t}\partial_\alpha, \qquad Y := \partial_s = \frac{\partial x^\alpha}{\partial s}\partial_\alpha,$$

которые коммутируют между собой, $[u, Y] = 0$. После замены в коммутаторе частных производных на ковариантные, получим равенство

$$u^\beta \nabla_\beta Y^\alpha = Y^\beta \nabla_\beta u^\alpha + u^\beta Y^\gamma T_{\beta\gamma}{}^\alpha, \tag{17.11}$$

где $T_{\beta\gamma}{}^\alpha = \Gamma_{\beta\gamma}{}^\alpha - \Gamma_{\gamma\beta}{}^\alpha$ – тензор кручения.

Определение. Векторное поле $Y \in \mathcal{X}(\mathbb{M}, \Sigma)$ называется *вектором девиации* геодезических. Векторные поля с компонентами:

$$V^\alpha = u^\beta \nabla_\beta Y^\alpha, \tag{17.12}$$

$$A^\alpha = u^\beta \nabla_\beta V^\alpha, \tag{17.13}$$

называются соответственно *скоростью* и *ускорением девиации* близких геодезических. □

Векторное поле девиации определяет расположение близких геодезических на поверхности Σ, а векторные поля скорости и ускорения девиации характеризуют, как меняются расположение геодезических при движении вдоль геодезической.

Предложение 17.1.5. *Ускорение девиации геодезических определяется тензором кривизны и кручения:*

$$A^\alpha = u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\gamma u^\delta \nabla_\delta \left(Y^\beta T_{\gamma\beta}{}^\alpha \right). \tag{17.14}$$

Доказательство. Прямая проверка:

$$\begin{aligned}
&u^\gamma \nabla_\gamma (u^\beta \nabla_\beta Y^\alpha) = u^\gamma \nabla_\gamma (Y^\beta \nabla_\beta u^\alpha + u^\beta Y^\delta T_{\beta\delta}{}^\alpha) = \\
&= u^\gamma \nabla_\gamma Y^\beta \nabla_\beta u^\alpha + u^\gamma Y^\beta \nabla_\gamma \nabla_\beta u^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\
&= u^\gamma \nabla_\gamma Y^\beta \nabla_\beta u^\alpha + u^\gamma Y^\beta \nabla_\beta \nabla_\gamma u^\alpha + u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha - u^\gamma Y^\beta T_{\gamma\beta}{}^\delta \nabla_\delta u^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\
&= Y^\gamma \nabla_\gamma (u^\beta \nabla_\beta u^\alpha) + u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha) = \\
&= u^\gamma u^\delta Y^\beta R_{\gamma\beta\delta}{}^\alpha + u^\beta u^\gamma \nabla_\gamma (Y^\delta T_{\beta\delta}{}^\alpha),
\end{aligned}$$

где мы учли равенство (17.11), уравнения геодезических (17.2) и выражение коммутатора ковариантных производных через тензоры кривизны и кручения (7.81). □

В (псевдо-)римановой геометрии уравнение (17.14) упрощается

$$A^\alpha = u^\gamma u^\delta Y^\beta \widetilde{R}_{\gamma\beta\delta}{}^\alpha. \tag{17.15}$$

В литературе по общей теории относительности оно известно, как *уравнение девиации геодезических*. Это уравнение можно переписать в виде

$$\nabla_u^2 Y^\alpha + u^\delta u^\beta Y^\gamma \widetilde{R}_{\gamma\delta\beta}{}^\alpha = 0.$$

Оно совпадает с равенством нулю второй вариации действия для экстремалей (17.33), которое будет получено позже.

17.2 Экстремали

Другим выделенным типом кривых в аффинной геометрии $(\mathbb{M}, g, \Gamma)$ являются экстремали, которые определяются как линии экстремальной длины. Рассмотрим произвольную достаточно гладкую кривую $\gamma = x(t) \in \mathbb{M}$, $t \in [t_1, t_2]$. Предположим, что квадрат вектора скорости кривой, $u^\alpha := \dot{x}^\alpha(t)$, всюду отличен от нуля, $u^2 := u^\alpha u^\beta g_{\alpha\beta} \neq 0$.

Замечание. Для римановой метрики это условие автоматически выполняется, поскольку вектор скорости предполагается отличным от нуля. Если метрика не является знакоопределенной, то это условие нетривиально. Например, для лоренцевой метрики это условие равносильно тому, что мы рассматриваем либо времени-, либо пространственноподобные кривые. □

Пусть кривая соединяет две точки $p = \{x^\alpha(t_1)\}$ и $q = \{x^\alpha(t_2)\}$. Тогда длина этой кривой задается интегралом

$$S = \int_p^q ds, \qquad ds = dt\sqrt{|g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta|} = dt\sqrt{|u^2|}. \tag{17.16}$$

Этот функционал инвариантен относительно общих преобразований координат x^α и произвольной перепараметризации кривой $t \to \tau(t)$.

Определение. *Экстремалью*, соединяющей две точки (псевдо-)риманова многообразия $p, q \in \mathbb{M}$, если она существует, называется неизотропная кривая γ класса $\mathcal{C}^2([t_1, t_2])$, для которой функционал (17.16) принимает экстремальное значение. □

Замечание. Экстремали в римановом пространстве обобщают понятие прямой в евклидовом пространстве, сохраняя свойство быть линиями минимальной (экстремальной) длины, соединяющей две точки. □

Пример 17.2.1. Не любые две точки линейно связного многообразия $p, q \in \mathbb{M}$, на котором задана риманова метрика, можно соединить кривой наименьшей длины. Рассмотрим кольцо на евклидовой плоскости $\mathbb{R}^2$, рис.17.2. Это – риманово некомпактное связное неодносвязное двумерное многообразие. Если отрезок, соединяющий точки p и q, пересекает вырезанную дырку, то эти точки нельзя соединить кривой наименьшей длины. Действительно, для любой кривой γ, соединяющей точки p и q, всегда найдется кривая γ' меньшей длины. Построение ясно из рисунка. □

Замечание. Если метрика на многообразии $\mathbb{M}$ не является знакоопределенной, то существуют изотропные кривые, для которых функционал длины (17.16) равен нулю и данное выше определение экстремалей не проходит. Определение изотропных экстремалей будет дано ниже. □

Найдем уравнения, которым должны удовлетворять координатные функции $x^\alpha(t)$ для того, чтобы кривая $x(t)$ была экстремалью. Пусть экстремаль соединяет две точки p и q многообразия. Выберем такую карту на многообразии, которая целиком содержит данную экстремаль. Для этого достаточно взять объединение всех достаточно малых шаров, центры которых лежат на экстремали. Пусть в этой карте экстремаль и ее вариация задаются набором функций $x^\alpha(t)$ и $\delta x^\alpha(t)$. Мы предполагаем, что вариации кривой в конечных точках $p, q \in \mathbb{M}$ равны нулю, $\delta x^\alpha(p) = \delta x^\alpha(q) = 0$,

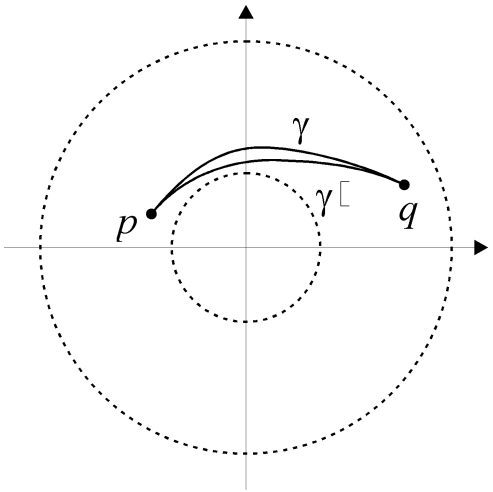


Рис. 17.2: Точки p и q нельзя соединить экстремалью.

и поэтому вкладом граничных членов при интегрировании по частям можно пренебречь. Вариация длины кривой (17.16) с точностью до знака имеет вид

$$\delta S = \int \frac{dt}{2\sqrt{|u^2|}} \left[2\delta(\dot{x}^\alpha)\dot{x}^\beta g_{\alpha\beta} + \dot{x}^\alpha \dot{x}^\beta \delta g_{\alpha\beta} \right].$$

Проинтегрируем первое слагаемое по частям и воспользуемся равенством

$$\delta g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} \delta x^\alpha.$$

Тогда вариация длины кривой принимает вид

$$\delta S = -\int dt \left[\frac{d}{dt}\left(\frac{\dot{x}^\beta g_{\alpha\beta}}{\sqrt{|u^2|}} \right) - \frac{1}{2\sqrt{|u^2|}} \dot{x}^\beta \dot{x}^\gamma \partial_\alpha g_{\beta\gamma} \right] \delta x^\alpha.$$

Поскольку

$$ds = \sqrt{|u|^2}dt, \quad \text{и} \quad \dot{x}^\alpha = \sqrt{|u^2|}\frac{dx^\alpha}{ds},$$

то вариацию длины кривой можно переписать в виде

$$\delta S = -\int ds \left(\frac{d^2x^\beta}{ds^2} + \widetilde{\Gamma}_{\gamma\delta}{}^{\beta} \frac{dx^\gamma}{ds}\frac{dx^\delta}{ds} \right) g_{\alpha\beta}\delta x^\alpha,$$

где $\widetilde{\Gamma}_{\beta\gamma}{}^{\alpha}$ – символы Кристоффеля (7.22). Фактически, на этом этапе мы воспользовались инвариантностью интеграла (17.16) относительно перепараметризации кривой, выбрав длину кривой в качестве параметра, $t \to s$. Параметр s вдоль экстремали называется *каноническим*. В дальнейшем канонический параметр мы будем обозначать буквой t. Таким образом, мы доказали следующее утверждение.

Теорема 17.2.1. *Для того, чтобы кривая $x(t)$ была экстремалью, необходимо и достаточно, чтобы координатные функции $x^\alpha(t)$ удовлетворяли системе обыкновенных дифференциальных уравнений*

$$\ddot{x}^\alpha = -\widetilde{\Gamma}_{\beta\gamma}{}^{\alpha} \dot{x}^\beta \dot{x}^\gamma, \tag{17.17}$$

где точка обозначает дифференцирование по каноническому параметру t.

Замечание. Параметр t, по которому проводится дифференцирование в уравнении (17.17) также, как и для геодезических, называется каноническим или аффинным. Он определен с точностью до аффинного преобразования. Таким образом, канонический параметр в общем случае не равен, а пропорционален длине экстремали. □

Пример 17.2.2. Пусть в некоторой области $\mathbb{U} \subset \mathbb{M}$ (псевдо-)риманова многообразия существует такая система координат, в которой метрика имеет постоянные компоненты. Тогда в этой области символы Кристоффеля равны нулю и, следовательно, уравнения (17.17) принимают вид $\ddot{x}^\alpha = 0$. Поэтому прямые и только они являются экстремалями в рассматриваемой области $\mathbb{U}$. □

Мы получили критерий того, что неизотропная кривая является экстремалью. При выводе уравнений (17.17) из вариационного принципа для действия (17.16) существенно используется условие $u^2 \neq 0$, которое исключает изотропные (светоподобные) экстремали. Поэтому изотропные экстремали определим не с помощью функционала длины, а непосредственно уравнениями (17.17). Для любой изотропной кривой $u^2 = 0$ и интеграл (17.16) равен нулю. В то же время уравнения (17.17) имеют смысл.

Определение. *Изотропной экстремалью* называется изотропная кривая γ класса $\mathcal{C}^2([t_1, t_2])$, которая задана функциями $x^\alpha(t)$, удовлетворяющими системе обыкновенных дифференциальных уравнений (17.17). □

Это определение корректно, т.к. квадрат вектора скорости экстремали постоянен. При этом не всякая изотропная кривая является экстремалью.

Пример 17.2.3. В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ любая изотропная кривая имеет вид (2.179). В то же время экстремалями являются прямые линии и только они. Следовательно, не всякая изотропная кривая является экстремалью. Отметим, что в двумерном пространстве-времени любая изотропная кривая является экстремалью, что нетрудно доказать в изотермических координатах (конформной калибровке). □

Определения изотропных и неизотропных экстремалей можно объединить.

Определение. Кривая γ класса $\mathcal{C}^2([t_1, t_2])$ называется *экстремалью*, если она задается функциями $x^\alpha(t)$, удовлетворяющими системе дифференциальных уравнений (17.17). Другими словами, экстремалью на (псевдо-)римановом многообразии $(\mathbb{M}, g)$ называется геодезическая для связности Леви-Чивиты. □

Замечание. Мы начали с другого определения экстремалей, т.к. в последнем определении теряется основное свойство неизотропных экстремалей реализовывать экстремумы функционала длины (17.16). □

Уравнение для экстремалей (17.17) можно переписать в явно ковариантном виде

$$\dot{x}^\beta \widetilde{\nabla}_\beta \dot{x}^\alpha = 0. \tag{17.18}$$

Это просто уравнение (17.2) для связности Леви–Чивиты. Опустив в этом уравнении индекс α и расписав ковариантную производную, получим альтернативную форму уравнения для экстремалей

$$\frac{\partial}{\partial t}(g_{\alpha\beta}\dot{x}^\beta) - \frac{1}{2}\partial_\alpha g_{\beta\gamma}\dot{x}^\beta \dot{x}^\gamma = 0. \tag{17.19}$$

Уравнения (17.17) определяются только метрикой, поскольку функционал длины не зависит от аффинной связности. Тем самым экстремали не зависят от того заданы ли на многообразии $\mathbb{M}$ тензоры кручения и неметричности или нет. Сравнение уравнений (17.17) с уравнением для геодезических (17.3) показывает, что экстремали являются геодезическими линиями по отношению к параллельному переносу, определяемому символами Кристоффеля. Это означает, что все свойства геодезических справедливы также и для экстремалей. В частности, канонический параметр вдоль экстремали инвариантен относительно преобразования координат. При произвольной параметризации уравнение для экстремалей имеет вид (17.5).

Из предложения 17.1.3 следует, что если метрика g на многообразии является гладкой функцией, то экстремали являются гладкими кривыми класса $\mathcal{C}^\infty$.

Через произвольную точку $x \in \mathbb{M}$ в направлении X^α проходит одна и только одна экстремаль, поскольку она определяется системой обыкновенных дифференциальных уравнений второго порядка (17.17). Это значит что любую экстремаль, которая заканчивается в некоторой точке $q \in \mathbb{M}$ можно продолжить. Действительно, если она заканчивается в точке $q \in \mathbb{M}$ при значении канонического параметра t_2, то существует единственная экстремаль, проходящая через q и имеющая тот же касательный вектор.

Определение. Экстремаль γ в $\mathbb{M}$ называется *полной*, если ее можно продолжить до бесконечного значения канонического параметра в обе стороны, $t \in (-\infty, \infty)$. □

Замечание. При продолжении экстремали возможны два случая. Во-первых, она может оказаться полной и иметь бесконечную длину. К полным экстремалям мы относим также и замкнутые экстремали, которые имеют конечную длину. Хотя их длина конечна, но канонический параметр продолжается до бесконечности, что соответствует бесконечному числу проходов вдоль экстремали. Во-вторых, при конечном значении канонического параметра экстремаль может попасть в такую точку многообразия, в которой один из геометрических инвариантов, например, скалярная кривизна, неопределен. Эта точка является сингулярной, и продолжение экстремали через нее не имеет смысла. □

Очевидно, что в (псевдо-)римановой геометрии экстремали и геодезические совпадают, поскольку совпадают уравнения (17.3) и (17.17) при $\Gamma_{\alpha\beta}{}^\gamma = \widetilde{\Gamma}_{\alpha\beta}{}^\gamma$. В общем случае экстремали и геодезические не совпадают и их поведение должно исследоваться отдельно. Критерий совпадения экстремалей и геодезических дает следующая

Теорема 17.2.2. *В аффинной геометрии* $(\mathbb{M}, g, \Gamma)$ *экстремали и геодезические совпадают тогда и только тогда, когда кручение и неметричность удовлетворяют соотношению*

$$\frac{1}{2}(T_{\alpha\beta\gamma} - T_{\gamma\alpha\beta}) = Q_{\alpha\beta\gamma}. \tag{17.20}$$

Доказательство. Экстремали и геодезические совпадают если и только если симметричная часть компонент связности равна символам Кристоффеля, $\Gamma_{\{\beta\gamma\}}{}^\alpha = \widetilde{\Gamma}_{\beta\gamma}{}^\alpha$. Тогда из (17.9) следует соотношение между тензором кручения и неметричностью:

$$T_{\alpha\beta\gamma} - T_{\gamma\alpha\beta} = -Q_{\beta\gamma\alpha} - Q_{\gamma\alpha\beta} + Q_{\alpha\beta\gamma}. \tag{17.21}$$

Сделав циклическую перестановку индексов $\alpha\beta\gamma \mapsto \beta\gamma\alpha$ и сложив полученное уравнение с (17.21) получим (17.20). Обратно. Подстановка тензора неметричности (17.20) в уравнение (17.21), как легко проверить, приводит к тождеству. □

Следствие. В геометрии Римана–Картана, где $Q = 0$, экстремали и геодезические совпадают тогда и только тогда, когда тензор кручения с опущенным индексом антисимметричен по всем трем индексам, $T_{\alpha\beta\gamma} = T_{[\alpha\beta\gamma]}$.

17.3 Интегрирование уравнений для экстремалей и геодезических

Уравнения для экстремалей и геодезических в ряде случаев имеют первые интегралы, наличие которых существенно упрощает их исследование. Начнем с универсального закона сохранения. Из сравнения уравнений (17.17) и (17.3) следует, что экстремаль является геодезической линией для связности, определяемой символами Кристоффеля. Поскольку при параллельном переносе и в римановой геометрии, и в геометрии Римана–Картана длины векторов не меняется, то отсюда сразу следует, что длина вектора скорости $u = \{\dot{u}^\alpha\}$ вдоль экстремалей и геодезических постоянна.

Предложение 17.3.1. *В геометрии Римана–Картана и (псевдо-)римановой геометрии для уравнений геодезических (17.3) и экстремалей (17.17) существует первый интеграл*

$$C_0 = u^2 := \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta} = \mathsf{const}, \tag{17.22}$$

квадратичный по первым производным (скоростям).

Доказательство. Следствие предложения 17.1.4. Можно доказать и формально, продифференцировав уравнение (17.22) по каноническому параметру и воспользовавшись уравнением для геодезических или экстремалей. □

Рассмотрим как меняется квадрат длины касательного вектора $C_0(t) = \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}$ вдоль геодезической линии в аффинной геометрии общего вида. Дифференцируя это соотношение по каноническому параметру и используя уравнение для геодезических, получим

$$\frac{du^2}{dt} = \dot{x}^\alpha \dot{x}^\beta \dot{x}^\gamma Q_{\alpha\beta\gamma}. \tag{17.23}$$

Отсюда следует, что $C_0 = \mathsf{const}$ при нулевом тензоре неметричности. В геометрии Римана–Картана–Вейля уравнение (17.23) принимает вид

$$\frac{du^2}{dt} = \dot{x}^\alpha W_\alpha C_0(t). \tag{17.24}$$

Первый интеграл (17.22) имеет кинематический характер и существует для любой экстремали и геодезической в (псевдо-)римановой геометрии и геометрии Римана–Картана. Если метрика имеет лоренцеву сигнатуру, то экстремали и геодезические можно разделить на три класса: времениподобные, $C_0 > 0$, изотропные или светоподобные, $C_0 = 0$, и пространственноподобные, $C_0 < 0$. Поскольку канонический параметр определен с точностью до аффинных преобразований, то для времениподобных и пространственноподобных экстремалей его всегда можно выбрать таким образом, что $C_0 = \pm 1$. В этом случае для времениподобных экстремалей канонический параметр называется *собственным временем*, а для пространственноподобных – *длиной* экстремали.

Замечание. Отметим, что, если некоторая кривая, не обязательно экстремаль или геодезическая, имеет определенный тип, то вдоль нее параметр всегда можно выбрать таким образом, что будет выполнено условие (17.22). Для изотропных кривых равенство (17.22), очевидно, выполняется. Допустим, что кривая имеет определенный тип в некоторой области, т.е. $C_0(t) \neq 0$. Тогда, вводя новый параметр $s(t)$, получим

$$C_0(t) = \left(\frac{ds}{dt}\right)^2 \dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta},$$

где точка обозначает дифференцирование по s. Уравнение

$$\frac{ds}{dt} = \sqrt{|C_0|}$$

всегда имеет решение. Поэтому условие (17.22) будет выполнено относительно нового параметра s. □

Существование других первых интегралов связано с инфинитезимальными симметриями метрики, которые определяются векторными полями Киллинга.

Предложение 17.3.2. *Если метрика на многообразии имеет один или несколько векторов Киллинга $K_i = \{K_i^\alpha\}$, $i = 1, \ldots, m$, то для каждого вектора Киллинга имеется свой интеграл движения и для экстремалей, и для геодезических*

$$C_i = K_i^\alpha \dot{x}^\beta g_{\alpha\beta} = \mathsf{const}, \qquad i = 1, \ldots, m, \tag{17.25}$$

который линеен по компонентам скорости.

Доказательство. Дифференцируем соотношения (17.25) по каноническому параметру и используем уравнения (17.3) или (17.17). □

17.4 Вторая вариация уравнений для экстремалей

Допустим, что метрика положительно определена, т.е. многообразие риманово. Экстремали на (псевдо-)римановом многообразии $(\mathbb{M}, g)$ определяются действием (17.16). Тогда любая экстремаль удовлетворяет уравнениям Эйлера–Лагранжа (17.17). Это свойство является необходимым условием для того, чтобы экстремаль была линией наименьшей или наибольшей длины. Для того, чтобы найти достаточное условие реализации минимума или максимума функционалов (17.16) или (17.37) необходимо исследовать вторую вариацию функционала.

Напомним общее определение второй вариации. Пусть задана достаточно гладкая кривая $\gamma = \{x^\alpha(t)\} \in \mathbb{M}$, где $t \in [t_1, t_2]$, соединяющая две фиксированные точки многообразия $p = \gamma(t_1)$ и $q = \gamma(t_2)$. Рассмотрим функционал

$$S[\gamma] = \int_\gamma dt L(x, \dot{x}). \tag{17.26}$$

Допустим, что заданы два произвольных векторных поля $X, Y \in \mathcal{X}(\mathbb{M}, \gamma)$, которые определены на кривой γ и равны нулю в точках p и q. Обозначим через λ и μ два вещественных параметра.

Определение. Назовем *первой вариацией* функционала (17.26) частную производную

$$\frac{\partial}{\partial\lambda}S[\gamma+\lambda X]\Big|_{\lambda=0}=\int_\gamma dt\frac{\delta S}{\delta x^\alpha}X^\alpha, \tag{17.27}$$

Частная производная

$$G_\gamma[X,Y]=G_\gamma[Y,X]=\frac{\partial^2}{\partial\lambda\partial\mu}S_\gamma[\gamma+\lambda X+\mu Y]\Big|_{\lambda=0,\ \mu=0} \tag{17.28}$$

называется *второй вариацией* функционала (17.26). □

Поскольку векторное поле X произвольно, то равенство нулю первой вариации эквивалентно уравнениям Эйлера–Лагранжа, которые мы записываем в виде

$$\frac{\delta S}{\delta x^\alpha}=\frac{\partial L}{\partial x^\alpha}-\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial\dot x^\alpha}\right)=0. \tag{17.29}$$

Предложение 17.4.1. *Если кривая $\gamma=\{x^\alpha(t)\}$ удовлетворяет уравнениям Эйлера–Лагранжа (17.29), то вторая вариация функционала (17.26) имеет вид*

$$G_\gamma[X,Y]=-\int_{t_1}^{t_2}dt\,\left(J_{\alpha\beta}X^\beta\right)Y^\alpha, \tag{17.30}$$

где

$$J_{\alpha\beta}X^\beta=\frac{d}{dt}\left(\frac{\partial^2L}{\partial\dot x^\alpha\partial\dot x^\beta}\dot X^\beta+\frac{\partial^2L}{\partial\dot x^\alpha\partial x^\beta}X^\beta\right)-\frac{\partial^2L}{\partial x^\alpha\partial\dot x^\beta}\dot X^\beta-\frac{\partial^2L}{\partial x^\alpha\partial x^\beta}X^\beta. \tag{17.31}$$

Доказательство. Прямые вычисления с учетом выражения для первой вариации

$$\begin{aligned}G_\gamma[X,Y]=&\frac{\partial}{\partial\mu}\int dt\left(\frac{\partial L}{\partial x^\beta}X^\beta+\frac{\partial L}{\partial\dot x^\beta}\dot X^\beta\right)\Big|_{\mu=0}=\\&=\int dt\left(\frac{\partial^2L}{\partial x^\alpha\partial x^\beta}X^\beta Y^\alpha+\frac{\partial^2L}{\partial\dot x^\alpha\partial x^\beta}X^\beta\dot Y^\alpha+\frac{\partial^2L}{\partial x^\alpha\partial\dot x^\beta}\dot X^\beta Y^\alpha+\frac{\partial^2L}{\partial\dot x^\alpha\partial\dot x^\beta}\dot X^\beta\dot Y^\alpha\right).\end{aligned}$$

После интегрирования по частям двух слагаемых, пропорциональных $\dot Y^\alpha$, получаем (17.31). □

Определение. Линейный оператор $J=\{J^\alpha{}_\beta:=g^{\alpha\gamma}J_{\gamma\beta}\}$,

$$J:\quad\mathcal{X}(\mathbb{M},\gamma)\to\mathcal{X}(\mathbb{M},\gamma),$$

который действует на векторные поля X, определенные вдоль кривой γ, по правилу (17.31), называется *оператором Якоби.* □

Теперь применим формулу (17.30) для вычисления второй вариации действия для экстремалей. Во-первых, отметим, что действие для экстремалей в виде (17.16) не является единственным. Рассмотрим действие

$$S=\int_\gamma dt\frac{1}{2}g_{\alpha\beta}\dot x^\alpha\dot x^\beta, \tag{17.32}$$

Предложение 17.4.2. *Действие (17.32) имеет тот же набор экстремалей, что и действие (17.16).*

Доказательство. Вариация действия (17.32) по $x^\alpha(t)$ приводит к уравнениям (17.17). □

В отличие от действия (17.16) функционал (17.32) не инвариантен относительно перепараметризации кривой. Это значит, что параметр t в действии (17.32) является каноническим.

Пусть $\gamma = \{x^\alpha(t)\}$, $t \in [t_1, t_2]$, – экстремаль и $u = \{\dot{x}^\alpha\}$ – вектор скорости экстремали.

Теорема 17.4.1. *Вторая вариация действия для экстремалей (17.32) имеет вид*

$$G_\gamma(X,Y) = -\int_{t_1}^{t_2} dt \left(\nabla_u^2 Y^\alpha + u^\delta u^\beta Y^\gamma \widetilde{R}_{\gamma\delta\beta}{}^\alpha\right) X_\alpha, \tag{17.33}$$

где

$$\begin{aligned}
\nabla_u Y^\alpha &= u^\beta \left(\partial_\beta Y^\alpha + \widetilde{\Gamma}_{\beta\gamma}{}^\alpha Y^\gamma\right) = \dot{Y}^\alpha + u^\beta \widetilde{\Gamma}_{\beta\gamma}{}^\alpha Y^\gamma, \\
\nabla_u^2 Y^\alpha &= u^\delta \nabla_\delta \left(\dot{Y}^\alpha + u^\beta \widetilde{\Gamma}_{\beta\gamma}{}^\alpha Y^\gamma\right) = \\
&= \ddot{Y}^\alpha + 2u^\beta \widetilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{Y}^\gamma + u^\delta u^\epsilon Y^\gamma \left(\partial_\delta \widetilde{\Gamma}_{\epsilon\gamma}{}^\alpha - \widetilde{\Gamma}_{\delta\epsilon}{}^\beta \widetilde{\Gamma}_{\gamma\beta}{}^\alpha + \widetilde{\Gamma}_{\delta\gamma}{}^\beta \widetilde{\Gamma}_{\epsilon\beta}{}^\alpha\right)
\end{aligned} \tag{17.34}$$

– ковариантные производные вдоль вектора скорости и $\widetilde{R}_{\gamma\delta\beta}{}^\alpha$ – тензор кривизны.

Доказательство. Первая вариация действия (17.32) имеет вид

$$\left.\frac{\partial S[\gamma + \lambda X]}{\partial \lambda}\right|_{\lambda=0} = -\int_{t_1}^{t_2} dt \left(\ddot{x}^\alpha + \widetilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma\right) X_\alpha.$$

Поэтому

$$\begin{aligned}
&G_\gamma(X,Y) = \\
&= -\frac{\partial}{\partial \mu} \int_{t_1}^{t_2} dt \left(\ddot{x}^\alpha + \mu \ddot{Y}^\alpha + \widetilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{x}^\gamma + 2\mu \widetilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{Y}^\gamma + \mu \partial_\epsilon \widetilde{\Gamma}_{\gamma\delta}{}^\alpha \dot{x}^\beta \dot{x}^\gamma Y^\epsilon\right) X_\alpha = \\
&= -\int_{t_1}^{t_2} dt \left(\ddot{Y}^\alpha + 2\widetilde{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}^\beta \dot{Y}^\gamma + \partial_\epsilon \widetilde{\Gamma}_{\gamma\delta}{}^\alpha \dot{x}^\beta \dot{x}^\gamma Y^\epsilon\right),
\end{aligned}$$

где мы оставили только линейные по μ слагаемые и учли, что вклад производной $\partial X_\alpha / \partial \mu$ равен нулю при выполнении уравнений для экстремалей $\nabla_u \dot{x}^\alpha = 0$. Исключив из полученного выражения производные $\ddot{Y}^\alpha$ и $\dot{Y}^\alpha$ с помощью второго из соотношений (17.34), получим (17.33). □

Ясно, что билинейная форма (17.33) симметрична, $G_\gamma(X,Y) = G_\gamma(Y,X)$, т.к. ковариантное дифференцирование можно перекинуть на вектор X, а тензор кривизны симметричен относительно перестановки пар индексов: $\widetilde{R}_{\alpha\beta\gamma\delta} = \widetilde{R}_{\gamma\delta\alpha\beta}$ (7.76).

Предположим, что метрика $g_{\alpha\beta}$ риманова. Тогда условие минимальности экстремали γ, соединяющей точки p и q, состоит в том, что квадратичная форма $G_\gamma(X,X)$ положительно определена для всех векторных полей, определенных на экстремали γ и обращающихся в нуль на ее концах.

Определение. Векторное поле X, которое определено на экстремали γ, соединяющей точки p и q, называется *якобиевым*, если оно удовлетворяет *уравнению Якоби*

$$J^\alpha{}_\beta X^\beta = 0, \tag{17.35}$$

где J – оператор Якоби (17.31), и обращается в нуль на концах p и q. Точки p и q называются *сопряженными* вдоль экстремали γ, если существует ненулевое якобиево поле X вдоль γ. □

Для действия (17.32) уравнение Якоби принимает вид

$$\nabla_u^2 X^\alpha + u^\delta u^\beta X^\gamma \widetilde{R}_{\gamma\delta\beta}{}^\alpha = 0. \tag{17.36}$$

Предложение 17.4.3. *Билинейная форма $G_\gamma(X,Y)$ невырождена тогда и только тогда, когда концевые точки p и q экстремали γ не сопряжены вдоль γ.*

Доказательство. Пусть X и Y произвольные векторные поля вдоль γ, которые обращаются в нуль в концевых точках. Напомним, что билинейная форма $G_\gamma(X,Y)$ называется невырожденной, если не существует такого векторного поля X, что $G_\gamma(X,Y) = 0$ для всех Y. Если поле X якобиево, то $G_\gamma(X,Y) = 0$ при любом Y.

Обратно. Допустим, что для вектора X выполнено уравнение $G_\gamma(X,Y) = 0$ для всех Y. Положим $Y = f(t)JX$, где функция $f(t)$ неотрицательна и обращается в нуль на концах экстремали. Тогда из выражения для второй вариации (17.30) имеем

$$G_\gamma(X,Y) = -\int_{t_1}^{t_2} dt\, fg(JX, JX) = 0.$$

Отсюда следует, что $JX = 0$. Поэтому концы сопряжены. □

Теперь сформулируем необходимое условие минимальности экстремали.

Теорема 17.4.2. *Если экстремаль γ, соединяющая точки p и q, содержит внутри себя пару сопряженных точек p' и q', то она не является минимальной.*

Доказательство. См., например, [**?**]. □

Теорема 17.4.3. *Для достаточно малых отрезков $l = [t_1, t_2]$ экстремали задают минимум функционала (17.32). Поэтому каждая экстремаль является кратчайшей линией в классе дважды дифференцируемых кривых, соединяющих достаточно близкие точки p и q.*

Доказательство. Экстремаль γ реализует минимум функционала (17.32), если форма $G_\gamma(X,X)$ положительна для всех ненулевых векторных полей $X \in \mathcal{X}(\mathbb{M},\gamma)$, которые обращаются в нуль на концах p и q. Из формулы для второй вариации (17.33) следует, что

$$\begin{aligned} G_\gamma(X,X) &= -\int_p^q dt \left[(\nabla_u^2 X, X) - u^\alpha X^\beta u^\gamma X^\delta \widetilde{R}_{\alpha\beta\gamma\delta}\right] \\ &= \int_p^q dt \left[(\nabla_u X, \nabla_u X) + u^\alpha X^\beta u^\gamma X^\delta \widetilde{R}_{\alpha\beta\gamma\delta}\right], \end{aligned}$$

где мы проинтегрировали первое слагаемое по частям с учетом равенств $X(p) = X(q) = 0$. Можно доказать, что для достаточно коротких экстремалей справедлива оценка

$$\left|\int_p^q dt\, u^\alpha X^\beta u^\gamma X^\delta \widetilde{R}_{\alpha\beta\gamma\delta}\right| = \mathsf{O}(l) \int_p^q dt (\nabla_u X, \nabla_u X)$$

при $l \to 0$. Поскольку квадратичная форма $(\nabla_u X, \nabla_u X)$ положительна, то для достаточно коротких экстремалей положительна и форма $G_\gamma(X,X)$. □

17.5 Уравнение Гамильтона–Якоби для экстремалей

Уравнения для экстремалей $\gamma = x(t)$ вытекают из вариационного принципа для действия (17.16). Важным обстоятельством является то, что уравнения для экстремалей являются уравнениями Эйлера–Лагранжа также и для другого действия (17.32). А именно, рассмотрим действие

$$S_{\mathrm{m}} = \int_a^b dt L_{\mathrm{m}}, \qquad L_{\mathrm{m}} = -\frac{1}{2} m g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta, \tag{17.37}$$

где точка обозначает дифференцирование по каноническому параметру t и $m = \mathsf{const}$ – постоянная, имеющая физический смысл массы точечной частицы. Это действие совпадает с (17.32) при $m = -1$. Действие (17.37) приводит к уравнениям для экстремалей, в которых переменная t уже является каноническим параметром. Это согласуется с тем обстоятельством, что рассмотренное действие инвариантно относительно общих преобразований координат и сдвигов параметра t. Для сравнения напомним, что исходное действие для экстремалей (17.16) инвариантно также относительно произвольных преобразований параметра t вдоль экстремали.

Действие (17.37) имеет простой физический смысл. Предположим, что метрика имеет лоренцеву сигнатуру и зафиксируем временну́ю калибровку (21.166)

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \qquad \mathsf{sign}\, g_{\mu\nu} = (- \ldots -). \tag{17.38}$$

Символы Кристоффеля для этой метрики были вычислены ранее (21.172). Предположим также, что пространственная часть метрики $g_{\mu\nu}$ не зависит от времени x^0. Тогда уравнения для экстремалей расщепляются:

$$\ddot{x}^0 = 0, \tag{17.39}$$

$$\ddot{x}^\mu = -\widetilde{\Gamma}_{\nu\rho}{}^{\mu} \dot{x}^\nu \dot{x}^\rho. \tag{17.40}$$

Из первого уравнения следует, что, не ограничивая общности, канонический параметр t можно отождествить с временем $x^0 = ct$, где $c = \mathsf{const}$ – скорость света. Тогда лагранжиан (17.37) имеет прямой физический смысл – с точностью до аддитивной постоянной это кинетическая энергия точечной частицы, которая движется в римановом пространстве со статической метрикой $g_{\mu\nu}(x)$. Несмотря на то, что потенциальная энергия частицы равна нулю, ее траекториями уже не будут прямые линии, если метрика нетривиально зависит от точки пространства.

Вернемся к исходному действию (17.37) до фиксирования временно́й калибровки. Переформулируем эту лагранжеву систему на гамильтоновом языке, рассматривая канонический параметр t в качестве параметра эволюции. Под временем мы подразумеваем координату x^0 и, соответственно, предполагаем, что $g_{00} > 0$. Кроме этого мы предполагаем, что все сечения постоянного времени $x^0 = \mathsf{const}$ пространственноподобны. Импульс, сопряженный координатам x^α, и гамильтониан системы равны

$$\begin{aligned} p_\alpha &= \frac{\partial L_{\mathrm{m}}}{\partial \dot{x}^\alpha} = -m g_{\alpha\beta} \dot{x}^\beta, \\ H_{\mathrm{m}} &= -\frac{1}{2m} g^{\alpha\beta} p_\alpha p_\beta. \end{aligned} \tag{17.41}$$

Для действия (17.37) связи на канонические переменные отсутствуют, т.к. метрика рассматривается как внешнее поле и варьирование по ней не проводится. Соответствующие уравнения Гамильтона (уравнения движения) имеют вид

$$\dot{x}^\alpha = [x^\alpha, H_{\rm m}] = -\frac{1}{m} g^{\alpha\beta} p_\beta, \tag{17.42}$$

$$\dot{p}_\alpha = [p_\alpha, H_{\rm m}] = \frac{1}{2m} \partial_\alpha g^{\beta\gamma} p_\beta p_\gamma. \tag{17.43}$$

Дифференцируя первое из этих уравнений по каноническому параметру и исключая импульсы p_α и производные $\dot{p}_\alpha$ с помощью уравнений движения, нетрудно проверить, что система уравнений (17.42), (17.43) эквивалентна системе уравнений для экстремалей (17.17). Тем самым мы переписали уравнения для экстремалей в виде канонической системы уравнений движения.

Ранее было доказано, что длина касательного вектора к экстремали постоянна (17.22). В гамильтоновом форме это утверждение имеет вид

$$C_0 = \frac{1}{m^2} g^{\alpha\beta} p_\alpha p_\beta = \mathsf{const}. \tag{17.44}$$

По своей физической сути это есть закон сохранения энергии точечной частицы. В данном случае только кинетической, т.к. потенциальная энергия тождественно равна нулю.

В дальнейшем нам понадобятся гамильтоновы уравнения для нулевых экстремалей, где в качестве параметра эволюции выбрано время x^0, а не канонический параметр t. Они получаются следующим образом. Для нулевых экстремалей интеграл движения (17.44) принимает вид

$$g^{00} p_0^2 + 2 g^{0\mu} p_0 p_\mu + g^{\mu\nu} p_\mu p_\nu = 0.$$

Это квадратное уравнение решается относительно p_0:

$$p_0 = N^\mu p_\mu \pm N\hat{p}, \qquad \hat{p} = \sqrt{-\hat{g}^{\mu\nu} p_\mu p_\nu}, \tag{17.45}$$

где $N = 1/\sqrt{g^{00}} \neq 0$ и $N^\mu = \hat{g}^{\mu\nu} g_{0\nu}$ – функции хода и сдвига, используемые в АДМ параметризации метрики (22.5), а $\hat{g}^{\mu\nu}$ – метрика, обратная к пространственной метрике $g_{\mu\nu}$. Если частица движется, то $p_\mu \neq 0$ и $\hat{p} \neq 0$. Тогда из уравнения (17.42) находим производную координаты x^0 по каноническому параметру

$$\dot{x}^0 = \mp \frac{\hat{p}}{mN}.$$

Отсюда следует, что выбор знака в (17.45) соответствует выбору взаимной ориентации канонического параметра t и времени x^0. После этого канонические уравнения (17.42), (17.43) можно записать в виде системы уравнений движения только для пространственных координат и импульсов

$$\begin{aligned} \partial_0 x^\mu &= \frac{\dot{x}^\mu}{\dot{x}^0} = -N^\mu \pm \frac{N \hat{g}^{\mu\nu} p_\nu}{\hat{p}}, \\ \partial_0 p_\mu &= \frac{\dot{p}_\mu}{\dot{x}^0} = \partial_\mu N^\nu p_\nu \pm \partial_\mu N \hat{p} \mp \frac{N \partial_\mu \hat{g}^{\nu\rho} p_\nu p_\rho}{2\hat{p}}. \end{aligned} \tag{17.46}$$

Таким образом, из системы канонических уравнений (17.42), (17.43) для нулевых экстремалей мы исключили в явном виде канонический параметр t и нулевую компоненту импульса p_0.

Предложение 17.5.1. *Система уравнений для пространственных компонент канонически сопряженных переменных x^μ и p_μ (17.46) является гамильтоновой. При этом эволюция системы уравнений рассматривается по отношению к времени x^0, а гамильтонианом является выражение для p_0 (17.45).*

Доказательство. Прямое сравнение уравнений

$$\partial_0 x^\mu = [x^\mu, p_0],$$
$$\partial_0 p_\mu = [p_\mu, p_0].$$

□

Число гамильтоновых уравнений движения, определяющих экстремаль, сократилось с $2n$ в (17.42), (17.43) до $2n-2$ в (17.46). Это достигнуто за счет использования интеграла движения (17.44) и выбора специального параметра эволюции x^0.

Продолжим анализ гамильтоновой формы уравнений для экстремалей. Функция действия $S_{\mathrm{m}}(x,t)$ (20.32) удовлетворяет уравнению Гамильтона–Якоби (20.34)

$$\frac{\partial S_{\mathrm{m}}}{\partial t} - \frac{1}{2m} g^{\alpha\beta} \frac{\partial S_{\mathrm{m}}}{\partial x^\alpha} \frac{\partial S_{\mathrm{m}}}{\partial x^\beta} = 0. \tag{17.47}$$

Поскольку гамильтониан (17.41) не зависит от параметра t явно, то функция действия имеет вид

$$S_{\mathrm{m}}(x^\alpha, t) = \frac{mC_0}{2} t + W_{\mathrm{m}}(x^\alpha), \qquad C_0 = \mathsf{const},$$

где укороченная функция действия W_{m} удовлетворяет укороченному уравнению Гамильтона–Якоби

$$g^{\alpha\beta} \frac{\partial W_{\mathrm{m}}}{\partial x^\alpha} \frac{\partial W_{\mathrm{m}}}{\partial x^\beta} = m^2 C_0. \tag{17.48}$$

Так как $p_\alpha = \partial W_{\mathrm{m}} / \partial x^\alpha$, то постоянная C_0 равна длине касательного вектора к экстремали

$$g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = \frac{1}{m^2} g^{\alpha\beta} p_\alpha p_\beta = C_0.$$

Отметим, что для экстремалей нулевой длины $C_0 = 0$, и укороченное действие совпадает с полным $W_{\mathrm{m}} = S_{\mathrm{m}}$.

17.6 Волновое уравнение

Пусть задано многообразие $\mathbb{M}$, $\mathsf{dim}\,\mathbb{M} = n$, с метрикой лоренцевой сигнатуры, $\mathsf{sign}\, g_{\alpha\beta} = (+ - \ldots -)$. Рассмотрим волновое уравнение для скалярного поля $\varphi(x) \in \mathcal{C}^2(\mathbb{M})$:

$$g^{\alpha\beta} \widetilde{\nabla}_\alpha \widetilde{\nabla}_\beta \varphi = g^{\alpha\beta} \partial_\alpha \partial_\beta \varphi - g^{\alpha\beta} \widetilde{\Gamma}_{\alpha\beta}{}^\gamma \partial_\gamma \varphi = 0, \tag{17.49}$$

или, эквивалентно,

$$\frac{1}{\sqrt{|g|}} \partial_\alpha \left(\sqrt{|g|} g^{\alpha\beta} \partial_\beta \varphi \right) = 0, \qquad g = \mathsf{det}\, g_{\alpha\beta},$$

где мы воспользовались тождеством (7.61). Это – линейное дифференциальное уравнение в частных производных второго порядка с переменными коэффициентами гиперболического типа, т.к. метрика имеет лоренцеву сигнатуру. Важным понятием

в теории дифференциальных уравнений является характеристика (или характеристическая поверхность), которая для дифференциальных уравнений второго порядка определяется квадратичной формой $\sqrt{|g|}g^{\alpha\beta}$, задающей вторые производные от неизвестной функции $\varphi(x)$. В рассматриваемом случае квадратичная форма вторых производных задается обратной метрикой $g^{\alpha\beta}$.

Определение. *Характеристикой* гиперболического дифференциального уравнения второго порядка (17.49) называется $\mathcal{C}^1$ гиперповерхность в многообразии $\mathbb{M}$, которая задается уравнением

$$W(x) = 0, \tag{17.50}$$

где функция W на поверхности $W = 0$ удовлетворяет условию

$$g^{\alpha\beta}\partial_\alpha W \partial_\beta W\Big|_{W=0} = 0. \tag{17.51}$$

При этом требуется, чтобы по крайней мере одна из частных производных $\partial_\alpha W$ была отлична от нуля на гиперповерхности (17.50). □

Отметим, что в определении характеристики важна гиперболичность, т.к. при положительно или отрицательно определенной метрике уравнение (17.51) не имеет вещественных решений. Это условие для характеристики совпадает с укороченным уравнением Гамильтона–Якоби для экстремалей (17.48) при $C_0 = 0$. Это значит, что характеристика соответствует укороченной функции действия для экстремалей нулевой длины. Напомним, что для экстремалей нулевой длины укороченная и полная функции действия совпадают. Однако условие (17.51) является более слабым, т.к. мы требуем выполнения (17.51) только на характеристике, а не во всем пространстве-времени.

Из определения характеристик следует, что они являются изотропными гиперповерхностями, рассмотренными в разделе 21.12.3. Метрика, индуцированная на таких поверхностях, по-определению, вырождена, и все нормальные векторы изотропны (теорема 21.12.7).

Характеристики обладают следующим важным свойством. Допустим, что каждая гиперповерхность $W(x) - y^0 = 0$, где $-\epsilon < y^0 < \epsilon$, $\epsilon > 0$, есть характеристика уравнения (17.49). Другими словами, мы требуем, чтобы уравнение характеристик (17.51) выполнялось не только на самой характеристике, но и в некоторой ее окрестности. Поскольку на каждой характеристике по крайней мере одна из частных производных $\partial_\alpha W$ отлична от нуля, то это семейство заполняет некоторую достаточно малую область, через каждую точку которой проходит одна и только одна характеристика. Тогда можно перейти в новую систему координат $x^\alpha \to y^\alpha(x)$, где $y^0 = W$. При этом обратная метрика преобразуется по тензорному закону

$$g^{\alpha\beta}(y) = \frac{\partial y^\alpha}{\partial x^\gamma}\frac{\partial y^\beta}{\partial x^\delta}g^{\gamma\delta}(x).$$

Отсюда следует, что в новой системе координат $g^{00} = 0$. Это обстоятельство имеет важное следствие. Рассмотрим задачу Коши для волнового уравнения (17.49) в новой системе координат. Пусть на характеристике заданы начальные условия

$$\varphi|_{y^0=0} = \varphi_0, \qquad \partial_0\varphi|_{y^0=0} = \varphi_1,$$

где φ_0 и φ_1 – достаточно гладкие функции от "пространственных" координат $\{y^\mu\} = (y^1, \ldots, y^{n-1})$. По начальным данным на характеристике при $y^0 = 0$ можно вычислить все частные производные по "пространственным" координатам $\partial_\mu\varphi, \partial^2_{\mu\nu}\varphi, \ldots$

и все частные производные с одной производной по "времени" $\partial^2_{0\mu}\varphi, \partial^3_{0\mu\nu}\varphi, \ldots$. Для определения эволюции скалярного поля необходимо знать вторые производные по "времени" $\partial^2_{00}\varphi$. Если задача Коши корректно поставлена, то вторые производные по "времени" находятся из волнового уравнения. На характеристике это не так, поскольку $g^{00} = 0$. Вместо определения второй производной $\partial^2_{00}\varphi$ волновое уравнение накладывает ограничение (связь) на возможный выбор начальных условий:

$$g^{0\mu}\partial_\mu\varphi_1 - g^{\alpha\beta}\widetilde{\Gamma}_{\alpha\beta}{}^0\varphi_1 - g^{\alpha\beta}\widetilde{\Gamma}_{\alpha\beta}{}^\mu\partial_\mu\varphi_0 = 0.$$

Для нахождения вторых производных $\partial^2_{00}\varphi$ волновое уравнение нужно продифференцировать по y^0. Однако после этой процедуры вторые производные не будут определены однозначно. Таким образом, при постановке задачи Коши на характеристике, во-первых, начальные данные нельзя задавать произвольно, и, во-вторых, волновое уравнение не определяет эволюцию поля единственным образом. Тем самым задача Коши на характеристике не допускает корректную постановку. Выше мы взяли слова "пространственный" и "время" в кавычки, т.к. характеристика является изотропной поверхностью и не может быть пространственноподобной, а координата y^0 светоподобна и не может играть роль времени.

Поскольку вторые производные от неизвестной функции по одну сторону характеристики не определяются уравнением (17.49) и значениями функции φ по другую сторону от характеристики, то они могут иметь разрывы. Это значит, что решения уравнения (17.49) могут иметь разрывы производных, которые распространяются в пространстве вдоль характеристик.

Для корректной постановки задачи Коши для волнового уравнения (17.49) на многообразии с метрикой лоренцевой сигнатуры по x^0 координаты выбираются таким образом, что $g^{00} > 0$, а сечение $x^0 = 0$ – пространственноподобным. Поскольку пространственноподобное сечение не может быть изотропным, то оно не может быть также характеристикой. Оно не может также касаться характеристики, потому что в этом случае один из касательных векторов был бы изотропным, что противоречит отрицательной определенности метрики, индуцированной на гиперповерхности.

Замечание. В разделе 5.2 мы определили время x^0 как такую координату, что касательное векторное поле ∂_0 является времениподобным. Это соответствует условию $g_{00} > 0$. Отметим, что условия $g_{00} > 0$ и $g^{00} > 0$ неэквивалентны. □

Уравнение характеристик (17.51) представляет собой нелинейное дифференциальное уравнение в частных производных первого порядка. В общем случае оно очень сложно и, как правило, имеет решения с особыми точками. Характеристика, например, может быть не гладкой поверхностью.

Относительно частной производной по времени $\partial_0 W$ уравнение характеристик является алгебраическим квадратным уравнением, и имеет не более двух вещественных корней. Допустим, что метрика имеет лоренцеву сигнатуру и $g_{00} > 0$, тогда оно имеет два вещественных корня разных знаков

$$\partial_0 W = N^\mu\partial_\mu W \pm N\sqrt{-\hat{g}^{\mu\nu}\partial_\mu W\partial_\nu W}, \tag{17.52}$$

где использована АДМ параметризаций метрики (22.5). Здесь мы предполагаем, что все сечения $x^0 = \mathsf{const}$ являются пространственноподобными и, следовательно, метрика $g_{\mu\nu}$ и ее обратная $\hat{g}^{\mu\nu}$ отрицательно определены. Обозначим $p_\mu = \partial_\mu W$. Введем функцию Гамильтона в фазовом пространстве $\{x^\mu, p_\mu\}$ (см. главу 20)

$$H(x^0, x^\mu, p_\mu) = -N^\mu p_\mu \mp N\sqrt{-\hat{g}^{\mu\nu}p_\mu p_\nu}, \tag{17.53}$$

где зависимость от координат x^0 и x^μ входит через метрику $g^{\alpha\beta} = g^{\alpha\beta}(x^0, x^\mu)$, и время x^0 рассматривается в качестве параметра эволюции.

Определение. Траектории $x^\mu(x^0)$ в конфигурационном пространстве, соответствующем гамильтониану (17.53), называются *бихарактеристиками* волнового уравнения (17.49). □

Предложение 17.6.1. *Бихарактеристики волнового уравнения (17.49) совпадают с нулевыми экстремалями для метрики $g_{\alpha\beta}$.*

Доказательство. Следствие предложения 17.5.1. □

Единственное отличие бихарактеристик и нулевых экстремалей сводится к тому, что экстремали определяются в параметрическом виде $x^\alpha(t)$, а бихарактеристики задаются в виде явной зависимости координат $x^\mu(x^0)$.

Гамильтониан (17.53) совпадает с временно́й компонентой импульса (17.45). Функция $W(x^0, x^\mu)$, определяющая характеристику, является действием для бихарактеристик.

Замечание. Характеристики и бихарактеристики имеют физическую интерпретацию. Из теории дифференциальных уравнений известно, что если в момент времени x^0 в точке $\boldsymbol{x}$ произошло некоторое возмущение решения φ волнового уравнения (17.49), то оно будет распространяться в виде волны. При этом сечения $x^0 = \mathsf{const}$ соответствующей характеристической поверхности (характеристического коноида с вершиной в точке $\{x^0, \boldsymbol{x}\}$) являются фронтом волны в момент времени x^0, а бихарактеристики – это лучи, вдоль которых распространяется волна. □

Предложение 17.6.2. *Если бихарактеристика касается характеристики в некоторой точке, то она целиком лежит на этой характеристике.*

Доказательство. Для нулевых экстремалей лагранжиан можно выбрать в виде

$$L = \frac{1}{2} g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta$$

и равен нулю $L = 0$. Следовательно, функция действия $S(x, t)$ и укороченная функция действия $W = S + Et$ равны нулю, т.к. для нулевых экстремалей $E = 0$. Тем самым для бихарактеристик $W = 0$ и они целиком лежат на какой-либо из характеристик. □

17.7 Приближение эйконала

Посмотрим на волновое уравнение (17.49) с другой точки зрения. Любое решение волнового уравнения (17.49) можно представить в виде[1]

$$\varphi = A\, \mathsf{e}^{iW/\hbar}, \tag{17.54}$$

[1] В этом представлении $\hbar \in \mathbb{R}$ рассматривается, как вещественный параметр. Обозначение продиктовано аналогией с квантовой механикой, где $\hbar$ – это постоянная Планка.

где $A(x) \neq 0$ и $W(x)$ – амплитуда и фаза волны. Тогда волновое уравнение примет вид

$$g^{\alpha\beta}\frac{\partial^2_{\alpha\beta}A}{A} + \frac{1}{\hbar}2ig^{\alpha\beta}\frac{\partial_\alpha A}{A}\partial_\beta W + \frac{1}{\hbar}ig^{\alpha\beta}\partial^2_{\alpha\beta}W - \\ - \frac{1}{\hbar^2}g^{\alpha\beta}\partial_\alpha W\partial_\beta W - g^{\alpha\beta}\Gamma_{\alpha\beta}{}^{\gamma}\left(\frac{\partial_\gamma A}{A} + \frac{1}{\hbar}i\partial_\gamma W\right) = 0.$$

Тогда в пределе $\hbar \to 0$ волновое уравнение сводится к уравнению на фазу

$$g^{\alpha\beta}\partial_\alpha W\partial_\beta W = 0. \tag{17.55}$$

Оно совпадает с уравнением для характеристик (17.51), однако должно выполняться во всем пространстве-времени.

Определение. Для волновых решений предел $\hbar \to 0$ означает, что относительное изменение амплитуды мало, по сравнению с изменением фазы. Этот предел называют *приближением эйконала*. Фазу $W(x)$ называют *эйконалом*, а уравнение (17.55) – *уравнением эйконала*. В электродинамике это приближение называют также приближением *геометрической оптики*. Для плоских волн оно соответствует высоким частотам излучения и маленьким длинам волн. □

Параметр $\hbar$ был введен с единственной целью – определить эйкональное приближение. Поэтому в дальнейшем, для краткости, мы включим его в определение фазы $W/\hbar \to W$.

Пример 17.7.1. Рассмотрим волновое уравнение в трехмерном пространстве-времени Минковского $\mathbb{R}^{1,2}$ в декартовой системе координат $\{x^\alpha\} = \{x^0, x^1, x^2\}$

$$\partial_0^2\varphi - \partial_1^2\varphi + \partial_2^2\varphi = 0.$$

В данном примере трехмерное пространство Минковского выбрано для наглядности. Все следующие формулы естественным образом переносятся на пространство Минковского произвольной размерности $\mathbb{R}^{1,n-1}$.

Уравнение характеристик имеет вид

$$\Big((\partial_0 W)^2 - (\partial_1 W)^2 - (\partial_2 W)^2\Big)\Big|_{W=0} = 0. \tag{17.56}$$

Это уравнение допускает два семейства характеристик (изотропные поверхности в примере 21.12.2).

Первое семейство характеристик

$$W = (x^0 - x^0_0)^2 - (x^1 - x^1_0)^2 - (x^2 - x^2_0)^2 = 0,$$

где $x_0 = \{x^0_0, x^1_0, x^2_0\}$ – три произвольных вещественных параметра (координаты точки x_0). Нетрудно проверить, что

$$(\partial_0 W)^2 - (\partial_1 W)^2 - (\partial_2 W)^2 = 4W.$$

Это значит, что уравнение характеристик (17.56) выполняется только на характеристиках, а не во всем пространстве-времени. Для каждой точки x_0 характеристики первого семейства состоят из двух конусов (прошлого, $x^0 < x^0_0$, и будущего, $x^0 > x^0_0$) с общей вершиной в точке x_0, как показано на рис. 17.3,*а*.

Рис. 17.3: Характеристиками первого семейства являются конусы с вершинами в произвольной точке x_0 (a). Характеристиками второго семейства являются плоскости, касательные к конусам (b).

Второе семейство характеристик

$$W = x^0 + k_\mu x^\mu - C, \qquad \mu = 1, 2$$

параметризуется единичным вектором в пространстве $|\boldsymbol{k}| = \sqrt{k_1^2 + k_2^2} = 1$ и постоянной C. В этом случае уравнение характеристик выполняется во всем пространстве, а не только на поверхности $W = 0$. Характеристики второго семейства параметризуются двумя параметрами, и представляют собой плоскости, перпендикулярные ковектору $(1, k_1, k_2)$ нулевой длины и пересекающие плоскость $x^0 = 0$ по прямой $k_\mu x^\mu = C$. Эти характеристики касаются характеристик первого семейства (конусов).

Напомним, что нулевыми экстремалями в пространстве Минковского являются прямые и только они с нулевым касательным вектором. Мы видим, что если произвольная нулевая экстремаль (или бихарактеристика) в некоторой точке касается характеристики, то она целиком принадлежит этой характеристике. Кроме того, через каждую регулярную точку характеристики (исключение составляют вершины конусов первого семейства) проходит одна и только одна нулевая экстремаль. Поэтому нулевые экстремали полностью заметают характеристические поверхности. Отметим, что одна и та же нулевая экстремаль может принадлежать разным характеристическим конусам.

Из уравнения для характеристик (17.56) можно найти производную по времени

$$\partial_0 W = \pm\sqrt{(\partial_1 W)^2 + (\partial_2 W)^2}.$$

Отсюда следует выражение для гамильтониана, определяющего бихарактеристики (траектории)

$$H = \pm\sqrt{p_1^2 + p_2^2} = \pm\hat{p}.$$

Уравнения движения для бихарактеристик имеют вид

$$\begin{aligned} &\dot{x}^\mu = \pm\frac{p^\mu}{\hat{p}}, \\ &\dot{p}_\mu = 0, \qquad\quad \Rightarrow \quad p_\mu = \mathsf{const}, \end{aligned}$$

где точка обозначает дифференцирование по времени $t = x^0$. Отсюда следует, что импульс бихарактеристик постоянен, а траектории – это прямые $x^\mu = \pm n^\mu t + x_0^\mu$, проходящие через все точки пространства $x_0 \in \mathbb{R}^2$ во всех возможных направлениях $n^\mu = p^\mu/\hat{p}$. Поскольку пространственный вектор n^μ имеет единичную длину, то бихарактеристики в пространстве-времени совпадают с нулевыми экстремалями.

В пространстве Минковского уравнение (17.49) допускает решение в виде плоской волны

$$\varphi = A_0\, \mathsf{e}^{ik_\alpha x^\alpha} = A_0\, \mathsf{e}^{i(\omega t - \boldsymbol{kx})}, \qquad A_0 = \mathsf{const} \neq 0,$$

где $\boldsymbol{kx} = -k_\mu x^\mu$. Сравнение этого выражения с (17.54) дает выражения для *частоты* ω и *волнового вектора* $\boldsymbol{k}$ (в рассматриваемом примере мы, для краткости, полагаем $\hbar = 1$)

$$\omega = k_0 = \partial_0 W, \qquad \boldsymbol{k} = \{\partial_1 W, \partial_2 W\}$$

Тогда волновое уравнение (17.51) сводится к соотношению между частотой и волновым вектором

$$g^{\alpha\beta}k_\alpha k_\beta = \omega^2 - \boldsymbol{k}^2 = 0,$$

где $\boldsymbol{k}^2 = -g^{\mu\nu}k_\mu k_\nu$. Для плоской волны в пространстве Минковского эйконал $W(x) = x^\alpha k_\alpha$ является линейной функцией от координат, а частота и волновой вектор постоянны. □

В рассмотренном примере плоских волн зависимость эйконала от координат была линейной. В общем случае эта зависимость является более сложной. Тогда частота и волновой вектор определяются соотношениями

$$\omega = \partial_0 W, \qquad k_\mu = \partial_\mu W, \tag{17.57}$$

и зависят от точки пространства-времени. Уравнение эйконала (17.55) определяет зависимость частоты от волнового вектора

$$g^{00}\omega^2 + 2g^{0\mu}\omega k_\mu + g^{\mu\nu}k_\mu k_\nu = 0.$$

Эта зависимость называется *дисперсией*, а производная

$$v_{\mathrm{g}}^\mu = \frac{\partial\omega}{\partial k_\mu}$$

называется *групповой скоростью*. Сравнение уравнения, определяющего дисперсию, с (17.52) показывает, что частота ω и волновой вектор k_μ по сути дела совпадают, соответственно, с гамильтонианом и импульсами бихарактеристик волнового уравнения.

17.8 Гармонические координаты

Для исследования волнового уравнения (17.49) на многообразии $\mathbb{M}$, $\dim\mathbb{M} = n$, с метрикой лоренцевой сигнатуры $g_{\alpha\beta}$ удобно использовать гармонические координаты. Как отмечалось в разделе 2.5 функции перехода к новой системе координат являются скалярными полями на $\mathbb{M}$. Поэтому рассмотрим волновое уравнение (17.49) и допустим, что в некоторой области $\mathbb{U} \subset \mathbb{M}$ оно имеет n функционально независимых решений. Пронумеруем эти решения φ^{A}, $\mathrm{A}, \mathrm{B}, \ldots = 0, 1, \ldots, n-1$. Тогда в области $\mathbb{U}$ решения волнового уравнения задают систему координат $x^{\mathrm{A}} = \varphi^{\mathrm{A}}$, которая называется *гармонической*.

Замечание. Эта система координат была введена Т. де Дондером [?] и К. Ланцосом [?] и получила физическую интерпретацию в работах В. А. Фока [?, ?]. □

Гармоническая система координат обладает следующим важным свойством.

Предложение 17.8.1. *Система координат является гармонической тогда и только тогда, когда выполнены условия*

$$\partial_{\mathrm{A}}(\sqrt{|g|}g^{\mathrm{AB}}) = 0 \quad\Leftrightarrow\quad \Xi^{\mathrm{C}} = g^{\mathrm{AB}}\Gamma_{\mathrm{AB}}{}^{\mathrm{C}} = 0, \tag{17.58}$$

где $g = \det g_{\mathrm{AB}}$.

Доказательство. Справедливо тождество

$$\partial_\alpha \left[\det \left(\frac{\partial x}{\partial \varphi} \right) \frac{\partial \varphi^{\mathrm{A}}}{\partial x^\beta} \right] \frac{\partial x^\alpha}{\partial \varphi^{\mathrm{A}}} = 0,$$

где $\partial x/\partial \varphi$ – обратная матрица Якоби преобразования координат (2.57). Это тождество доказывается прямым дифференцированием с учетом правила дифференцирования определителя матрицы (28.20). С учетом доказанного равенства и правила преобразования координат имеем

$$\partial_{\mathrm{A}}(\sqrt{|g|}g^{\mathrm{AB}}) = \frac{\partial x^\gamma}{\partial \varphi^{\mathrm{A}}} \partial_\gamma \left[\sqrt{|\overset{\circ}{g}|} \det \left(\frac{\partial x}{\partial \varphi} \right) \frac{\partial \varphi^{\mathrm{A}}}{\partial x^\alpha} \frac{\partial \varphi^{\mathrm{B}}}{\partial x^\beta} g^{\alpha\beta} \right] =$$

$$= \det \left(\frac{\partial x}{\partial \varphi} \right) \partial_\alpha \left(\sqrt{|\overset{\circ}{g}|} g^{\alpha\beta} \partial_\beta \varphi^{\mathrm{B}} \right) = 0,$$

где $\overset{\circ}{g} = \det g_{\alpha\beta}$. Таким образом, из гармоничности координат следует первое равенство (17.58) и наоборот. Эквивалентность равенств (17.58) между собой проверяется прямой проверкой. □

Уравнения на метрику (17.58) называются *условиями гармоничности*. В дальнейшем будем, как обычно, обозначать гармоническую систему координат снова греческими буквами x^α, предполагая, что условия гармоничности (17.58) выполнены. В гармонической системе координат оператор Лапласа–Бельтрами, действующий на функцию f, принимает простой вид

$$\triangle f = g^{\alpha\beta} \partial^2_{\alpha\beta} f \tag{17.59}$$

и не содержит первых частных производных $\partial_\alpha f$.

Пример 17.8.1. В пространстве-времени Минковского $\mathbb{R}^{1,n-1}$ декартова система координат является гармонической. □

Допустим, что на произвольном лоренцевом многообразии $\mathbb{M}$, которое топологически совпадает с $\mathbb{R}^{1,n-1}$, выбрана гармоническая система координат x^α. Тогда произвольная линейная комбинация координат удовлетворяет волновому уравнению (17.59). Можно доказать, что все гармонические системы координат на $\mathbb{M}$, которые являются асимптотически декартовыми, связаны между собой преобразованиями Лоренца [**?**]. В. А. Фок придавал гармоническим системам координат в общей теории относительности физический смысл, считая, что на произвольном лоренцевом многообразии $\mathbb{M}$ они выделены и играют ту же роль, что и декартовы координаты в пространстве-времени Минковского.

В переменных Картана условие гармоничности (17.58) имеет простой вид

$$\tilde{\omega}^a = g^{\alpha\beta} \partial_\alpha e_\beta{}^a = 0.$$

17.9 Нормальные, геодезические или римановы координаты

Нормальные координаты, которые называются также геодезическими или римановыми играют большую роль при изучении свойств метрики и связности, а также в

приложениях в математической физике. Сначала мы введем эту систему координат с помощью рядов, что является более наглядным и важным для приложений. Затем дадим определение с помощью экспоненциального отображения и сформулируем ряд утверждений, связанных с полнотой римановых многообразий.

Пример 17.9.1. Сферические координаты (см. раздел 8.1) являются нормальными координатами в трехмерном евклидовом пространстве $\mathbb{R}^3$. □

В определенном смысле нормальные координаты являются обобщением сферической системы координат на общий случай многообразия с заданной аффинной геометрией.

17.9.1 Нормальные координаты. Локальное рассмотрение.

Пусть на многообразии $\mathbb{M}$ задана аффинная геометрия $(\mathbb{M}, g, \Gamma)$. В настоящем разделе мы предполагаем, что все геометрические объекты (в частности метрика и связность) являются вещественно аналитическими, т.е. их компоненты в окрестности каждой точки $x_0 \in \mathbb{M}$ представимы в виде сходящихся степенных рядов. Мы рассмотрим специальную систему координат в окрестности точки x_0, которая является аналогом декартовой системы координат в евклидовом пространстве и имеет многочисленные приложения. В литературе эта система координат встречается под разными названиями: геодезические, римановы или нормальные координаты, и обычно строится в (псевдо-)римановом пространстве. Мы построим такую систему координат в более общем случае произвольной аффинной геометрии.

Сначала мы рассмотрим геодезические линии вблизи точки x_0. Пусть на многообразии задана кривая $x(t)$, проходящая через точку x_0 в заданном направлении,

$$x^\alpha(0) = x_0^\alpha, \qquad \dot{x}^\alpha(0) = \dot{x}_0^\alpha. \tag{17.60}$$

Для того, чтобы кривая была геодезической, функции $x^\alpha(t)$ должны удовлетворять уравнениям для геодезических (17.3) с начальными условиями (17.60). При достаточно малых t будем искать решение уравнений для геодезических в виде степенного ряда

$$x^\alpha = x_0^\alpha + \dot{x}_0^\alpha t + \frac{1}{2}\ddot{x}_0^\alpha t^2 + \frac{1}{3!}\dddot{x}_0^\alpha\, t^3 + \dots \tag{17.61}$$

Первые два члена определяются начальными условиями (17.60), а все последующие – уравнениями (17.3). В нулевом порядке по t из уравнений для геодезических следует

$$\ddot{x}_0^\alpha = -\overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha \dot{x}_0^\beta \dot{x}_0^\gamma, \qquad \text{где } \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha = \Gamma_{\beta\gamma}{}^\alpha(x_0).$$

Первый порядок уравнений по t определяет кубический член разложения в (17.61)

$$\dddot{x}_0^\alpha = (-\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha + 2\overset{\circ}{\Gamma}_{\beta\gamma}{}^\epsilon \overset{\circ}{\Gamma}_{\{\delta\epsilon\}}{}^\alpha)\dot{x}_0^\beta \dot{x}_0^\gamma \dot{x}_0^\delta,$$

где, для краткости, мы допускаем некоторую вольность в обозначениях:

$$\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha = \partial_\delta \Gamma_{\beta\gamma}{}^\alpha|_{x=x_0}\,,$$

и фигурные скобки обозначают симметризацию по индексам.

Следовательно, решение задачи Коши для геодезических в третьем порядке по t имеет вид

$$\begin{aligned} x^\alpha = x_0^\alpha + \dot{x}_0^\alpha t - \frac{1}{2}\overset{\circ}{\Gamma}_{0\beta\gamma}{}^\alpha \dot{x}_0^\beta \dot{x}_0^\gamma t^2 \\ - \frac{1}{6}(\partial_\delta \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha - 2\overset{\circ}{\Gamma}_{\beta\gamma}{}^\epsilon \overset{\circ}{\Gamma}_{\{\delta\epsilon\}}{}^\alpha)\dot{x}_0^\beta \dot{x}_0^\gamma \dot{x}_0^\delta t^3 + \dots . \end{aligned} \tag{17.62}$$

Коэффициенты при более высоких степенях t^k, $k \geq 4$ пропорциональны $(\dot{x}_0)^k t^k$ с коэффициентами, зависящими от аффинной связности и их частных производных вплоть до $k-2$ порядка, вычисленными в точке x_0. Интервал сходимости ряда (17.62) определяется начальными данными и компонентами аффинной связности. Мы, конечно, предполагаем, что он больше нуля.

Перейдем к определению нормальных координат в окрестности $\mathbb{U}_0 \subset \mathbb{M}$ точки x_0. Пусть $\mathbb{M}$ – многообразие с заданной аффинной связностью без кручения, т.е. $\Gamma_{\beta\gamma}{}^\alpha = \Gamma_{\gamma\beta}{}^\alpha$. Перепишем закон преобразования компонент аффинной связности (7.5) в виде

$$\Gamma_{\beta'\gamma'}{}^{\alpha'} = \partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma (\Gamma_{\beta\gamma}{}^\alpha \partial_\alpha x^{\alpha'} - \partial^2_{\beta\gamma} x^{\alpha'}).$$

Совершим преобразование координат, которое задается квадратичным полиномом с постоянными коэффициентами

$$x^{\alpha'} = B_\alpha{}^{\alpha'}(x^\alpha - x_0^\alpha) + \frac{1}{2}\overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha B_\alpha{}^{\alpha'}(x^\beta - x_0^\beta)(x^\gamma - x_0^\gamma), \tag{17.63}$$

где $B_\alpha{}^{\alpha'}$ – произвольная невырожденная матрица. Тогда в новой системе координат компоненты аффинной связности будут равны

$$\Gamma_{\beta'\gamma'}{}^{\alpha'} = \partial_{\beta'} x^\beta \partial_{\gamma'} x^\gamma \left[\Gamma_{\beta\gamma}{}^\alpha \left(B_\alpha{}^{\alpha'} + \overset{\circ}{\Gamma}_{\alpha\delta}{}^\epsilon B_\epsilon{}^{\alpha'}(x^\delta - x_0^\delta)\right) - \overset{\circ}{\Gamma}_{\beta\gamma}{}^\alpha B_\alpha{}^{\alpha'}\right].$$

Отсюда следует, что в точке x_0 компоненты аффинной связности без кручения обращаются в нуль.

Замечание. Конечно, если бы связность обладала кручением, то, поскольку кручение является тензором, его нельзя обратить в нуль никаким преобразованием координат даже в одной точке. □

Выше мы доказали, что для произвольной аффинной связности в произвольной заданной точке x_0 всегда можно обратить в нуль симметричную часть связности. Для этого достаточно выбрать соответствующим образом только квадратичные члены в функциях преобразований координат (17.63). Данная система координат существует в окрестности произвольной точки $x_0 \in \mathbb{M}$ и определена неоднозначно. Во-первых, матрица $B_\alpha{}^{\alpha'}$ в (17.63) является невырожденной, а в остальном произвольна. Во-вторых, к правилу преобразования координат (17.63) можно добавить произвольные слагаемые третьей и более высокой степени по $(x - x_0)$. Этот произвол в выборе системы координат можно использовать для дальнейшей специализации системы координат.

Геометрический смысл построенной системы координат состоит в том, что в достаточно малой окрестности точки x_0 свойства многообразия близки к свойствам аффинного пространства, т.к. при параллельном переносе компоненты тензоров в линейном приближении по вектору смещения не меняются. Как и декартовы координаты в евклидовом пространстве, данные координаты в точке x_0 определены, в частности, с точностью до линейных преобразований.

Можно доказать, что симметричную часть компонент аффинной связности можно обратить в нуль не только в фиксированной точке, но и вдоль произвольной кривой γ на многообразии $\mathbb{M}$ [**?**]. Соответствующая система координат называется *геодезической вдоль кривой* $\gamma \in \mathbb{M}$.

Оставшийся произвол в выборе системы координат можно использовать для дальнейшей специализации геометрических объектов. Воспользуемся свободой добавления высших степеней $(x - x_0)^k$, $k \geq 3$, в закон преобразования координат (17.63) и покажем, что в произвольной точке $x_0 \in \mathbb{M}$ можно обратить в нуль не только симметричные части самих компонент аффинной связности, но и все их полностью симметризованные частные производные.

Теорема 17.9.1. *Если компоненты аффинной связности вещественно аналитичны, то в окрестности произвольной точки $x_0 \in \mathbb{M}$ существует такая система координат, что выполнены равенства*

$$\mathring{\Gamma}_{\{\gamma_1\gamma_2\}}{}^{\alpha} = 0, \qquad \partial_{\{\gamma_3}\mathring{\Gamma}_{\gamma_1\gamma_2\}}{}^{\alpha} = 0, \qquad \ldots \qquad \partial^{k-2}_{\{\gamma_3\ldots\gamma_k}\mathring{\Gamma}_{\gamma_1\gamma_2\}}{}^{\alpha} = 0, \tag{17.64}$$

где фигурные скобки обозначают симметризацию по всем индексам, заключенным между ними.

Доказательство. Перепишем закон преобразования компонент аффинной связности (6.31) в виде

$$\Gamma_{\beta'\gamma'}{}^{\alpha'} A_{\alpha'}{}^{\alpha} = A_{\beta'}{}^{\beta} A_{\gamma'}{}^{\gamma} \Gamma_{\beta\gamma}{}^{\alpha} + \partial_{\beta'} A_{\gamma'}{}^{\alpha}, \tag{17.65}$$

где матрица $A_{\alpha'}{}^{\alpha}(x) = J^{-1}{}^{\alpha}_{\alpha'} = \partial_{\alpha'} x^{\alpha}$ является обратным якобианом преобразования координат. Дифференцирование этого соотношения по $x^{\delta'}$ приводит к равенству

$$\begin{aligned} \partial_{\delta'}\Gamma_{\beta'\gamma'}{}^{\alpha'} A_{\alpha'}{}^{\alpha} + \Gamma_{\beta'\gamma'}{}^{\alpha'} \partial_{\delta'} A_{\alpha'}{}^{\alpha} = {} & \partial_{\delta'} A_{\beta'}{}^{\beta} A_{\gamma'}{}^{\gamma} \Gamma_{\beta\gamma}{}^{\alpha} + A_{\beta'}{}^{\beta} \partial_{\delta'} A_{\gamma'}{}^{\gamma} \Gamma_{\beta\gamma}{}^{\alpha} \\ & + A_{\beta'}{}^{\beta} A_{\gamma'}{}^{\gamma} \partial_{\delta'} \Gamma_{\beta\gamma}{}^{\alpha} + \partial^2_{\delta'\beta'} A_{\gamma'}{}^{\alpha}. \end{aligned} \tag{17.66}$$

Последовательное дифференцирование этого соотношения приводит к равенствам, содержащим старшие производные от компонент аффинной связности и обратной матрицы Якоби $A_{\alpha'}{}^{\alpha}$. Рассматривая эти тождества в точке x^0, мы докажем возможность выбора системы координат, в которой выполнены равенства (17.64), по теории возмущений.

Совершим преобразование координат $x^{\alpha} = x^{\alpha}(y)$, которое вблизи точки $x_0 \in \mathbb{M}$ задается степенным рядом

$$x^{\alpha} = x^{\alpha}_0 + \left.\frac{\partial x^{\alpha}}{\partial y^{\gamma}}\right|_0 y^{\gamma} + \frac{1}{2} \left.\frac{\partial^2 x^{\alpha}}{\partial y^{\gamma_1} \partial y^{\gamma_2}}\right|_0 y^{\gamma_1} y^{\gamma_2} + \ldots \tag{17.67}$$

Здесь для новой системы координат мы используем букву y, чтобы избежать штрихов у индексов. Выберем первые три члена разложения в виде

$$\begin{aligned} \left.\frac{\partial x^{\alpha}}{\partial y^{\gamma}}\right|_0 &= \delta^{\alpha}_{\gamma}, \\ \left.\frac{\partial^2 x^{\alpha}}{\partial y^{\gamma_1} \partial y^{\gamma_2}}\right|_0 &= -\mathring{\Gamma}_{\{\gamma_1\gamma_2\}}{}^{\alpha}, \\ \left.\frac{\partial^3 x^{\alpha}}{\partial y^{\gamma_1} \partial y^{\gamma_2} \partial y^{\gamma_3}}\right|_0 &= 2\mathring{\Gamma}_{\{\gamma_1\gamma_2}{}^{\delta}\mathring{\Gamma}_{\delta\gamma_3\}}{}^{\alpha} - \partial_{\{\gamma_3}\mathring{\Gamma}_{\gamma_1\gamma_2\}}{}^{\alpha}. \end{aligned}$$

Тогда из уравнений (17.65), (17.66) следует, что в новой системе координат

$$\overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}{}^{\alpha} = 0, \qquad \partial_{\{\gamma_3}\overset{\circ}{\Gamma}_{\gamma_1\gamma_2\}}{}^{\alpha} = 0.$$

Выбор квадратичного члена ряда (17.67) уже был использован ранее. Выбор кубического члена разложения позволил обратить в нуль симметризованную первую частную производную аффинной связности в той же точке.

В дифференциальные тождества, получаемые последовательным дифференцированием (17.65), максимальная степень производной от обратной матрицы Якоби $A_{\alpha'}{}^{\alpha}$ всегда входит линейно и на единицу превышает максимальную степень производной от компонент аффинной связности. Это значит, что коэффициенты ряда (17.67) всегда можно подобрать таким образом, чтобы обратить в нуль все симметризованные частные производные от коэффициентов аффинной связности (17.64). В этом нет ничего удивительного, т.к. члены ряда (17.67), начиная с квадратичного, находятся во взаимно однозначном соответствии с условиями на аффинную связность (17.64). □

Нетрудно проверить, что ряд (17.67), определяющий преобразование координат, в точности совпадает с рядом (17.62), определяющим геодезические линии, если положить $y^\alpha = \dot{x}_0^\alpha t$. Это означает, что в новой системе координат прямые линии $y^\alpha = X_0^\alpha t$, где $t \in \mathbb{R}$, и X_0^α – произвольные числа, из которых по-крайней мере одно отлично от нуля, являются геодезическими линиями на $\mathbb{M}$. В этом месте прослеживается аналогия с декартовой системой координат в евклидовом пространстве: геодезические являются прямыми линиями.

Построенная система координат является нормальной системой координат в окрестности $\mathbb{U}_0$ точки $x_0 \in \mathbb{M}$, определяется только симметричной частью компонент аффинной связностью и никак от метрики не зависит. При этом у нас остался произвол в выборе первых двух членов разложения функций преобразования координат (17.67). Слагаемое нулевого порядка определим так, чтобы точка x_0 отображается в начало координат $y_0 = (0, \ldots, 0)$. Слагаемое первого порядка по y^α можно использовать для фиксирования значения метрики в начале координат. Очевидно, что его всегда можно подобрать таким образом, чтобы метрика в точке x_0 была диагональной

$$\overset{\circ}{g}_{\alpha\beta} = \eta_{\alpha\beta}. \tag{17.68}$$

Таким образом мы однозначно фиксировали все члены разложения функций преобразования координат (17.67). Это доказывает следующее утверждение.

Теорема 17.9.2. *Пусть в окрестности произвольной точки $x_0 \in \mathbb{M}$ многообразия, на котором задана аффинная геометрия, метрика и связность заданы вещественно аналитическими функциями. Тогда существует такая система координат y^α, что точка x_0 соответствует началу координат, а также выполнены равенства (17.64) и (17.68). Такая система координат определена однозначно.*

Замечание. В доказательстве теоремы нигде не использовалась сигнатура метрики. Поэтому сформулированная теорема справедлива как для римановых, так и для псевдоримановых метрик. □

Определение. Система координат y^α в теореме 17.9.2 называется *нормальной*, *геодезической* или *римановой*. □

Предложение 17.9.1. *Система координат y^α в некоторой окрестности начала координат является нормальной тогда и только тогда, когда выполнены условия:*

$$\Gamma_{\beta\gamma}{}^{\alpha}(y) y^\beta y^\gamma = 0 \quad \text{и} \quad g_{\alpha\beta}(0) = \eta_{\alpha\beta}. \tag{17.69}$$

Доказательство. В нормальной системе координат геодезическая линия, проходящая через точку x_0 в направлении u_0, является прямой и задается параметрически в виде

$$y^\alpha(t) = u_0^\alpha t. \tag{17.70}$$

Подстановка этого выражения в уравнение для геодезических линий (17.3) дает

$$u_0^\alpha u_0^\beta \Gamma_{\alpha\beta}{}^\gamma = 0.$$

Умножив это уравнение на t^2, получим (17.69).

Обратно. Пусть выполнено уравнение (17.69). Разложим компоненты связности в ряд Тейлора вблизи начала координат. Тогда уравнение (17.69) эквивалентно цепочке равенств (17.64). Затем разложим уравнение геодезических и сами геодезические в ряды Тейлора по t. Приравняв нулю коэффициенты при одинаковых степенях t, получаем, что прямые линии и только они являются геодезическими линиями, проходящими через начало координат. □

Замечание. Сравнение первого условия в (17.69) со вторым условием в (17.58) показывает, что нормальные координаты не являются гармоническими. □

Нормальная система координат обладает рядом замечательных свойств. В частности, в окрестности начала координат произвольное тензорное поле можно представить в виде ряда, коэффициенты которого зависят от ковариантных производных этого поля и тензоров кривизны, кручения и их ковариантных производных, вычисленных в точке x_0. Доказательство этого утверждения проводится конструктивно, путем явного построения соответствующего разложения.

Для доказательства нам понадобится предварительный результат.

Лемма 17.9.1. *Пусть связность* Γ *на* $\mathbb{M}$ *вещественно аналитична. Тогда в окрестности начала нормальной системы координат все производные*

$$\partial^{k-1}_{\{\gamma_1\dots\gamma_{k-1}} \mathring{\Gamma}_{\gamma_k\}\alpha}{}^\beta, \tag{17.71}$$

где, в отличие от (17.64), симметризация проводится только по части нижних индексов, выражаются через тензоры кривизны, кручения и их ковариантные производные.

Доказательство. Лемма доказывается прямыми, но громоздкими вычислениями. Поэтому мы ограничимся только первыми двумя слагаемыми.

Используя свойства (17.64), нетрудно доказать равенства

$$\begin{aligned}
\partial_{\{\gamma_1} \mathring{\Gamma}_{\gamma_2\}\alpha}{}^\beta &= \frac{1}{3}\left(\partial_{\{\gamma_1} \mathring{\Gamma}_{\alpha\gamma_2\}}{}^\beta - \partial_\alpha \mathring{\Gamma}_{\{\gamma_1\gamma_2\}}{}^\beta + 2\partial_{\{\gamma_1} \mathring{T}_{\gamma_2\}\alpha}{}^\beta\right), \\
\partial^2_{\{\gamma_1\gamma_2} \mathring{\Gamma}_{\gamma_3\}\alpha}{}^\beta &= \frac{1}{3}\left(\partial^2_{\{\gamma_1\gamma_2} \mathring{\Gamma}_{\alpha\gamma_3\}}{}^\beta - \partial^2_{\alpha\{\gamma_1} \mathring{\Gamma}_{\gamma_2\gamma_3\}}{}^\beta + 2\partial^2_{\{\gamma_1\gamma_2} \mathring{T}_{\gamma_3\}\alpha}{}^\beta\right).
\end{aligned} \tag{17.72}$$

Здесь и в дальнейшем фигурные скобки обозначают симметризацию *только* по индексам γ.

В точке x_0 аффинная связность полностью определяется тензором кручения

$$\mathring{\Gamma}_{\gamma\alpha}{}^\beta = \frac{1}{2}\mathring{T}_{\gamma\alpha}{}^\beta. \tag{17.73}$$

Далее, прямые вычисления приводят к следующим равенствам

$$\begin{aligned}
\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2\}}{}^{\beta} &= \partial_\alpha \overset{\circ}{\Gamma}_{\{\gamma_1\gamma_2\}}{}^{\beta} - \partial_{\{\gamma_1}\overset{\circ}{\Gamma}_{\alpha\gamma_2\}}{}^{\delta} - \frac{1}{4}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\overset{\circ}{T}_{\gamma_2\}\delta}{}^{\beta},\\
\nabla_{\{\gamma_1}\overset{\circ}{R}_{\alpha\gamma_2\gamma_3\}}{}^{\beta} &= \partial^2_{\alpha\{\gamma_1}\overset{\circ}{\Gamma}_{\gamma_2\gamma_3\}}{}^{\beta} - \partial^2_{\{\gamma_1\gamma_2}\overset{\circ}{\Gamma}_{\alpha\gamma_3\}}{}^{\beta} + \frac{2}{3}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta} + \frac{2}{3}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_2\gamma_3\}}{}^{\beta}\\
&\quad + \frac{1}{6}\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\alpha}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta} - \frac{1}{3}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\nabla_{\gamma_2}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta},\\
\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\}\alpha}{}^{\beta} &= \partial_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\}\alpha}{}^{\beta},\\
\nabla_{\{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{T}_{\gamma_3\}\alpha}{}^{\beta} &= \partial^2_{\{\gamma_1\gamma_2}\overset{\circ}{T}_{\gamma_3\}\alpha}{}^{\beta} + \frac{1}{6}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta} + \frac{1}{6}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_2\gamma_3\}}{}^{\beta}\\
&\quad + \frac{1}{6}\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\alpha}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\gamma} + \frac{1}{6}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\nabla_{\gamma_2}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta}.
\end{aligned}$$

Используя полученные формулы, правые части (17.72) можно записать в ковариантном виде

$$\partial_{\{\gamma_1}\Gamma_{\gamma_2\}\alpha}{}^{\beta} = \frac{1}{3}\left(-\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2\}}{}^{\beta} + 2\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\}\alpha}{}^{\beta} - \frac{1}{4}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\overset{\circ}{T}_{\gamma_2\}\delta}{}^{\beta}\right), \tag{17.74}$$

$$\begin{aligned}
\partial^2_{\{\gamma_1\gamma_2}\overset{\circ}{\Gamma}_{\gamma_3\}\alpha}{}^{\beta} &= \frac{1}{3}\left(-\nabla_{\{\gamma_1}\overset{\circ}{R}_{\alpha\gamma_2\gamma_3\}}{}^{\beta} + \frac{1}{3}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta} + \frac{1}{3}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_2\gamma_3\}}{}^{\beta}\right.\\
&\quad \left. + 2\nabla_{\{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{T}_{\gamma_3\}\alpha}{}^{\beta} - \frac{1}{6}\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\alpha}{}^{\delta}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta} - \frac{2}{3}\overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\delta}\nabla_{\gamma_2}\overset{\circ}{T}_{\gamma_3\}\delta}{}^{\beta}\right),
\end{aligned} \tag{17.75}$$

где, напомним, симметризация проводится только по индексам γ. Аналогичным образом все частные производные от связности вида (17.71) можно выразить через ковариантные объекты. □

Теперь можно доказать утверждение о разложении произвольного тензорного поля в окрестности начала нормальной системы координат в ряд Тейлора. Для определенности рассмотрим произвольный ковариантный тензор второго ранга $A_{\alpha\beta}$ в нормальных координатах и разложим его в ряд Тейлора вблизи начала координат:

$$A_{\alpha\beta}(y) = \overset{\circ}{A}_{\alpha\beta} + \partial_\gamma \overset{\circ}{A}_{\alpha\beta}\, y^\gamma + \frac{1}{2}\partial^2_{\gamma_1\gamma_2}\overset{\circ}{A}_{\alpha\beta}\, y^{\gamma_1}y^{\gamma_2} + \dots \tag{17.76}$$

Эта процедура явно нековариантна, потому что координаты y^γ не являются компонентами вектора, и коэффициенты этого ряда также нековариантны. Основным достоинством нормальных координат является то, что все коэффициенты этого ряда тем не менее можно выразить через ковариантные величины. Для этого необходимо выразить все частные производные через ковариантные и выразить компоненты аффинной связности в начале координат через тензор кривизны, кручения и их ковариантные производные. Начнем с первой производной

$$\begin{aligned}
\partial_\gamma \overset{\circ}{A}_{\alpha\beta} &= \nabla_\gamma \overset{\circ}{A}_{\alpha\beta} + \overset{\circ}{\Gamma}_{\gamma\alpha}{}^{\delta}\overset{\circ}{A}_{\delta\beta} + \overset{\circ}{\Gamma}_{\gamma\beta}{}^{\delta}\overset{\circ}{A}_{\alpha\delta}\\
&= \nabla_\gamma \overset{\circ}{A}_{\alpha\beta} + \frac{1}{2}\overset{\circ}{T}_{\gamma\alpha}{}^{\delta}\overset{\circ}{A}_{\delta\beta} + \frac{1}{2}\overset{\circ}{T}_{\gamma\beta}{}^{\delta}\overset{\circ}{A}_{\alpha\delta},
\end{aligned}$$

где мы воспользовались формулами (17.73). Прямые, но более громоздкие выкладки позволяют представить вторые частные производные также в ковариантном виде:

$$\begin{aligned}
\partial^2_{\gamma_1\gamma_2}\overset{\circ}{A}_{\alpha\beta} &= \nabla_{\{\gamma_1}\nabla_{\gamma_2\}}\overset{\circ}{A}_{\alpha\beta} + \overset{\circ}{T}_{\{\gamma_1\alpha}{}^{\delta}\nabla_{\gamma_2\}}\overset{\circ}{A}_{\delta\beta} + \overset{\circ}{T}_{\{\gamma_1\beta}{}^{\delta}\nabla_{\gamma_2\}}\overset{\circ}{A}_{\alpha\delta} + \frac{1}{2}\overset{\circ}{T}_{\{\gamma_1\alpha}{}^{\delta}\overset{\circ}{T}_{\gamma_2\}}{}^{\epsilon}\overset{\circ}{A}_{\delta\epsilon}\\
&\quad + \frac{1}{3}\left(-\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2\}}{}^{\delta} + 2\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\}\alpha}{}^{\delta} - \overset{\circ}{T}_{\alpha\{\gamma_1}{}^{\epsilon}\overset{\circ}{T}_{\gamma_2\}\epsilon}{}^{\delta}\right)\overset{\circ}{A}_{\delta\beta}\\
&\quad + \frac{1}{3}\left(-\overset{\circ}{R}_{\beta\{\gamma_1\gamma_2\}}{}^{\delta} + 2\nabla_{\{\gamma_1}\overset{\circ}{T}_{\gamma_2\}\beta}{}^{\delta} - \overset{\circ}{T}_{\beta\{\gamma_1}{}^{\epsilon}\overset{\circ}{T}_{\gamma_2\}\epsilon}{}^{\delta}\right)\overset{\circ}{A}_{\alpha\delta}.
\end{aligned}$$

Эту процедуру можно продолжить вплоть до произвольного порядка. Однако уже в третьем порядке формулы настолько громоздки, что нет смысла их приводить в явном виде.

Аналогичное представление справедливо для тензорного поля произвольного ранга и типа.

Теорема 17.9.3. *Компоненты произвольного вещественно аналитического тензорного поля в некоторой окрестности начала нормальной системы координат представимы в виде ряда Тейлора, коэффициенты которого определяются ковариантными производными компонент данного тензорного поля, а также тензором кривизны, тензором кручения и их ковариантными производными, взятыми в начале координат.*

17.9.2 Нормальные координаты в (псевдо-)римановом пространстве

Основное достоинство нормальных координат заключается в том, что вещественно аналитические тензорные поля в окрестности $\mathbb{U}_0 \subset \mathbb{M}$ произвольной точки $x_0 \in \mathbb{M}$ представляются в виде ряда Тейлора, коэффициенты которого задаются только ковариантными объектами: ковариантными производными данного тензорного поля, а также тензорами кривизны и кручения и их ковариантными производными. В предыдущем разделе были явно вычислены первые два члена этого ряда для ковариантного тензора второго ранга. Эти члены содержат много слагаемых с тензором кручения. Поэтому в (псевдо-)римановой геометрии, где кручение тождественно равно нулю, формулы упрощаются, и можно продвинуться значительно дальше в вычислениях.

Кроме того, если на многообразии $\mathbb{M}$ задана аффинная геометрия $(\mathbb{M}, g, \Gamma)$, то в качестве определяющего пучка кривых, проходящих через точку $x_0 \in \mathbb{M}$ можно выбрать не геодезические линии, а экстремали, которые определяются исключительно метрикой. Это также дает основание рассмотреть нормальные координаты в (псевдо-)римановой геометрии более подробно.

Два слова об обозначениях. Мы часто используем знак тильды для геометрических объектов в (псевдо-)римановой геометрии. Поскольку значок окружности над символом уже используется для обозначения геометрических объектов, рассматриваемых в точке x_0, то, чтобы не загромождать обозначений, знак тильды в настоящем разделе мы опустим.

В (псевдо-)римановой геометрии в нормальных координатах $\overset{\circ}{\Gamma}_{\alpha\beta}{}^{\gamma} = 0$, и ковариантная производная произвольного тензора в этой точке совпадает с частной производной. При этом в выражении для тензора кривизны пропадают квадратичные слагаемые по связности:

$$\overset{\circ}{R}_{\alpha\beta\gamma}{}^{\delta} = \partial_\alpha \overset{\circ}{\Gamma}_{\beta\gamma}{}^{\delta} - \partial_\beta \overset{\circ}{\Gamma}_{\alpha\gamma}{}^{\delta}$$

или

$$\overset{\circ}{R}_{\alpha\beta\gamma\delta} = \frac{1}{2}(\partial^2_{\alpha\gamma}\overset{\circ}{g}_{\beta\delta} - \partial^2_{\alpha\delta}\overset{\circ}{g}_{\beta\gamma} - \partial^2_{\beta\gamma}\overset{\circ}{g}_{\alpha\delta} + \partial^2_{\beta\delta}\overset{\circ}{g}_{\alpha\gamma}).$$

Вычисления, аналогичные тем, что привели к формулам (17.74), (17.75), в (псевдо-

)римановой геометрии дают

$$\begin{aligned}
\partial_{\{\gamma_1}\overset{\circ}{\Gamma}_{\gamma_2\}\alpha}{}^{\beta} &= -\frac{1}{3}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2\}}{}^{\beta},\\
\partial^2_{\{\gamma_1\gamma_2}\overset{\circ}{\Gamma}_{\gamma_3\}\alpha}{}^{\beta} &= -\frac{1}{2}\nabla_{\{\gamma_1}\overset{\circ}{R}_{\alpha\gamma_2\gamma_3\}}{}^{\beta},\\
\partial^3_{\{\gamma_1\gamma_2\gamma_3}\overset{\circ}{\Gamma}_{\gamma_4\}\alpha}{}^{\beta} &= -\frac{3}{5}\nabla_{\{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{R}_{\alpha\gamma_3\gamma_4\}}{}^{\beta} - \frac{2}{15}\overset{\circ}{R}_{\alpha\{\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_3\gamma_4\}}{}^{\beta},
\end{aligned} \tag{17.77}$$

где симметризация проводится только по индексам γ. Здесь мы дополнительно вычислили третью симметризованную производную от связности.

Разложим ковариантный тензор $A_{\alpha_1\ldots\alpha_s}$ произвольного ранга s в в окрестности точки x_0 в ряд по нормальным координатам:

$$A_{\alpha_1\ldots\alpha_s} = \overset{\circ}{A}_{\alpha_1\ldots\alpha_s} + \partial_\gamma\overset{\circ}{A}_{\alpha_1\ldots\alpha_s}\,y^\gamma + \frac{1}{2}\partial^2_{\gamma_1\gamma_2}\overset{\circ}{A}_{\alpha_1\ldots\alpha_s}\,y^{\gamma_1}y^{\gamma_2} + \ldots,$$

где $x_0 = \{y^\alpha = 0\}$. Преобразовав частные производные в ковариантные и исключив слагаемые со связностью с помощью формул (17.77), получим следующий ряд

$$\begin{aligned}
A_{\alpha_1\ldots\alpha_s} &= \overset{\circ}{A}_{\alpha_1\ldots\alpha_s} + \nabla_\gamma\overset{\circ}{A}_{\alpha_1\ldots\alpha_s}y^\gamma\\
&+ \frac{1}{2!}\left(\nabla_{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{A}_{\alpha_1\ldots\alpha_s} - \frac{1}{3}\sum_{k=1}^{s}\overset{\circ}{R}_{\alpha_k\gamma_1\gamma_2}{}^{\beta_k}\overset{\circ}{A}_{\alpha_1\ldots\beta_k\ldots\alpha_s}\right)y^{\gamma_1}y^{\gamma_2}\\
&+ \frac{1}{3!}\left(\nabla_{\gamma_1}\nabla_{\gamma_2}\nabla_{\gamma_3}\overset{\circ}{A}_{\alpha_1\ldots\alpha_s} - \sum_{k=1}^{s}\overset{\circ}{R}_{\alpha_k\gamma_1\gamma_2}{}^{\beta_k}\nabla_{\gamma_3}\overset{\circ}{A}_{\alpha_1\ldots\beta_k\ldots\alpha_s}\right.\\
&\qquad\left. -\frac{1}{2}\sum_{k=1}^{s}\nabla_{\gamma_1}\overset{\circ}{R}_{\alpha_k\gamma_2\gamma_3}{}^{\beta_k}\overset{\circ}{A}_{\alpha_1\ldots\beta_k\ldots\alpha_s}\right)y^{\gamma_1}y^{\gamma_2}y^{\gamma_3} + \ldots
\end{aligned} \tag{17.78}$$

Под знаком суммы подразумевается суммирование по немому индексу β_k, который стоит на k-том месте в $\overset{\circ}{A}_{\alpha_1\ldots\beta_k\ldots\alpha_s}$. Симметризацию по индексам γ можно не указывать, т.к. происходит свертка с симметричным произведением $y^{\gamma_1}\ldots y^{\gamma_k}$.

Аналогичные ряды можно построить для тензоров, содержащих произвольные наборы ковариантных и контравариантных индексов. Во всех случаях коэффициенты ряда вплоть до любого порядка могут быть выражены только через ковариантные объекты в точке x_0.

Нормальные координаты особенно удобны при анализе (псевдо-)римановой метрики. В этом случае возникает дополнительное упрощение, поскольку все ковариантные производные от метрики тождественно равны нулю. Несложные вычисления показывают, что ряд (17.78) для метрики в нормальных координатах принимает вид

$$\begin{aligned}
g_{\alpha\beta} &= \eta_{\alpha\beta} - \frac{1}{3}\overset{\circ}{R}_{\alpha\gamma_1\gamma_2\beta}\,y^{\gamma_1}y^{\gamma_2} - \frac{1}{3!}\nabla_{\gamma_1}\overset{\circ}{R}_{\alpha\gamma_2\gamma_3\beta}\,y^{\gamma_1}y^{\gamma_2}y^{\gamma_3}\\
&+ \frac{1}{5!}\left(-6\nabla_{\gamma_1}\nabla_{\gamma_2}\overset{\circ}{R}_{\alpha\gamma_3\gamma_4\beta} + \frac{16}{3}\overset{\circ}{R}_{\alpha\gamma_1\gamma_2}{}^{\delta}\overset{\circ}{R}_{\delta\gamma_3\gamma_4\beta}\right)y^{\gamma_1}y^{\gamma_2}y^{\gamma_3}y^{\gamma_4} + \ldots,
\end{aligned} \tag{17.79}$$

где мы также вычислили слагаемое четвертого порядка.

17.9.3 (Псевдо-)римановы пространства постоянной кривизны

Выражение для метрики в нормальных координатах (17.79) принимает особенно простой вид для пространств постоянной кривизны, которые определяются равенством $\nabla_\epsilon R_{\alpha\beta\gamma}{}^{\delta} = 0$.

Предложение 17.9.2. *В пространстве постоянной кривизны (псевдо-)риманова метрика в окрестности произвольной точки $x_0 \in \mathbb{M}$ в нормальных координатах y^α имеет вид*

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{1}{2}\sum_{k=1}^{\infty} \frac{(-1)^k 2^{2k+2}}{(2k+2)!} V_\alpha{}^{\epsilon_1} V_{\epsilon_1}{}^{\epsilon_2} \dots V_{\epsilon_{k-1}}{}^{\epsilon_k} \eta_{\epsilon_k\beta}, \tag{17.80}$$

где

$$V_{\epsilon_{k-1}}{}^{\epsilon_k} = \overset{\circ}{R}_{\epsilon_{k-1}\gamma_1\gamma_2}{}^{\epsilon_k}\, y^{\gamma_1} y^{\gamma_2}, \qquad \epsilon_0 = \alpha.$$

Доказательство. Сначала проверяем, что первые два члена суммы действительно совпадают с (17.79) для пространств постоянной кривизны. Далее доказательство проводится по индукции. □

Ряд для метрики (17.80) можно просуммировать для пространств постоянной кривизны специального вида (**??**). В этом случае в начале координат справедливо равенство

$$\overset{\circ}{R}_{\alpha\beta\gamma\delta} = -\frac{2K}{n(n-1)}(\eta_{\alpha\gamma}\eta_{\beta\delta} - \eta_{\alpha\delta}\eta_{\beta\gamma}), \qquad n = \mathsf{dim}\,\mathbb{M}, \quad K = \mathsf{const}. \tag{17.81}$$

Тогда матрица $V_\alpha{}^\beta$ пропорциональна проекционному оператору:

$$V_\alpha{}^\beta = as\left(\delta_\alpha^\beta - \frac{y_\alpha y^\beta}{s}\right),$$

где

$$a = \frac{2K}{n(n-1)}, \qquad s = y^\alpha y_\alpha, \qquad y_\alpha = y^\beta \eta_{\alpha\beta},$$

и ряд для метрики принимает вид

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{1}{2}\left(\eta_{\alpha\beta} - \frac{y_\alpha y_\beta}{s}\right)\sum_{k=1}^{\infty} \frac{2^{2k+2}}{(2k+2)!}(-as)^k. \tag{17.82}$$

Это выражение для метрики определено и для $s = 0$, т.к. ряд начинается с линейного по s члена. Теперь ряд можно просуммировать, что приводит к следующему выражению для метрики

$$g_{\alpha\beta} = f(s)\Pi^{\mathrm{T}}_{\alpha\beta} + \Pi^{\mathrm{L}}_{\alpha\beta} = f\eta_{\alpha\beta} + (1-f)\frac{y_\alpha y_\beta}{s}, \tag{17.83}$$

где метрика представлена в виде суммы проекционных операторов

$$\Pi^{\mathrm{T}}_{\alpha\beta} = \eta_{\alpha\beta} - \frac{y_\alpha y_\beta}{s}, \qquad \Pi^{\mathrm{L}}_{\alpha\beta} = \frac{y_\alpha y_\beta}{s},$$

а функция $f(s)$ определена рядом

$$f(s) := \sum_{k=0}^{\infty} \frac{2^{2k+2}}{(2k+2)!}(-as)^k, \tag{17.84}$$

который сходится на всей комплексной плоскости s. Тем самым доказано следующее утверждение.

Теорема 17.9.4. *Пусть на многообразии задана (псевдо-)риманова метрика g класса $\mathcal{C}^2$ такая, что многообразие является пространством постоянной кривизны специального вида:*

$$R_{\alpha\beta\gamma\delta} = -\frac{2K}{n(n-1)}(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}). \tag{17.85}$$

Тогда метрика g на $\mathbb{M}$ вещественно аналитична.

Доказательство. Прямая проверка показывает, что вещественно аналитическая метрика (17.83), (17.84) описывает многообразие постоянной кривизны (17.85). Единственность метрики следует из единственности решения задачи Коши для уравнений геодезических. □

Ряд (17.84) можно просуммировать. Для псевдоримановых пространств положительной кривизны $K > 0, a > 0$, и функция $f(s)$ имеет вид

$$f(s) = \begin{cases} \dfrac{\sin^2\sqrt{as}}{as}, & s > 0, \\ 1, & s = 0, \\ -\dfrac{\operatorname{sh}^2\sqrt{-as}}{as}, & s < 0. \end{cases} \tag{17.86}$$

Функция $f(s)$, как нетрудно проверить, аналитична. Напомним, что $\sin ix = i \operatorname{sh} x$.

Для псевдоримановых пространств отрицательной кривизны $K < 0, a < 0$ имеем

$$f(s) = \begin{cases} -\dfrac{\operatorname{sh}^2\sqrt{-as}}{as}, & s > 0, \\ 1, & s = 0, \\ \dfrac{\sin^2\sqrt{as}}{as}, & s < 0. \end{cases} \tag{17.87}$$

Отметим, что на световом конусе, $s = 0$, метрика в обоих случаях совпадает с метрикой Минковского.

Если кривизна псевдориманова пространства постоянной кривизны равна нулю, $K = 0$, то $f = 1$.

Для римановых пространств постоянной кривизны всегда $s > 0$, и

$$f(s) = \begin{cases} \dfrac{\sin^2\sqrt{as}}{as}, & K > 0, \\ 1, & K = 0, \\ -\dfrac{\operatorname{sh}^2\sqrt{-as}}{as}, & K < 0. \end{cases} \tag{17.88}$$

Тот факт, что функция $f(s)$ действительно приводит к ряду (17.82) проверяется прямой проверкой. Не зная функции $f(s)$, ряд (17.82) можно просуммировать следующим образом. Представление (17.82) задает тензорную структуру метрики. В разделе 16.4 для метрики более общего вида, которая параметризуется двумя произвольными функциями $f(s)$ и $g(s)$, были вычислены символы Кристоффеля и тензор кривизны. Рассматриваемому случаю соответствует условие

$$g = \frac{(f + f's)^2}{f - afs} = 1,$$

где штрих обозначает дифференцирование по s. Последнее равенство задает обыкновенное дифференциальное уравнение первого порядка на $f(s)$. Это уравнение легко решается после подстановки

$$U = fs \tag{17.89}$$

Постоянная интегрирования находится из условия $U(0) = 0$ (ограниченность метрики).

Рассмотрим римановы пространства постоянной кривизны, для которых $g_{\alpha\beta}(0) = \delta_{\alpha\beta}$. Метрика (17.83) принимает особенно простой вид в сферической системе координат евклидова пространства $\mathbb{R}^n$. В сферических координатах $s = y^\alpha y^\beta \delta_{\alpha\beta} = r^2$, где r – радиальная координата, и справедливы тождества

$$\Pi^{\text{L}}_{\alpha\beta} dy^\alpha dy^\beta = dr^2, \quad \Pi^{\text{T}}_{\alpha\beta} dy^\alpha dy^\beta = r^2 d\Omega,$$

где $d\Omega$ – элемент телесного угла в евклидовом пространстве $\mathbb{R}^n$. Таким образом метрику пространств постоянной кривизны в нормальных координатах можно записать в виде

$$ds^2 = dr^2 + r^2 f d\Omega. \tag{17.90}$$

Проанализируем пространство $\mathbb{R}^n$ с метрикой (17.90) подробнее. Объем сферы радиуса r равен (**??**) в $\mathbb{R}^n$ с метрикой (17.90) равен

$$S^{n-1}_r = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} (r^2 f)^{(n-1)/2}.$$

Мы видим, что нули функции $f(r^2) = 0$ определяют те сферы в $\mathbb{R}^n$, которые на самом деле являются точками пространства постоянной кривизны.

Нормальные координаты для многообразий постоянной кривизны определены для всех $\{y^\alpha\} \in \mathbb{R}^n$. При этом все геодезические, проходящие через точку x_0, полны, т.к. канонический параметр пробегает всю вещественную прямую $t \in (-\infty, \infty)$.

Проведенное рассмотрение доказывает

Теорема 17.9.5. *Нормальные координаты на (псевдо-)римановом многообразии $\mathbb{M}$ постоянной кривизны вида (17.85) в каждой точке $x_0 \in \mathbb{M}$ задают гладкое сюръективное отображение*

$$\mathbb{R}^n \to \mathbb{M}. \tag{17.91}$$

Для римановых пространств нулевой кривизны нормальные координаты совпадают с декартовыми.

Замечание. В дальнейшем мы увидим, что отображение (17.91) не является накрытием. □

В общем случае отображение (17.91) не является взаимно однозначным. Поэтому в области определения нормальных координат $\mathbb{R}^n$ можно задать отношение эквивалентности, отождествив те точки, которые отображаются на одну и ту же точку из $\mathbb{M}$. Таким образом пространство постоянной кривизны вида (17.85) можно рассматривать как евклидово пространство $\mathbb{R}^n$, в котором задано некоторое отношение эквивалентности между точками.

Пример 17.9.2. Рассмотрим двумерную сферу $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ единичного радиуса в качестве пространства постоянной положительной кривизны. В этом случае $K = 1, a = 1, n = 2$. Функция $f(r^2)$ в полярных координатах на плоскости $\mathbb{R}^2$ имеет вид

$$f = \frac{\sin^2 r}{r^2}.$$

Это соответствует метрике

$$ds^2 = dr^2 + \sin^2 r d\varphi^2.$$

Длина окружности на плоскости $\mathbb{R}^2$ радиуса r с центром в начале координат равна

$$L = \int_0^{2\pi} d\varphi \sin r = 2\pi \sin r.$$

Отсюда следует, что окружности радиусов $r = \pi k$, $k = 1, 2, \ldots$ отображаются в одну точку сферы. При этом плоскость $\mathbb{R}^2$ бесконечное число раз "накрывает" сферу $\mathbb{S}^2$. Если условиться, что начало координат соответствует южному полюсу сферы, то при отображении $\mathbb{R}^2 \to \mathbb{S}^2$ все окружности радиуса $r = 2\pi k$, $k = 0, 1, \ldots$, соответствуют южному полюсу, а все окружности радиуса $r = \pi + 2\pi k$ – северному. В рассматриваемом случае между точками евклидовой плоскости возникает отношение эквивалентности

$$y^\alpha \sim y^\alpha + \frac{y^\alpha}{r} 2\pi k, \qquad k = 0, 1, \ldots . \qquad \square$$

Нормальные координаты были определены таким образом, что геодезические линии в них совпадают с прямыми. В (псевдо-)римановом пространстве экстремали совпадают с геодезическими и поэтому также являются прямыми. Проверим это для пространств постоянной кривизны, которые были рассмотрены выше. Уравнения для экстремалей, определяемых метрикой (17.83) можно проинтегрировать. Из выражения для символов Кристоффеля (16.31), которые в рассматриваемом случае имеют вид

$$\Gamma_{\alpha\beta}{}^\gamma = \frac{f'}{f}\left(y_\alpha \Pi_\beta^{\mathrm{T}\gamma} + y_\beta \Pi_\alpha^{\mathrm{T}\gamma}\right) + \frac{1 - f - f's}{s} y^\gamma \Pi^{\mathrm{T}}_{\alpha\beta}, \tag{17.92}$$

следуют уравнения для экстремалей (17.17)

$$\ddot{y}^\alpha = -2\frac{f'}{f}\dot{y}^\alpha y_\beta \dot{y}^\beta + 2\frac{f'}{f} y^\alpha \frac{(y_\beta \dot{y}^\beta)^2}{s} + \frac{1 - f - f's}{s^2} y^\alpha (y_\beta \dot{y}^\beta)^2 - \frac{1 - f - f's}{s} y^\alpha (\dot{y}_\beta \dot{y}^\beta). \tag{17.93}$$

Эти уравнения имеют интеграл (17.22)

$$C_0 = \dot{y}^\alpha \dot{y}^\beta g_{\alpha\beta} = \dot{y}_\alpha \dot{y}^\alpha f - \frac{(y_\alpha \dot{y}^\alpha)^2}{s}(f - 1) = \mathrm{const}.$$

Прямая проверка показывает, что все прямые линии, проходящие через начало координат,

$$y^\alpha = v^\alpha t, \qquad v^\alpha = \mathrm{const}, \quad t \in \mathbb{R},$$

являются экстремалями. Эти экстремали, очевидно, полны. Отметим, что символы Кристоффеля (17.92) равны нулю только в начале координат. В близких точках они отличны от нуля, и среди экстремалей прямыми являются только те, которые проходят через начало координат.

Уравнения для экстремалей (17.93) можно проинтегрировать и в общем случае. То есть найти те экстремали, которые не проходят через начало координат. Для этой цели рассмотрим зависимость $s = y^\alpha y^\beta \eta_{\alpha\beta}$ от t. Учитывая равенства

$$\dot{s} = 2 y_\alpha \dot{y}^\alpha,$$
$$\ddot{s} = 2 y_\alpha \ddot{y}^\alpha + 2 \dot{y}_\alpha \dot{y}^\alpha,$$

из уравнений для экстремалей получаем обыкновенное уравнение на $s(t)$:

$$\ddot{s} = \left(\frac{1}{2s} - \frac{U'}{2U}\right)\dot{s}^2 + \frac{2C_0 U' s}{U},$$

где функция $U(s)$ определена равенством (17.89)

$$U(s) = \frac{\sin^2\sqrt{as}}{a}.$$

Рассмотрим случай $C_0 a > 0$. Введя новую переменную $z^2 = as$, при $as > 0$, и растянув канонический параметр $t \to t/\sqrt{C_0 a}$, приходим к уравнению

$$\ddot{z} = (1 - \dot{z}^2)\operatorname{ctg} z,$$

которое можно явно проинтегрировать. Общее решение этого уравнения имеет вид

$$\cos\sqrt{as} = \sqrt{1 - \frac{1}{C_1}}\sin(t + t_0), \qquad C_1 = \text{const} > 1, \quad t_0 = \text{const},$$

Аналогично можно рассмотреть все остальные случаи.

Таким образом для пространств постоянной кривизны специального вида (17.81) в нормальных координатах можно найти и проанализировать поведение всех экстремалей. В дальнейшем мы рассмотрим эту задачу в ряде конкретных случаев.

17.9.4 Нормальные координаты и экспоненциальное отображение

Пусть задана связность Γ на многообразии $\mathbb{M}$ и $\gamma = x(t)$ – геодезическая. Если t – канонический параметр вдоль геодезической, то координатные функции $\{x^\alpha(t)\}$ удовлетворяют системе уравнений (17.3). Зафиксируем канонический параметр t каким либо образом и начальную точку $x_0 = x(0)$. Тогда касательный вектор $u_0 := \dot{x}_0 \in \mathbb{T}_0(\mathbb{M})$ к геодезической в точке x_0 определен однозначно. Верно также обратное утверждение. Если задана точка $x_0 \in \mathbb{M}$ и касательный вектор $u_0 \in \mathbb{T}_0(\mathbb{M})$, то существует единственная геодезическая, проходящая через x_0 с начальным вектором u_0, при этом канонический параметр определен однозначно. Таким образом каждая геодезическая однозначно определена парой (x_0, u_0).

В разделе 3.6.5 было определено экспоненциальное отображение для гладких полных векторных полей на многообразии. Геодезическая линия $x(t)$ является интегральной кривой для векторного поля скорости $u(t) := \dot{x}(t)$. Тогда для полных геодезических определено экспоненциальное отображение

$$\exp tu_0 : \quad x_0 \mapsto x(t), \tag{17.94}$$

которое отображает начальную точку x_0 в точку $x(t)$. Если вектор скорости u_0 принимает все возможные направления в касательном пространстве $\mathbb{T}_0(\mathbb{M})$, то экспоненциальное отображение (17.94) можно рассматривать, как отображение касательного пространства в многообразие

$$\exp tu_0 : \quad \mathbb{T}_0(\mathbb{M}) \ni tu_0 \mapsto x(t) \in \mathbb{M}, \tag{17.95}$$

для которого мы сохраним прежнее обозначение.

Замечание. В разделе 3.6.5 экспоненциальное отображение было определено для каждого полного векторного поля. При этом вопрос о том, каким образом данное векторное поле задано, не рассматривался. Если векторное поле одно, то экспоненциальное отображение нельзя рассматривать, как отображение касательного пространства (17.95), т.к. в точке x_0 имеется только один вектор. В рассматриваемом

случае ситуация другая. На многообразии $\mathbb{M}$ задана связность Γ, которая определяет все множество геодезических. Для фиксированной точки x_0 мы рассматриваем множество геодезических, проходящих через данную точку во всех возможных направлениях. То есть вектор скорости u_0 в (17.95) принимает все возможные направления. Следовательно, экспоненциальное отображение можно рассматривать, как отображение касательного пространства. □

По построению, каждая прямая в касательном пространстве $\mathbb{T}_0(\mathbb{M})$, проходящая через начало координат, отображается в соответствующую геодезическую. Ясно, что такое отображение можно построить для каждой точки $x_0 \in \mathbb{M}$.

Если геодезическая неполна, то экспоненциальное отображение (17.95) определено только для некоторого интервала значений канонического параметра $-\epsilon_1 < t < \epsilon_2$, где $\epsilon_{1,2} > 0$. В результате экспоненциальное отображение будет определено в некоторой окрестности начала координат касательного пространства. Если связность на $\mathbb{M}$ класса $\mathcal{C}^\infty$, то экспоненциальное отображение будет того же класса гладкости $\mathcal{C}^\infty$. Поскольку дифференциал экспоненциального отображения в точке x_0 невырожден, то существует окрестность $\mathbb{U}_0 \ni x_0$, такая, что экспоненциальное отображение (17.95) является диффеоморфизмом $\mathbb{U}_0 \to \mathbb{V}_0 \subset \mathbb{T}_0(\mathbb{M})$, где $\mathbb{V}_0$ – некоторая окрестность касательного пространства, содержащая начало координат. Выберем координатный репер $\{\partial_\alpha\}$ в точке x_0 и предположим, что векторы скорости принимают значение на единичной сфере

$$\sum_{\alpha=1}^{n} (u_0^\alpha)^2 = 1.$$

Теперь отождествим касательное пространство $\mathbb{T}_0(\mathbb{M})$ с евклидовым пространством $\mathbb{R}^n$ естественным образом, отождествив координаты касательного вектора $\{tu_0^\alpha\}$ с декартовыми координатами точки $\{y^\alpha\}$ в $\mathbb{R}^n$. В результате получим координатную систему, определенную на $\mathbb{U}_0$.

Определение. Система координат в окрестности $\mathbb{U}_0 \subset \mathbb{M}$ точки x_0, определенная экспоненциальным отображением (17.95),

$$\varphi: \quad \mathbb{M} \supset \mathbb{U}_0 \ni \{x^\alpha(t)\} \mapsto \{y^\alpha := tu_0^\alpha\} \in \mathbb{V}_0 \subset \mathbb{R}^n,$$

называется *нормальной*. □

Замечание. Подчеркнем, что нормальная система координат определена исключительно связностью, а не метрикой, которой вообще может не быть на многообразии. □

Если на многообразии $\mathbb{M}$ помимо связности Γ задана также метрика g, то исходную систему координат x^α в окрестности точки x_0 можно всегда выбрать таким образом, что координатный базис ∂_α будем ортонормальным в данной точке x_0, т.е. $g_{\alpha\beta}(x_0) = \eta_{\alpha\beta}$. Мы всегда предполагаем, что для нормальной системы координат при наличии метрики данное условие выполнено.

Теорема 17.9.6 (Уайтхед). *Пусть y^α – нормальная система координат в окрестности точки $x_0 \in \mathbb{M}$. Определим окрестность $\mathbb{U}_0(\rho)$ точки x_0 равенством $\sum_\alpha (y^\alpha)^2 < \rho^2$. Тогда существует положительное число r такое, что если $0 < \rho < r$, то:*

1) *окрестность $\mathbb{U}_0(\rho)$ является геодезически выпуклой, т.е. любые две точки из $\mathbb{U}_0(\rho)$ можно соединить геодезической, целиком лежащей в $\mathbb{U}_0(\rho)$;*
2) *каждая точка из $\mathbb{U}_0(\rho)$ имеет нормальную координатную окрестность, содержащую $\mathbb{U}_0(\rho)$.*

Доказательство. См. [?]. □

Замечание. Данная теорема справедлива для произвольных достаточно гладких связностей, независимо от того задана ли на многообразии метрика или нет. □

17.10 Полнота римановых многообразий

Пусть задано риманово многообразие $(\mathbb{M}, g)$. В настоящем разделе положительная определенность метрики существенна, и мы будем ее предполагать.

Определение. Рассмотрим две произвольные точки многообразия $p, q \in \mathbb{M}$. *Расстоянием* между этими точками называется нижняя грань интегралов

$$l(p,q) = \inf \int_p^q ds, \qquad ds = \sqrt{\dot{x}^\alpha \dot{x}^\beta g_{\alpha\beta}} \tag{17.96}$$

по всем кусочно дифференцируемым кривым класса $\mathcal{C}^1$, соединяющим эти точки. □

Функция расстояния определяет отображение

$$l: \quad \mathbb{M} \times \mathbb{M} \ni (p,q) \mapsto l(p,q) \in \mathbb{R}_+.$$

Нетрудно проверить, что все свойства топологической метрики, рассмотренной в разделе (2.3.2) выполнены и, следовательно, функция $l(p,q)$ задает топологическую метрику на $\mathbb{M}$.

Определение. Множество точек

$$\mathbb{B}_r(p) = \{x \in \mathbb{M}: \quad l(x,p) < r\} \tag{17.97}$$

называется *шаром* радиуса r с центром в точке $p \in \mathbb{M}$. □

Предложение 17.10.1. *Функция расстояния (17.96) непрерывна и семейство метрических шаров для всех $p \in \mathbb{M}$ и $r \in \mathbb{R}_+$ образует базу исходной топологии многообразия.*

Доказательство. Следует из непрерывной зависимости интеграла от пределов интегрирования. □

Для связностей Леви–Чивиты на римановых пространствах справедлива

Теорема 17.10.1 (Майерс, Стинрод). *Пусть $(\mathbb{M}_1, g_1)$ и $(\mathbb{M}_2, g_2)$ – римановы многообразия. Пусть l_1 и l_2 – функции расстояния на $\mathbb{M}_1$ и $\mathbb{M}_2$ соответственно. Если задано отображение $f: \mathbb{M}_1 \to \mathbb{M}_2$ (которое не предполагается непрерывным или дифференцируемым), такое, что $l_1(p,q) = l_2\Big(f(p), f(q)\Big)$ для всех $p, q \in \mathbb{M}_1$, то f есть диффеоморфизм из $\mathbb{M}_1$ на $\mathbb{M}_2$, который отображает метрику g_1 в g_2.*

В частности, каждое отображение f из $\mathbb{M}$ на себя, сохраняющее функцию расстояния l, есть изометрия, т.е. отображение f сохраняет метрику g.

Доказательство. См. [?]. □

Напомним, что топологическое пространство $\mathbb{M}$ и, в частности, многообразие, называется метрически полным, если любая фундаментальная последовательность в $\mathbb{M}$ сходится к некоторой точке из $\mathbb{M}$.

С другой стороны, в разделе 17.2 мы определили экстремали как линии, для которых первая вариация интеграла (17.96) равна нулю. Экстремали являются одновременно геодезическими линиями для связности Леви–Чивиты. Эта связность является полной, если любую экстремаль можно продолжить в обе стороны до бесконечных значений канонического параметра. В этом случае мы говорим, что риманово многообразие полно.

Таким образом было введено два понятия полноты многообразия: метрическая полнота и полнота связности Леви–Чивиты. Оба эти понятия определяются одним геометрическим объектом – римановой метрикой и поэтому между ними существует тесная связь, которую мы здесь рассмотрим.

Теорема 17.10.2 (**Хопф, Ринов**). *Для связного риманова многообразия следующие условия эквивалентны:*

1) $(\mathbb{M}, g)$ *– полное риманово многообразие;*
2) $(\mathbb{M}, l)$ *– полное метрическое пространство;*
3) каждый замкнутый метрический шар $\overline{\mathbb{B}}_r(p)$ *в* $\mathbb{M}$ *компактен;*
4) для каждой точки $x_0 \in \mathbb{M}$ *экспоненциальное отображение (17.95) определено на всем касательном пространстве* $\mathbb{T}_0(\mathbb{M})$.

Доказательство. См. [**?**]. □

Теорема 17.10.3. *Если* $\mathbb{M}$ *– связное полное риманово многообразие, то* $\mathbb{M}$ *геодезически выпукло, т.е. любые две точки* $p, q \in \mathbb{M}$ *можно соединить минимизирующей экстремалью.*

Доказательство. См., например, [**?**]. □

Следствие. Если все геодезические, исходящие из любой выбранной точки p связного риманова многообразия полны, то $(\mathbb{M}, l)$ – геодезически и метрически полно.

Доказательство. См., например, [**?**]. □

Следствие. Каждое компактное риманово многообразие $(\mathbb{M}, g)$ метрически полно.

Доказательство. Следствие импликации 3) → 1) в теореме 17.10.2. □

Определение. Риманово пространство $(\mathbb{M}, g)$ называется *однородным*, если группа изометрий действует на $\mathbb{M}$ транзитивно. □

Теорема 17.10.4. *Каждое однородное риманово многообразие* $(\mathbb{M}, g)$ *полно.*

Доказательство. Пусть x – точка однородного риманова пространства $(\mathbb{M}, g)$. Тогда существует положительное число $r >$ такое, что для каждого единичного вектора $u \in \mathbb{T}_x(\mathbb{M})$ геодезическая $\exp tu$ определена для каждого $|t| \leq r$. Пусть $\gamma = x(t)$, $0 \leq t \leq s$, – произвольная геодезическая в $\mathbb{M}$ с каноническим параметром t. Покажем, что эта геодезическая γ может быть продолжена до геодезической, определенной, определенное при $0 \leq t \leq s + r$. Пусть $\imath$ – изометрия многообразия $\mathbb{M}$, которая отображает точку x в $x(s)$. Тогда дифференциал обратной изометрии $\imath_*^{-1}$ отображает касательный вектор $\dot{x}(s)$ в u

$$u = \imath_*^{-1}\dot{x}(s).$$

Поскольку $\exp tu$ есть геодезическая, проходящая через точку x, то $\imath(\exp tu)$ – геодезическая, проходящая через $x(s)$. Положим

$$x(s+t) := \imath(\exp tu) \quad \text{для} \quad 0 \le t \le r.$$

Тогда кривая $\gamma = x(t)$ при $0 \le t \le s+r$ является продолжением геодезической. □

Замечание. Доказанная теорема следует также из того общего факта, что каждое локально компактное метрическое однородное пространство полно. □

Теорема 17.10.5. *Пусть $\widetilde{\mathbb{M}}$ и $\mathbb{M}$ – связные римановы многообразия одинаковой размерности. Допустим, что существует изометрическое погружение $p: \widetilde{\mathbb{M}} \to \mathbb{M}$. Тогда:*

1) Если $\widetilde{\mathbb{M}}$ полно, то отображение $p: \widetilde{\mathbb{M}} \to \mathbb{M}$ является накрытием и $\mathbb{M}$ полно.

2) Обратно, если $p: \widetilde{\mathbb{M}} \to \mathbb{M}$ – накрытие и $\mathbb{M}$ полно, то $\widetilde{\mathbb{M}}$ также полно.

Доказательство. См., например, [**?**]. □

Следствие. Пусть $\widetilde{\mathbb{M}}$ и $\mathbb{M}$ – связные многообразия одинаковой размерности и $p: \widetilde{\mathbb{M}} \to \mathbb{M}$ – погружение. Тогда, если многообразие $\widetilde{\mathbb{M}}$ компактно, то $\mathbb{M}$ также компактно, а p – накрывающее отображение.

Доказательство. Возьмем любую риманову метрику g на $\mathbb{M}$. Тогда p^*g – единственная метрика на $\widetilde{\mathbb{M}}$ такая, что p – изометрическое погружение. Поскольку $\widetilde{\mathbb{M}}$ компактно, то, по следствию 17.10, оно полно. Тогда из теоремы 17.10.5 следует, что p – накрывающее отображение и отсюда $\mathbb{M}$ компактно. □

Замечание. В предыдущих теореме и следствии требование одинаковой размерности многообразий $\widetilde{\mathbb{M}}$ и $\mathbb{M}$ является излишним, т.к. погружение возможно только для многообразий одинаковой размерности. □

Определение. Говорят, что связное риманово пространство *непродолжаемо*, если его нельзя изометрически вложить в другое связное риманово пространство как собственное открытое подмногообразие. □

Теорема 17.10.5 показывает, что каждое полное связное риманово многообразие непродолжаемо. Обратное утверждение неверно.

Пример 17.10.1. Пусть $\mathbb{M}$ есть евклидова плоскость с выколотым началом координат, а $\widetilde{\mathbb{M}}$ – его универсальное накрывающее пространство. Как открытое подмногообразие евклидовой плоскости $\mathbb{M}$ имеет естественную евклидову метрику, которая, очевидно неполна. На универсальной накрывающей $\widetilde{\mathbb{M}}$ также задана естественная евклидова метрика. Риманово многообразие $\widetilde{\mathbb{M}}$ неполно по теореме 17.10.5. Может быть доказано, что $\widetilde{\mathbb{M}}$ непродолжаемо. Таким образом, неполное связное риманово многообразие в общем случае может быть непродолжаемо. □

Следствие. Пусть $\mathbb{G}$ – группа изометрий связного риманова многообразия $(\mathbb{M}, g)$. Если орбита $x\mathbb{G}$ точки $x \in \mathbb{M}$ содержит открытое подмножество из $\mathbb{M}$, то орбита $x\mathbb{G}$ совпадает со всем $\mathbb{M}$. Тем самым риманово многообразие $\mathbb{M}$ однородно.

17.11 Формулы Френе

Рассмотрим произвольную гладкую кривую $\gamma = \{x^\alpha(t)\}$ на трехмерном многообразии $\mathbb{M}$, $\dim \mathbb{M} = 3$, с положительно определенной метрикой g. Тогда длина кривой (7.8) отлична от нуля. Выберем длину кривой в качестве канонического параметра t вдоль кривой. Предположим также, что на многообразии задана метрическая связность Γ (геометрия Римана–Картана). Единичный касательный вектор к кривой определяется вектором скорости

$$u^\alpha := \dot{x}^\alpha,$$

где точка обозначает дифференцирование по каноническому параметру t. Дифференцируя тождество $u^2 = 1$, получим

$$\nabla_\beta u^2 = 2u_\alpha \nabla_\beta u^\alpha = 0. \tag{17.98}$$

С учетом этого равенства нетрудно показать, что ковариантная производная (7.9) от u^α вдоль кривой γ (ускорение кривой),

$$\frac{Du^\alpha}{dt} = u^\beta \nabla_\beta u^\alpha = \frac{1}{\rho} n^\alpha, \tag{17.99}$$

где $\rho(t)$ – некоторая функция вдоль кривой, ортогональна вектору скорости. Здесь мы предполагаем, что $u^\beta \nabla_\beta u^\alpha \neq 0$, т.е. кривая γ не является геодезической. Вектор n^α всегда можно выбрать единичным:

$$(u, n) = 0, \qquad n^2 = 1. \tag{17.100}$$

Определение. Единичное векторное поле $n^\alpha(s) \in \mathcal{X}(\mathbb{M}, \gamma)$, определенное вдоль кривой, называется *главной нормалью кривой*. Функция $1/\rho(t)$ называется *кривизной кривой*. □

Поскольку пространство трехмерно, то дополним векторы u^α и n^α до ортонормированного базиса в касательном пространстве с помощью вектора *бинормали* к кривой, определяемого соотношениями

$$(u, b) = 0, \qquad (n, b) = 0, \qquad b^2 = 1. \tag{17.101}$$

Ковариантные производные от n^α и b^α можно разложить по этому базису с некоторыми коэффициентами

$$\begin{aligned} \frac{Dn^\alpha}{dt} &= au^\alpha + bn^\alpha + cb^\alpha, \\ \frac{Db^\alpha}{dt} &= a'u^\alpha + b'n^\alpha + c'b^\alpha. \end{aligned} \tag{17.102}$$

Дифференцируя тождества (17.100), (17.101) вдоль кривой, получим условия на коэффициенты разложения:

$$\frac{1}{\rho} + a = 0, \qquad b = 0, \qquad a' = 0, \qquad c + b' = 0, \qquad c' = 0.$$

Тогда из соотношений (17.99) и (17.102) следуют *формулы Френе*:

$$\begin{aligned} \frac{Du^\alpha}{dt} &= \frac{1}{\rho} n^\alpha, \\ \frac{Dn^\alpha}{dt} &= -\frac{1}{\rho} u^\alpha + \frac{1}{\tau} b^\alpha, \\ \frac{Db^\alpha}{dt} &= -\frac{1}{\tau} n^\alpha, \end{aligned} \tag{17.103}$$

где введено обозначение

$$c = -b' = \frac{1}{\tau}.$$

Функция $1/\tau(t)$ называется *кручением* кривой.

Если кривая задана, то при желании можно найти явные выражения для векторных полей u^α, n^α и b^α, а также вычислить кривизну и кручение кривой. В трехмерном евклидовом пространстве $\mathbb{R}^3$ верно также обратное утверждение (см. [**?**]). А именно, если известны кривизна и кручение как функции канонического параметра вдоль кривой, то можно восстановить кривую в $\mathbb{R}^3$ с точностью до движений (сдвигов, вращений и отражений) всего пространства. Таким образом кривизна и кручение кривой в трехмерном евклидовом пространстве представляют собой полный набор геометрических инвариантов кривой.

Пример 17.11.1. Рассмотрим спираль в трехмерном евклидовом пространстве $\mathbb{R}^3$, ось которой совпадает с осью z:

$$\begin{aligned} x &= R\cos\frac{t}{\sqrt{R^2+v^2}},\\ y &= R\sin\frac{t}{\sqrt{R^2+v^2}},\\ z &= \frac{vt}{\sqrt{R^2+v^2}}, \end{aligned} \tag{17.104}$$

где $R > 0$ и v – постоянные и параметр $t \in \mathbb{R}$ совпадает с длиной спирали. Касательный вектор к спирали имеет следующие компоненты:

$$\begin{aligned} u^x &= -\frac{R}{\sqrt{R^2+v^2}}\sin\frac{t}{\sqrt{R^2+v^2}},\\ u^y &= \frac{R}{\sqrt{R^2+v^2}}\cos\frac{t}{\sqrt{R^2+v^2}},\\ u^z &= \frac{v}{\sqrt{R^2+v^2}}. \end{aligned}$$

Дифференцирование этих равенств по t приводит к следующему вектору главной нормали:

$$\begin{aligned} n^x &= -\cos\frac{t}{\sqrt{R^2+v^2}},\\ n^y &= -\sin\frac{t}{\sqrt{R^2+v^2}},\\ n^z &= 0. \end{aligned}$$

При этом кривизна спирали равна

$$\frac{1}{\rho} = \frac{R}{R^2+v^2}.$$

Дальнейшее дифференцирования полученных равенств по t определяет вектор бинормали:

$$\begin{aligned} b^x &= \frac{v}{\sqrt{R^2+v^2}}\sin\frac{t}{\sqrt{R^2+v^2}},\\ b^y &= -\frac{v}{\sqrt{R^2+v^2}}\cos\frac{t}{\sqrt{R^2+v^2}},\\ b^z &= \frac{R}{\sqrt{R^2+v^2}} \end{aligned}$$

и кручение спирали

$$\frac{1}{\tau} = \frac{v}{R^2 + v^2}.$$

Таким образом вычислены все характеристики спирали.

Если $v = 0$, то спираль вырождается в окружность. Для окружности кривизна и кручение равны

$$\frac{1}{\rho} = \frac{1}{R}, \qquad \frac{1}{\tau} = 0. \qquad \square$$

Пример 17.11.2. Кривизна и кручение прямой линии в трехмерном евклидовом пространстве равны нулю.

Замечание. Кривизна $1/\rho$ и кручение $1/\tau$ кривой зависят от метрики g и тензора кручения T трехмерного многообразия $\mathbb{M}$, что следует из определения метрической связности. При выводе формул Френе условие метричности связности важно, т.к. при отличной от нуля неметричности формула (17.98) неверна и ковариантная производная (17.99) не будет ортогональна вектору скорости. Заметим также, что кривизна и кручение кривой являются понятиями, отличными от кривизны и кручения аффинной связности, введенных ранее. $\square$

Посмотрим на геодезические линии с точки зрения введенных выше понятий. Следующее утверждение очевидно.

Предложение 17.11.1. *Кривая γ в трехмерном пространстве Римана–Картана является геодезической тогда и только тогда, когда ее кривизна $1/\rho$ равна нулю.*

Для геодезической линии на $\mathbb{M}$ правая часть уравнения (17.99) равна нулю, и, следовательно, вектор нормали n к геодезической нельзя определить соотношением (17.99). Кручение геодезической линии также неопределено.

При смещении вдоль кривой, отличной от геодезической, на расстояние dt компоненты касательного вектора и вектора главной нормали получают приращение

$$Du^\alpha = \frac{dt}{\rho} n^\alpha, \qquad Dn^\alpha = -\frac{dt}{\rho} u^\alpha + \frac{dt}{\tau} b^\alpha.$$

Отсюда следует, что при параллельном переносе вдоль кривой на бесконечно малое расстояние касательный вектор и вектор главной нормали остаются в плоскостях, натянутых на векторы u^α и n^α, тогда и только тогда, когда кручение кривой равно нулю, $1/\tau = 0$. В этом случае векторы u^α и n^α поворачиваются на угол $d\varphi = -dt/\rho$.

Если рассматривать кривую γ на двумерном многообразии с заданной римановой метрикой и метрической связностью, то вектор бинормали тождественно равен нулю, а формулы Френе принимают вид

$$\begin{aligned} \frac{Du^\alpha}{dt} &= \frac{1}{\rho} n^\alpha, \\ \frac{Dn^\alpha}{dt} &= -\frac{1}{\rho} u^\alpha. \end{aligned}$$

В этом случае кривые нулевой кривизны и только они являются геодезическими.

Глава 18

Симплектические и пуассоновы многообразия

Симплектические и пуассоновы многообразия играют важную роль в дифференциальной геометрии в связи с применениями, в первую очередь, к гамильтоновой динамике, рассмотренной в главе 20.

18.1 Симплектические группы

Симплектические группы играют в симплектической геометрии ту же роль, что и группы вращений в евклидовой геометрии. Поскольку эти группы устроены намного сложнее, чем ортогональные и унитарные матричные группы, то мы посвятим симплектическим группам и их свойствам целый раздел. Более подробное изложение содержится в [**?**].

Определение. Рассмотрим антисимметричную $2n\times 2n$ матрицу (*каноническую симплектическую форму*),

$$\varpi = (\varpi_{\alpha\beta}) = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \tag{18.1}$$

где $\alpha, \beta = 1, \dots, 2n$ и $\mathbb{1}$ – единичная $n \times n$ матрица. Эта матрица определяет билинейную квадратичную форму в евклидовом пространстве $\mathbb{R}^{2n}$, рассматриваемом, как векторное пространство,

$$\varpi : \quad \mathbb{R}^{2n} \times \mathbb{R}^{2n} \ni X, Y \mapsto \varpi(X, Y) := X^\alpha Y^\beta \varpi_{\alpha\beta} \in \mathbb{R},$$

где X^α и Y^β – компоненты векторов в декартовой системе координат. Квадратные матрицы A размера $2n \times 2n$ с вещественными элементами, оставляющие каноническую симплектическую форму ϖ инвариантной,

$$A^{\mathrm{T}} \varpi A = \varpi, \tag{18.2}$$

образуют группу Ли $\mathbb{SP}(n, \mathbb{R})$, которая называется *вещественной симплектической группой.* □

Взятие определителя от обеих частей определения (18.2) приводит к равенству $\det A = \pm 1$, поскольку $\det \varpi = 1$. Отсюда следует, что для любой матрицы A существует обратная. Нетрудно убедиться, что обратная матрица также является симплектической, а также в том, что произведение двух симплектических матриц снова дает симплектическую матрицу. Тем самым все групповые аксиомы выполнены.

Каноническая симплектическая форма универсальна в следующем смысле. Если ω – произвольная невырожденная антисимметричная матрица, то из курса линейной алгебры известно, что за счет линейного преобразования базиса в $\mathbb{R}^{2n}$ ее всегда можно преобразовать к каноническому виду (18.1).

Предложение 18.1.1. *Алгебра Ли $\mathfrak{sp}(n,\mathbb{R})$ группы $\mathbb{SP}(n,\mathbb{R})$ состоит из матриц вида*

$$\begin{pmatrix} B & C \\ D & -B^{\mathrm{T}} \end{pmatrix} \in \mathfrak{sp}(n,\mathbb{R}), \tag{18.3}$$

где B – произвольная вещественная $n\times n$ матрица, а вещественные $n\times n$ матрицы C и D симметричны.

Доказательство. Вблизи единицы группы симплектическая матрица представима в виде

$$A = \mathsf{e}^{tM} = \mathbb{1} + tM + \ldots, \qquad t\in\mathbb{R},$$

где $M\in\mathfrak{sp}(n,\mathbb{R})$ – элемент алгебры Ли. Подставляя это разложение в (18.2) в линейном по t порядке получаем равенство

$$M^{\mathrm{T}}\varpi + \varpi M = 0. \tag{18.4}$$

Представим элемент алгебры в блочном виде

$$M = \begin{pmatrix} B & C \\ D & E \end{pmatrix}.$$

Тогда из (18.4) следуют равенства:

$$C = C^{\mathrm{T}}, \qquad D = D^{\mathrm{T}}, \qquad E = -B^{\mathrm{T}} \qquad \square$$

Следствие. Размерность симплектической группы $\mathbb{SP}(n,\mathbb{R})$ равна $n(2n+1)$. $\square$

Выше было отмечено, что $\det A = \pm 1$. Справедливо более сильное утверждение.

Предложение 18.1.2. *Определитель любой симплектической матрицы равен единице*

$$\det A = 1, \qquad A\in\mathbb{SP}(n,\mathbb{R}). \tag{18.5}$$

Доказательство. Любую симплектическую матрицу $A = (A_\alpha{}^\beta)\in\mathbb{SP}(n,\mathbb{R})$, $\alpha,\beta = 1,\ldots,2n$, можно рассматривать, как невырожденное линейное преобразование $2n$-мерного евклидова пространства $\mathbb{R}^{2n}$, которое оставляет каноническую 2-форму ϖ инвариантной. Следовательно, симплектическое преобразование сохраняет инвариантной и любую внешнюю степень формы ϖ. В частности, симплектическое преобразование сохраняет $2n$-форму объема ϖ^n. Известно, что при преобразовании координат форма объема умножается на якобиан преобразования координат. Таким образом для симплектических преобразований имеем

$$\varpi^n = \varpi^n \det A.$$

Отсюда следует равенство (18.5). $\square$

Симплектические группы $\mathbb{SP}(n,\mathbb{R})$ существенно отличаются от групп вращений евклидова пространства. В частности, они являются некомпактными. Продемонстрируем это на примере простейшей группы $\mathbb{SP}(1,\mathbb{R})$.

Теорема 18.1.1. *Группа* $\mathbb{SP}(1,\mathbb{R})$ *изоморфна группе* 2×2 *матриц* $\mathbb{SL}(2,\mathbb{R})$. *С топологической точки зрения эта группа трехмерна некомпактна и гомеоморфна прямому произведению окружности на двумерную плоскость,* $\mathbb{SP}(1,\mathbb{R})\approx\mathbb{S}^1\times\mathbb{R}^2$. *Она неодносвязна, и ее фундаментальная группа изоморфна группе целых чисел*

$$\pi\Big(\mathbb{SP}(1,\mathbb{R})\Big)\approx\mathbb{Z}.$$

Доказательство. В двумерном случае каноническая симплектическая форма с точностью до знака совпадает с полностью антисимметричным тензором $\varpi_{\alpha\beta}=-\varepsilon_{\alpha\beta}$ (28.64). Рассмотрим матрицу $A\in\mathbb{SP}(1,\mathbb{R})$, как линейное преобразование двумерной евклидовой плоскости $\mathbb{R}^2$. Тогда уравнение (18.2) запишется в виде

$$A_\alpha{}^\gamma A_\beta{}^\delta\varepsilon_{\gamma\delta}=\varepsilon_{\alpha\beta}.$$

Поскольку левая часть уравнения антисимметрична по индексам α и β, то она равна $\det A\,\varepsilon_{\alpha\beta}$. Отсюда следует, что уравнение (18.2) эквивалентно одному уравнению $\det A=1$. Таким образом группа $\mathbb{SP}(1,\mathbb{R})$ изоморфна группе вещественных 2×2 матриц с единичным определителем, т.е. группе $\mathbb{SL}(2,\mathbb{R})$.

Из курса линейной алгебры известно, что любое линейное преобразование плоскости, для которого $\det A=1$, можно однозначно представить в виде композиции двух преобразований: ортогонального поворота плоскости (группа $\mathbb{U}(1)\approx\mathbb{S}^1$) и преобразования, задающегося верхнетреугольной матрицей вида

$$\begin{pmatrix} a & b\\ 0 & 1/a\end{pmatrix},\qquad a>0.$$

Вещественные числа a и b можно рассматривать в качестве координат на полуплоскости $a>0$, которая гомеоморфна всей евклидовой плоскости. Таким образом мы получаем топологическое разложение группы $\mathbb{SP}(1,\mathbb{R})$ на прямое произведение окружности и двумерной плоскости. Прямое произведение $\mathbb{S}^1\times\mathbb{R}^2$ гомотопически эквивалентно окружности (стягивается к окружности), и поэтому фундаментальная группа $\mathbb{SP}(1,\mathbb{R})$ изоморфна фундаментальной группой окружности (теорема 11.3.4). □

Симплектические группы $\mathbb{SP}(n,\mathbb{R})$ при $n>1$ имеют более сложную структуру, и их описание выходит за рамки настоящей монографии. Отметим лишь, что все группы $\mathbb{SP}(n,\mathbb{R})$ некомпактны. В дальнейшем нам понадобятся другой класс групп Ли, который обозначается $\mathbb{SP}(n)$ и состоит из компактных симплектических групп. Они существуют и строятся, как подгруппы в комплексных симплектических группах $\mathbb{SP}(n)\subset\mathbb{SP}(n,\mathbb{C})$.

Рассмотрим $2n$-мерное комплексное пространство $\mathbb{C}^{2n}$ с координатами z^α, $\alpha=1,\dots,2n$. Каноническая симплектическая форма ϖ задает на $\mathbb{C}^{2n}$ симплектическую структуру (билинейную квадратичную форму), т.е. двум векторам X и Y ставится в соответствие число

$$\varpi:\quad \mathbb{C}^{2n}\times\mathbb{C}^{2n}\ni X,Y\mapsto\varpi(X,Y):=X^\alpha Y^\beta\varpi_{\alpha\beta}\in\mathbb{C}.$$

По-определению, симплектическая структура антисимметрична $\varpi(X,Y)=-\varpi(Y,X)$. Невырожденное комплексное линейное преобразование координат называется *симплектическим*, если оно сохраняет каноническую симплектическую структуру. Это преобразование задается комплексной $2n\times 2n$ матрицей $A\in\mathbb{SP}(2n,\mathbb{C})$, которая удовлетворяет тому же равенству (18.2), что и в вещественном случае. Единственное

отличие состоит в том, что элементами матрицы A теперь являются комплексные числа. Эти матрицы образуют *комплексную симплектическую группу* $\mathbb{SP}(n,\mathbb{C})$. Ясно, что группа $\mathbb{SP}(n,\mathbb{R})$ содержится в группе $\mathbb{SP}(n,\mathbb{C})$, как подгруппа вещественных симплектических преобразований.

Представление элементов алгебры Ли $\mathfrak{sp}(n,\mathbb{C})$ в виде (18.3) справедливо и для комплексных симплектических групп, только матрицы B, C и D будут комплексными. Поэтому размерность группы $\mathbb{SP}(n,\mathbb{C})$ в два раза больше размерности вещественной группы $\mathbb{SP}(n,\mathbb{R})$ и равна $2n(2n+1)$. Так же, как и в вещественном случае, группа $\mathbb{SP}(n,\mathbb{C})$ при всех n является некомпактной.

Предложение 18.1.3. *Симплектические группы $\mathbb{SP}(n,\mathbb{R})$ и $\mathbb{SP}(n,\mathbb{C})$ являются связными, т.е. состоят из одной компоненты.*

Предложение 18.1.4. *Характеристический полином*

$$f(\lambda) = \det(A - \lambda\mathbb{1}) = \sum_{k=1}^{2n} a_k \lambda^k$$

симплектического вещественного преобразования $A \in \mathbb{SP}(n,\mathbb{R})$ обладает свойством

$$f(\lambda) = \lambda^{2n} f(1/\lambda),$$

что означает симметричность его коэффициентов $a_k = a_{2n-k}$. В частности, если λ собственное число симплектического преобразования, то $1/\lambda$ также собственное число.

Доказательство. Из определения симплектического преобразования (18.2) следует, что $A = -\varpi A^{-1\mathrm{T}}\varpi$, т.к. $\varpi^2 = -\mathbb{1}$. Отсюда вытекает цепочка равенств

$$\begin{aligned} f(\lambda) &= \det(-\varpi A^{-1\mathrm{T}}\varpi - \lambda\mathbb{1}) = \det(-A^{-1\mathrm{T}} + \lambda\mathbb{1}) \\ &= \det(-\mathbb{1} + \lambda A) = \lambda^{2n} \det\left(A - \frac{1}{\lambda}\mathbb{1}\right), \end{aligned}$$

где мы воспользовались равенством $\det A = \det A^{\mathrm{T}} = 1$. □

Отметим, что у характеристического полинома не может быть нулевого собственного значения, т.к. $\det A = 1$. Поскольку характеристический полином является вещественным, то, если λ – комплексное собственное число, то $\bar{\lambda}$ – также собственное число. Таким образом, в случае общего положения собственные числа вещественного симплектического преобразования разбиваются на четверки $\lambda, \bar{\lambda}, 1/\lambda, 1/\bar{\lambda}$, т.е. собственные числа расположены на комплексной плоскости симметрично относительно вещественной оси и единичной окружности.

По-построению, комплексная симплектическая группа $\mathbb{SP}(n,\mathbb{C})$ содержит некомпактную вещественную подгруппу $\mathbb{SP}(n,\mathbb{R})$. Оказывается, что группа $\mathbb{SP}(n,\mathbb{C})$ содержит компактную подгруппу, которая называется компактной симплектической группой и обозначается через $\mathbb{SP}(n)$. Эту подгруппу удобно определить с использованием алгебры кватернионов $\mathbb{H}$ (см. приложение 28.3).

Определение. Рассмотрим n-мерное кватернионное пространство $\mathbb{H}^n$ с базисом e_{A}, $\mathrm{A} = 1, \ldots, n$. Каждый вектор $q \in \mathbb{H}^n$ однозначно представим в виде $q = q^{\mathrm{A}} e_{\mathrm{A}}$, где каждая координата является кватернионом $q^{\mathrm{A}} \in \mathbb{H}$. Каждый кватернион разлагается по базису $\{1, i, j, k\}$:

$$q^{\mathrm{A}} = a^{\mathrm{A}} + b^{\mathrm{A}} i + c^{\mathrm{A}} j + d^{\mathrm{A}} k.$$

Вещественная размерность $\mathbb{H}^n$ равна $4n$. Очевидно, что $\mathbb{H}^1 = \mathbb{H}$.

Рассмотрим в $\mathbb{H}^n$ симметричное вещественнозначное скалярное произведение

$$(q_1, q_2) := \mathsf{re}\, q_1\bar{q}_2 = \sum_{\mathrm{A}=1}^{n}(a_1^{\mathrm{A}}a_2^{\mathrm{A}} + b_1^{\mathrm{A}}b_2^{\mathrm{A}} + c_1^{\mathrm{A}}c_2^{\mathrm{A}} + d_1^{\mathrm{A}}d_2^{\mathrm{A}}). \tag{18.6}$$

Множество всех линейных кватернионных преобразований $A \in \mathbb{GL}(n, \mathbb{H})$ пространства $\mathbb{H}^n$, не меняющих начало координат и сохраняющих скалярное произведение

$$(Aq_1, Aq_2) = (q_1, q_2),$$

называется *симплектической компактной группой* $\mathbb{SP}(n)$. □

Кватернионное пространство $\mathbb{H}^n$ естественным образом отождествляется с евклидовым пространством $\mathbb{R}^{4n}$. Тогда скалярное произведение (18.6) совпадает с обычным евклидовым скалярным произведением в $\mathbb{R}^{4n}$. Это значит, что группа $\mathbb{SP}(n)$ является подгруппой в $\mathbb{O}(4n, \mathbb{R})$. Поскольку группа вращений $\mathbb{O}(4n, \mathbb{R})$ компактна, то и симплектическая группа $\mathbb{SP}(n)$ также компактна.

В приложении 28.3 показано, что алгебра кватернионов $\mathbb{H}$ естественным образом отождествляется с двумерным комплексным пространством $\mathbb{C}^2$. Выполняя эту операцию вдоль каждой из кватернионных координат, мы получим отождествление $\mathbb{H}^n$ с $\mathbb{C}^{2n}$. При отождествлении

$$q = z_1 + j\bar{z}_2, \qquad z_1 = a + ib, \quad z_2 = c + id,$$

квадратичная форма двух кватернионов $q_1 = z_{11} + jz_{12}$ и $q_2 = z_{21} + jz_{22}$ перейдет в сумму комплексных квадратичных форм

$$(q_1, q_2)_{\mathbb{H}} := q_1\bar{q}_2 = (q_1, q_2)_{\mathbb{C}} + \langle q_1, q_2\rangle_{\mathbb{C}},$$

где

$$(q_1, q_2)_{\mathbb{C}} = \sum_{\mathrm{A}}(z_{11}^{\mathrm{A}}\bar{z}_{21}^{\mathrm{A}} + z_{12}^{\mathrm{A}}\bar{z}_{22}^{\mathrm{A}}), \qquad \langle q_1, q_2\rangle_{\mathbb{C}} = \sum_{\mathrm{A}}(z_{12}^{\mathrm{A}}z_{21}^{\mathrm{A}} - z_{11}^{\mathrm{A}}z_{22}^{\mathrm{A}}).$$

Квадратичная форма $(q_1, q_2)_{\mathbb{C}}$ эрмитова, т.е. $\overline{(q_1, q_2)}_{\mathbb{C}} = (q_2, q_1)_{\mathbb{C}}$, а форма $\langle q_1, q_2\rangle_{\mathbb{C}}$ антисимметрична: $\langle q_1, q_2\rangle_{\mathbb{C}} = -\langle q_1, q_2\rangle_{\mathbb{C}}$. Отметим, что квадратичная форма $(q_1, q_2)_{\mathbb{H}}$ отличается от квадратичной формы (18.6) отсутствием знака реальной части.

Теорема 18.1.2. *Множество элементов $\mathbb{SP}(n)$ является связной компактной группой Ли вещественной размерности $n(2n+1)$. При отождествлении $\mathbb{H}^n$ с $\mathbb{C}^{2n}$ группа $\mathbb{SP}(n)$ вкладывается как подгруппа в унитарную группу $\mathbb{U}(2n)$. При этом вложении алгебра Ли $\mathfrak{sp}(n)$ группы $\mathbb{SP}(n)$ состоит из комплексных $2n \times 2n$ матриц вида*

$$\begin{pmatrix} Z_1 & Z_2 \\ -\bar{Z}_2 & \bar{Z}_1, \end{pmatrix} \tag{18.7}$$

где Z_1 – комплексная антиэрмитова $n \times n$ матрица, а Z_2 – комплексная симметричная $n \times n$ матрица. Если матрицы из унитарной группы $\mathbb{U}(2n)$ представить в виде

$$\begin{pmatrix} U_1 & U_2 \\ U_3 & U_4, \end{pmatrix}$$

где U_1, U_2, U_3 и U_4 – комплексные $n \times n$ матрицы, то подгруппа $\mathbb{SP}(n)$ в $\mathbb{U}(2n)$ состоит из унитарных матриц вида

$$\begin{pmatrix} U_1 & U_2 \\ -\overline{U}_2 & \overline{U}_1. \end{pmatrix} \tag{18.8}$$

При этом матрица (18.8) является унитарной тогда и только тогда, когда комплексные матрицы U_1 и U_2 удовлетворяют уравнениям $U_1U_1^\dagger + U_2U_2^\dagger = \mathbb{1}$ и $U_2U_1^{\mathrm{T}} = U_1U_2^{\mathrm{T}}$.

Доказательство. Проводится прямой проверкой [**?**]. □

Связь симплектических групп с другими матричными группами дается следующими двумя теоремами, доказательство которых дано, например, в [**?**].

Теорема 18.1.3. *Рассмотрим стандартные вложения групп $\mathbb{O}(n) \hookrightarrow \mathbb{U}(n)$, $\mathbb{U}(n) \hookrightarrow \mathbb{SO}(2n)$, $\mathbb{SP}(n) \hookrightarrow \mathbb{U}(2n)$. Тогда имеют место следующие соотношения:*

1) $\mathbb{SO}(2n) \cap \mathbb{SP}(n) = \mathbb{U}(n)$; здесь группы $\mathbb{SO}(2n)$ и $\mathbb{SP}(n)$ рассматриваются, как подгруппы в одной группе $\mathbb{U}(2n)$;

2) $\mathbb{SP}(n, \mathbb{C}) \cap \mathbb{U}(2n) = \mathbb{SP}(n)$;

3) $\mathbb{SP}(n, \mathbb{R}) \cap \mathbb{GL}(2n, \mathbb{C}) = \mathbb{U}(n)$;

4) $\mathbb{SP}(n, \mathbb{R}) \cap \mathbb{U}(2n) = \mathbb{U}(n)$.

Теорема 18.1.4. *Для компактных симплектических групп $\mathbb{SP}(1)$ и $\mathbb{SP}(2)$ имеют место изоморфизмы:*

$$\mathbb{SP}(1) \simeq \mathbb{SU}(2) \simeq \mathbb{SPIN}(3), \qquad \mathbb{SP}(2) \simeq \mathbb{SPIN}(5).$$

При больших $n > 2$ компактные симплектические группы $\mathbb{SP}(n)$ уже не сводятся к унитарным и ортогональным группам.

Теорема 18.1.5. *Группа $\mathbb{SP}(n)$ является максимальной компактной подгруппой в комплексной симплектической группе $\mathbb{SP}(n, \mathbb{C})$.*

Доказательство. См., например, []. □

Теорема 18.1.6. *Все компактные симплектические группы $\mathbb{SP}(n)$ односвязны.*

Доказательство. См., например, [**?**]. □

18.2 Симплектические многообразия

Определение. Многообразие $\mathbb{M}$ четной размерности, $\dim \mathbb{M} = 2n$, называется *симплектическим*, если на нем задана 2-форма

$$\omega = \frac{1}{2} dx^\alpha \wedge dx^\beta \omega_{\alpha\beta} \in \Lambda_2(\mathbb{M}),$$

удовлетворяющая двум условиям:

1) $\det \omega_{\alpha\beta} \neq 0, \quad \forall x \in \mathbb{M}$ – невырожденность,

2) $d\omega = 0$ – замкнутость.

Форма ω называется *симплектической*. □

Пусть на многообразии $\mathbb{M}$ задано два векторных поля $X, Y \in \mathcal{X}(\mathbb{M})$. Тогда форма ω называется невырожденной, если из условия $\omega(X,Y) = 0$ для всех $Y \in \mathcal{X}(\mathbb{M})$ следует $X = 0$. Легко проверить, что данное инвариантное определение невырожденности 2-формы эквивалентно условию $\det \omega_{\alpha\beta} \neq 0$, которое должно выполняться во всех картах и для всех точек $x \in \mathbb{M}$.

По-определению, компоненты симплектической формы антисимметричны относительно перестановки индексов. Это значит, что ее невырожденность возможна только на многообразиях четной размерности. Поэтому размерность многообразия включена в определение.

В координатах замкнутость симплектической формы ω записывается в виде дифференциального уравнения

$$\partial_\alpha \omega_{\beta\gamma} + \partial_\beta \omega_{\gamma\alpha} + \partial_\gamma \omega_{\alpha\beta} = 0. \tag{18.9}$$

Рассмотрим связь симплектических форм с ориентацией многообразий и формами объема.

Теорема 18.2.1. *Пусть ω – замкнутая 2-форма на многообразии $\mathbb{M}$, $\dim \mathbb{M} = 2n$. Для того, чтобы эта форма была симплектической необходимо и достаточно, чтобы $2n$-форма ω^n была формой объема на $\mathbb{M}$, т.е. нигде не обращалась в нуль.*

Доказательство. Следствие предложения 4.4.3. □

Следствие. Любое симплектическое многообразие $(\mathbb{M}, \omega)$ ориентируемо. □

Доказательство. Отличие от нуля формы ω^n согласно теореме 11.4.3 достаточно для ориентируемости. □

Обычно симплектическое многообразие $(\mathbb{M}, \omega)$ ориентируют формой объема υ с дополнительным множителем:

$$\upsilon = \frac{1}{n!}(-1)^{\frac{n(n+1)}{2}} \omega^n, \tag{18.10}$$

чтобы согласовать выбор канонических координат (см. ниже) с канонической ориентацией евклидова пространства.

Пример 18.2.1. Рассмотрим каноническую 2-форму (18.1) на евклидовом пространстве $\mathbb{R}^{2n}$

$$\varpi = \frac{1}{2} dx^\alpha \wedge dx^\beta \varpi_{\alpha\beta} = dx^{n+1} \wedge dx^1 + dx^{n+2} \wedge dx^2 + \ldots + dx^{2n} \wedge dx^n. \tag{18.11}$$

Она невырождена, т.к. $\det \varpi_{\alpha\beta} = 1$, и замкнута, поскольку компоненты формы постоянны. Тем самым форма ϖ определяет симплектическую форму на евклидовом пространстве. Поскольку

$$\varpi^n = (-1)^{\frac{n(n+1)}{2}} n! dx^1 \wedge \ldots \wedge dx^{2n},$$

то соответствующая форма объема

$$\upsilon = \frac{1}{n!}(-1)^{\frac{n(n+1)}{2}} \varpi^n = dx^1 \wedge \ldots \wedge dx^{2n} \tag{18.12}$$

является канонической формой объема евклидова пространства (4.73). □

Обратное утверждение о связи между наличием формы объема и симплектической формы неверно. Для двумерных многообразий это так: всякая ориентируемая поверхность допускает симплектическую структуру. Это просто полностью антисимметричный тензор второго ранга

$$\omega = \upsilon = \frac{1}{2} dx^\alpha \wedge dx^\beta \varepsilon_{\alpha\beta} = dx^1 \wedge dx^2 \sqrt{|g|}.$$

Для ориентируемых многообразий более высокой размерности $\dim \mathbb{M} = 2n$, $n > 1$, симплектическая форма существует не всегда.

Следующая конструкция позволяет строить симплектические структуры на произвольном кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$, $\dim \mathbb{M} = n$, которое имеет размерность $2n$.

Определение. Пусть $X_r \in \mathbb{T}_r\Big(\mathbb{T}^*(\mathbb{M})\Big)$ – произвольный касательный вектор к кокасательному расслоению в точке $r \in \mathbb{T}^*(\mathbb{M})$. Определим *каноническую линейную форму* Θ на $\mathbb{T}^*(\mathbb{M})$, которая называется *формой Лиувилля* или *относительным интегральным инвариантом Пуанкаре*, следующим соотношением

$$\Theta(X_r) = r(\pi_* X_r), \tag{18.13}$$

где π_* – дифференциал проекции кокасательного расслоения, $\pi : \mathbb{T}^*(\mathbb{M}) \to \mathbb{M}$. □

Теорема 18.2.2. *На кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$ к произвольному многообразию $\mathbb{M}$ существует симплектическая структура.*

Доказательство. Заметим, что r является линейной формой на $\mathbb{T}_{\pi(r)}(\mathbb{M})$ и $\pi_* X_r$ – касательный вектор в точке $\pi(r)$. Следовательно, форма Лиувилля (18.13) определена всюду на кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$. Пусть q^α – локальная система координат на $\mathbb{U} \subset \mathbb{M}$. Введем на $\pi^{-1}(\mathbb{U})$ координаты $(x^1, \ldots, x^{2n}) = (q^1, \ldots, q^n, p_1, \ldots, p_n)$ следующим образом. Пусть q^α, $\alpha = 1, \ldots, n$ – координаты на $\mathbb{U}$. Формы $dx^1, \ldots, dx^n$ порождают пространство 1-форм на $\mathbb{U}$. Поэтому для всякой 1-формы $r \in \pi^{-1}(\mathbb{U})$ имеем

$$r = dq^\alpha p_\alpha.$$

Разложим касательный вектор X_r по координатному базису (q, p):

$$X_r = X_r^\alpha \frac{\partial}{\partial q^\alpha} + X_{r\alpha} \frac{\partial}{\partial p_\alpha}.$$

Тогда

$$\pi_* X_r = X_r^\alpha \frac{\partial}{\partial x^\alpha},$$

и условие (18.13) принимает вид

$$\Theta(X_r) = X_r^\alpha p_\alpha.$$

Таким образом получаем выражение для формы Лиувилля в локальных координатах:

$$\Theta = dx^\alpha p_\alpha. \tag{18.14}$$

Внешняя производная от формы Лиувилля имеет вид

$$d\Theta = dp_\alpha \wedge dq^\alpha = \frac{1}{2} dx^{\mathrm{A}} \wedge dx^{\mathrm{B}} \varpi_{\mathrm{AB}}, \qquad \mathrm{A}, \mathrm{B} = 1, \ldots, 2n,$$

т.е. совпадает с канонической симплектической формой ϖ. Таким образом построена невырожденная замкнутая 2-форма $d\Theta$ на произвольном кокасательном расслоении. □

Форма Лиувилля естественна в том смысле, что для любого сечения $\sigma : \mathbb{M} \to \mathbb{T}^*(\mathbb{M})$ возврат отображения σ^* переводит 1-форму Θ в r: $\sigma^*(\Theta) = r$.

Следствие. Любое кокасательное расслоение $\mathbb{T}^*(\mathbb{M})$ ориентируемо. □

Замечание. В гамильтоновом подходе к описанию динамики точечных частиц (см. главу 20) многообразие $\mathbb{M}$ и его кокасательное расслоение $\mathbb{T}^*(\mathbb{M})$ отождествляются, соответственно, с конфигурационным и фазовым пространствами. □

Поскольку ранг симплектической 2-формы максимален и равен $2n$, то ее класс также равен $2n$. В этом случае теорема Дарбу (4.4.4) формулируется следующим образом.

Теорема 18.2.3 (Дарбу). *Пусть задано симплектическое многообразие $(\mathbb{M}, \omega)$. Тогда у каждой точки $x \in \mathbb{M}$ существует такая координатная окрестность $\mathbb{U}_x \subset \mathbb{M}$, в которой симплектическая форма принимает канонический вид*

$$\omega = \varpi.$$

Определение. Координаты, в которых симплектическая форма имеет канонический вид, называются *координатами Дарбу.* □

Замечание. В главе 5 мы видели, что метрику g на многообразии $\mathbb{M}$ за счет выбора системы координат можно привести к каноническому виду только в фиксированной точке $x \in \mathbb{M}$. Лучшее, что можно сделать в общем случае, это привести метрику к каноническому виду вдоль произвольной кривой $\gamma \in \mathbb{M}$. В этом отношении симплектические многообразия проще. Согласно теореме Дарбу симплектическую форму можно привести к каноническому виду не только в данной точке многообразия, но и в некоторой окрестности этой точки. □

Поскольку симплектическая структура $\omega_{\alpha\beta}(x)$ является невырожденной, то в каждой точке $x \in \mathbb{M}$ существует обратная матрица $\omega^{-1\,\alpha\beta}(x)$, которая также антисимметрична:

$$\omega^{-1\,\alpha\beta}\omega_{\beta\gamma} = \omega_{\gamma\beta}\omega^{-1\,\beta\alpha} = \delta^\alpha_\gamma, \qquad \omega^{-1\,\alpha\beta} = -\omega^{-1\,\beta\alpha}.$$

Замечание. Если на симплектическом многообразии $(\mathbb{M}, \omega)$ задана также метрика g, то в общем случае

$$\omega^{-1\,\alpha\beta} \neq \omega^{\alpha\beta} := g^{\alpha\gamma}g^{\beta\delta}\omega_{\gamma\delta}.$$ □

Пример 18.2.2. Для канонической симплектической структуры (18.1)

$$\varpi^{-1} = (\varpi^{-1\alpha\beta}) = -\varpi = \varpi^{\mathrm{T}} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}.$$ □

Симплектическая структура устанавливает взаимно однозначное соответствие сечений касательных и кокасательных расслоений. В компонентах данное соответствие задается простым правилом:

$$A_\alpha = X^\beta\omega_{\beta\alpha}, \quad X^\alpha = A_\beta\omega^{-1\,\beta\alpha}, \qquad X \in \mathcal{X}(\mathbb{M}), \quad A \in \Lambda_1(\mathbb{M}). \tag{18.15}$$

Замкнутость симплектической формы (18.9) для обратного тензора $\omega^{-1\,\alpha\beta}(x)$ можно переписать в виде

$$\omega^{-1\,\alpha\delta}\partial_\delta\omega^{-1\,\beta\gamma} + \omega^{-1\,\beta\delta}\partial_\delta\omega^{-1\,\gamma\alpha} + \omega^{-1\,\gamma\delta}\partial_\delta\omega^{-1\,\alpha\beta} = 0.$$

Для этого достаточно свернуть равенство (18.9) с обратными матрицами $\omega^{-1\,\alpha\beta}$.

Используя внутреннее умножение (4.23), формулы (18.15) можно переписать в инвариантном виде:

$$A = \mathsf{i}_{X_A}\omega = X_A\rfloor\omega,$$

где X_A – векторное поле, соответствующее 1-форме A.

Поскольку симплектическая структура на $\mathbb{M}$ устанавливает взаимно однозначное соответствие между множеством векторных полей $\mathcal{X}(\mathbb{M})$ и пространством 1-форм $\Lambda_1(\mathbb{M})$, то структуру алгебры Ли на $\mathcal{X}(\mathbb{M})$ можно перенести на $\Lambda_1(\mathbb{M})$. Для двух произвольных 1-форм $A, B \in \Lambda_1(\mathbb{M})$ положим

$$[A, B] = [X_A, X_B]\rfloor\omega. \tag{18.16}$$

Тем самым линейное пространство $\Lambda_1(\mathbb{M})$ приобретает структуру алгебры Ли. Таким образом коммутатору векторных полей ставится в соответствие *скобка Пуассона* 1-форм (18.16).

В компонентах $A = dx^\alpha A_\alpha$, $B = dx^\alpha B_\beta$ и

$$[A, B] = dx^\alpha \left[(A_\beta\partial_\gamma B_\alpha + B_\gamma\partial_\beta A_\alpha)\omega^{-1\beta\gamma} + A_\beta B_\gamma\partial_\alpha\omega^{-1\beta\gamma}\right]. \tag{18.17}$$

Эта формула получается прямыми вычислениями.

Скобку Пуассона двух 1-форм можно выразить через производную Ли:

$$\begin{aligned}\mathsf{L}_{X_A}B &= \mathsf{L}_{X_A}(X_B\rfloor\omega) = \\ &= \mathsf{L}_{X_A}X_B\rfloor\omega + X_B\rfloor(\mathsf{L}_{X_A}\omega) = \\ &= [X_A, X_B]\rfloor\omega + X_B\rfloor\,(X_A\rfloor d\omega + d(X_A\rfloor\omega)) = \\ &= [A, B] + X_B\rfloor dA,\end{aligned}$$

где мы воспользовались линейностью производной Ли, основной формулой гомотопии (4.42) и замкнутостью симплектической формы. Отсюда следует выражение для скобки Пуассона через производную Ли:

$$[A, B] = L_{X_A}B - X_B\rfloor dA. \tag{18.18}$$

Если 1-формы A и B замкнуты, то полученное выражение упрощается:

$$[A, B] = L_{X_A}B = -\mathsf{L}_{X_B}A.$$

Предложение 18.2.1. *Скобка Пуассона двух замкнутых форм является точной формой.*

Доказательство. Из основной формулы гомотопии (4.42) следует, что для любой замкнутой формы A справедливо равенство $\mathsf{L}_X A = d(X\rfloor A)$. □

В компонентах скобка Пуассона двух замкнутых 1-форм (18.17) принимает простой вид

$$[A, B] = dx^\alpha\partial_\alpha(\omega^{-1\beta\gamma}A_\beta B_\gamma).$$

Скобку Пуассона можно также определить в алгебре функций на $\mathbb{M}$. А именно, каждая симплектическая структура определяет скобку Пуассона двух дифференцируемых функций $f(x)$ и $g(x)$ класса $\mathcal{C}^2$:

$$[f, g] = \omega^{-1\,\alpha\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial g}{\partial x^\beta}. \tag{18.19}$$

Дважды непрерывная дифференцируемость функций необходима для выполнения тождеств Якоби.

Замечание. Поля $\omega^{-1\,\alpha\beta}(x)$ являются компонентами антисимметричного контравариантного тензора второго ранга. Они являются частным случаем пуассоновой структуры, рассмотренной в следующем разделе. □

Замечание. Суперсимметричные модели часто формулируются в суперпространстве $\mathbb{N}$. Не вдаваясь в детали определения супепространства, отметим следующее. Пусть на $\mathbb{N}$ задана локальная система координат x^i, θ^α $(i = 1, \ldots, m,\ \alpha = 1, \ldots, 2n)$, где x^i – обычные коммутирующие координаты на подмногообразии $\mathbb{M} \subset \mathbb{N}$, а координаты θ^α, являются антикоммутирующими: $\theta^\alpha\theta^\beta = -\theta^\beta\theta^\alpha$. Тогда симплектическая форма $\omega \in \Lambda_2(\mathbb{M})$ играет роль метрики для антикоммутирующих переменных. Поэтому симплектические группы и многообразия играют важную роль в суперсимметричных моделях. □

Пусть на симплектическом многообразии $(\mathbb{M}, \omega)$ задана также аффинная геометрия, т.е. метрика g и аффинная связность Γ. Назовем аффинную связность *согласованной с симплектической структурой*, если ковариантная производная симплектической формы равна нулю:

$$\nabla_\alpha \omega_{\beta\gamma} = \partial_\alpha \omega_{\beta\gamma} - \Gamma_{\alpha\beta}{}^\delta \omega_{\delta\gamma} - \Gamma_{\alpha\gamma}{}^\delta \omega_{\beta\delta} = 0. \tag{18.20}$$

Рассмотрим линейную комбинацию ковариантных производных, в которой слагаемые отличаются циклической перестановкой индексов

$$\nabla_\alpha \omega_{\beta\gamma} + \nabla_\beta \omega_{\gamma\alpha} + \nabla_\gamma \omega_{\alpha\beta}.$$

Слагаемые с производными от симплектической формы сокращаются ввиду замкнутости ω, и мы получаем уравнение на тензор кручения

$$T_{\alpha\beta}{}^\delta \omega_{\delta\gamma} + T_{\beta\gamma}{}^\delta \omega_{\delta\alpha} + T_{\gamma\alpha}{}^\delta \omega_{\delta\beta} = 0. \tag{18.21}$$

Это необходимое условие для того, чтобы аффинная связность была согласована с симплектической структурой, но не достаточное.

18.3 Пуассоновы многообразия

Рассмотрим множество скалярных полей $f, g, \ldots \in \mathcal{C}^\infty(\mathbb{M})$ (алгебру гладких функций) на многообразии $\mathbb{M}$ произвольной размерности n.

Определение. Билинейное отображение

$$J: \quad \mathcal{C}^\infty(\mathbb{M}) \times \mathcal{C}^\infty(\mathbb{M}) \ni f, g \mapsto [f, g] \in \mathcal{C}^\infty(\mathbb{M})$$

называется *пуассоновой структурой* или *скобкой Пуассона*, если выполнены следующие четыре условия:

1) $[af + bg, h] = a[f, h] + b[g, h], \quad a, b \in \mathbb{R}$ – линейность,
2) $[f, g] = -[g, f]$ – антисимметрия,
3) $[f, gh] = [f, g]h + g[f, h]$ – правило Лейбница,
4) $[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0$ – тождество Якоби.

Многообразие с заданной скобкой Пуассона называется *пуассоновым*. □

Из линейности скобки Пуассона и правила Лейбница следует, что скобка Пуассона постоянной функции $f = C = \mathsf{const}$ с произвольной функцией равна нулю:

$$[C, g] = 0.$$

Используя это свойство, получим явное выражение для скобки Пуассона в локальной системе координат x^α. С этой целью разложим функции f и g в ряды Тейлора в окрестности произвольной точки x_0:

$$f(x) = f_0 + (x^\alpha - x_0^\alpha)\left.\frac{\partial f}{\partial x^\alpha}\right|_{x=x_0} + \dots$$
$$g(x) = g_0 + (x^\alpha - x_0^\alpha)\left.\frac{\partial g}{\partial x^\alpha}\right|_{x=x_0} + \dots$$

Тогда

$$[f, g] = [x^\alpha - x_0^\alpha, x^\beta - x_0^\beta]\left.\frac{\partial f}{\partial x^\alpha}\frac{\partial g}{\partial x^\beta}\right|_{x=x_0} + \mathsf{o}(x - x_0) =$$
$$= [x^\alpha, x^\beta]\left.\frac{\partial f}{\partial x^\alpha}\frac{\partial g}{\partial x^\beta}\right|_{x=x_0} + \mathsf{o}(x - x_0).$$

Таким образом, в точке x_0 скобка определяется первым слагаемым. Поскольку точка x_0 произвольна, то получаем явное выражение для скобки Пуассона в координатах

$$[f, g] = J^{\alpha\beta}\frac{\partial f}{\partial x^\alpha}\frac{\partial g}{\partial x^\beta}, \tag{18.22}$$

где введены *структурные функции* пуассонова многообразия:

$$J^{\alpha\beta} = [x^\alpha, x^\beta] \tag{18.23}$$

Напомним, что каждую координату можно рассматривать, как функцию на многообразии $x^\alpha : \mathbb{M} \ni x \mapsto x^\alpha(x) \in \mathbb{R}$, поэтому для координатных функций скобка Пуассона (18.23) определена.

Скобка Пуассона (18.22) совпадает со скобкой Пуассона (18.19), введенной для симплектической структуры при $J^{\alpha\beta} = \omega^{-1\alpha\beta}$.

По-построению структурные функции $J^{\alpha\beta}(x)$ представляют собой компоненты антисимметричного тензора второго ранга, и выражение в правой части равенства (18.22) является функцией. Из определения скобки Пуассона следует, что структурные функции антисимметричны,

$$J^{\alpha\beta} = -J^{\beta\alpha},$$

и удовлетворяют тождеству Якоби:

$$J^{\alpha\delta}\partial_\delta J^{\beta\gamma} + J^{\beta\delta}\partial_\delta J^{\gamma\alpha} + J^{\gamma\delta}\partial_\delta J^{\alpha\beta} = 0, \tag{18.24}$$

где слагаемые отличаются циклической перестановкой индексов α, β, γ. Легко проверяется, что произвольный антисимметричный тензор второго ранга, удовлетворяющий тождеству Якоби, взаимно однозначно определяет скобку Пуассона (18.22).

Пример 18.3.1. Произвольная постоянная антисимметричная матрица определяет пуассонову структуру в некоторой системе координат. Действительно, тождества Якоби в этом случае выполняются. □

Пуассонова структура является билинейным дифференциальным оператором и записывается также в виде

$$J(f,g) := [f,g] = J^{\alpha\beta}\partial_\alpha f \partial_\beta g.$$

Поэтому для нее применяют иногда обозначение

$$J = \frac{1}{2} J^{\alpha\beta}\partial_\alpha \wedge \partial_\beta,$$

где знак $\wedge$ обозначает внешнее умножение (см. раздел 4).

Если на многообразии помимо пуассоновой структуры J задана также аффинная связность Γ, то тождество Якоби (18.24) можно переписать в явно ковариантном виде

$$\begin{aligned} &J^{\alpha\delta}\nabla_\delta J^{\beta\gamma} + J^{\beta\delta}\nabla_\delta J^{\gamma\alpha} + J^{\gamma\delta}\nabla_\delta J^{\alpha\beta} \\ &- J^{\alpha\delta}T_{\delta\epsilon}{}^{\beta} J^{\epsilon\gamma} - J^{\beta\delta}T_{\delta\epsilon}{}^{\gamma} J^{\epsilon\alpha} - J^{\gamma\delta}T_{\delta\epsilon}{}^{\alpha} J^{\epsilon\beta} = 0, \end{aligned}$$

где $T_{\alpha\beta}{}^{\gamma}$ – тензор кручения аффинной связности. Это доказывает, что тождества Якоби ковариантны, и их выполнение в одной системе координат влечет за собой их выполнение во всех других системах. Впрочем, это можно было бы и не доказывать, т.к. определение пуассоновой структуры было дано в инвариантном виде.

Если матрица структурных функций невырождена, $\det J^{\alpha\beta} \neq 0$, то она имеет обратную $J^{-1}_{\alpha\beta}$:

$$J^{\alpha\beta}J^{-1}_{\beta\gamma} = \delta^{\alpha}_{\gamma},$$

которая также антисимметрична. Это значит, что контравариантный тензор $J^{\alpha\beta}$ взаимно однозначно определяет 2-форму $dx^\alpha \wedge dx^\beta J^{-1}_{\alpha\beta}$ на многообразии. Из тождеств Якоби (18.24) следует, что 2-форма $J^{-1}_{\alpha\beta}$ замкнута,

$$\partial_\alpha J^{-1}_{\beta\gamma} + \partial_\beta J^{-1}_{\gamma\alpha} + \partial_\gamma J^{-1}_{\alpha\beta} = 0,$$

т.е. является симплектической (см. раздел 18.2). Легко проверяется и обратное утверждение: произвольная симплектическая форма $\omega_{\alpha\beta} = J^{-1}_{\alpha\beta}$ определяет пуассонову структуру на многообразии с невырожденными структурными функциями. При этом тождества Якоби следуют из условия замкнутости формы. Таким образом пуассоновы структуры с невырожденными структурными функциями находятся во взаимно однозначном соответствии с симплектическими формами на многообразии.

В общем случае структурные функции пуассоновой структуры могут быть вырождены. Это значит, в частности, что, в отличие от симплектической структуры, скобка Пуассона может быть также определена на многообразии нечетной размерности. *Рангом* пуассоновой структуры называется ранг матрицы $J^{\alpha\beta}$ структурных функций. Ввиду антисимметрии $J^{\alpha\beta}$ ранг пуассоновой структуры может быть только четным. В общем случае ранг пуассоновой структуры может меняться от точки к точке.

Если пуассонова структура вырождена, $\det J^{\alpha\beta} = 0$, то существуют *функции Казимира* $c \in \mathcal{C}^\infty(\mathbb{U})$, определенные, возможно, только в некоторой области $\mathbb{U} \subset \mathbb{M}$. Они определяются следующим равенством

$$[c,f] = 0, \quad \forall f \in \mathcal{C}^\infty(\mathbb{U}). \tag{18.25}$$

В координатах это равенство принимает вид

$$J^{\alpha\beta}\partial_\beta c = 0.$$

Очевидно, что любое решение этих уравнений определено с точностью до аддитивной постоянной. Если пуассонова структура невырождена, то константы $c = \mathsf{const}$ являются единственными функциями Казимира.

Пусть ранг матрицы $J^{\alpha\beta}$ постоянен на $\mathbb{M}$ и равен $\mathsf{rank}\, J^{\alpha\beta} = 2m < n$. Количество функционально независимых функций Казимира c^{A}, $\mathrm{A} = 1, \ldots, n-2m$, равно числу функционально независимых решений уравнения (18.25), т.е. разности между размерностью многообразия n и рангом пуассоновой структуры $2m$. Если все функции Казимира известны и определены на всем $\mathbb{M}$, то условия $c^{\mathrm{A}} = \mathsf{const}$ выделяет в $\mathbb{M}$ подмногообразие $\mathbb{U}$ размерности $2m$. В разделе 18.5 мы увидим, что сужение J на это подмногообразие приводит к невырожденной пуассоновой структуре на $\mathbb{U}$.

Из определения функций Казимира следует, в частности, что скобка Пуассона двух функций Казимира равна нулю:

$$[c^{\mathrm{A}}, c^{\mathrm{B}}] = J^{\alpha\beta}\partial_\alpha c^{\mathrm{A}}\partial_\beta c^{\mathrm{B}} = 0. \tag{18.26}$$

Пусть задано пуассоново многообразие $(\mathbb{M}, J)$. Наличие пуассоновой структуры на алгебре функций $\mathcal{C}^\infty(\mathbb{M})$ задает билинейное отображение, которое превращает множество функций в алгебру Ли. Эта алгебра Ли гомоморфно отображается в алгебру Ли векторных полей на $\mathbb{M}$ следующим образом. Скобка Пуассона координатных функций x^α с произвольной функцией f определяет компоненты векторного поля $X_f = X_f^\alpha \partial_\alpha$, где

$$X_f^\alpha = [f, x^\alpha] = -J^{\alpha\beta}\partial_\beta f.$$

Таким образом, мы имеем отображение множества гладких функций в алгебру Ли векторных полей

$$\mathcal{C}^\infty(\mathbb{M}) \ni f \mapsto X_f \in \mathcal{X}(\mathbb{M}). \tag{18.27}$$

Предложение 18.3.1. *Скобка Пуассона двух функций $f, g \in \mathcal{C}^\infty(\mathbb{M})$ связана с коммутатором соответствующих векторных полей соотношением*

$$[X_f, X_g] = X_{[f,g]}. \tag{18.28}$$

То есть отображение (18.27) является гомоморфизмом.

Доказательство. Простое следствие тождеств Якоби. □

Если пуассонова структура невырождена, то гомоморфизм (18.27) имеет нетривиальное ядро, состоящее из функций, постоянных на $\mathbb{M}$. Если пуассонова структура вырождена, то ядро отображения (18.28) включает также линейную оболочку функций Казимира.

Предложение 18.3.2. *Векторное поле X_f, порожденное произвольной функцией f, сохраняет скобку Пуассона. А именно, производная Ли вдоль векторного поля X_f от структурных функций равна нулю:*

$$\mathsf{L}_{X_f} J = 0.$$

Доказательство. Простая проверка. □

Обсудим связь пуассоновых структур с гамильтоновой динамикой точечных частиц. Рассмотрим произвольную функцию $H \in \mathcal{C}^\infty(\mathbb{M})$ и соответствующее векторное поле X_H. Тогда экспоненциальное отображение, определяемое системой уравнений

$$\dot{x}^\alpha = -X_H^\alpha = [x^\alpha, H] = J^{\alpha\beta}\partial_\beta H,$$

является ни чем иным, как уравнениями движения механической системы точечных частиц в канонической (гамильтоновой) форме. При этом точка обозначает дифференцирование по времени, и $H = H(x)$ – гамильтониан системы. Векторное поле

$$X_H = -J^{\alpha\beta}\partial_\beta H \partial_\alpha \tag{18.29}$$

называется *гамильтоновым*. Другими словами, каждую функцию на пуассоновом многообразии можно рассматривать, как гамильтониан некоторой механической системы. При этом каждому гамильтониану соответствует свое гамильтоново векторное поле.

Легко проверить, что для каждого гамильтонова векторного поля справедливо равенство

$$X_H f = [f, H]. \tag{18.30}$$

Пример 18.3.2. Механическая система, состоящая из N точечных частиц, в фазовом пространстве описывается координатами q^i и импульсами p_i, $i = 1, \cdots, \mathrm{N}$, со скобками Пуассона

$$[q^i, p_j] = \delta^i_j, \quad [q^i, q^j] = 0, \quad [p_i, p_j] = 0.$$

Фазовое пространство представляет собой 2N-мерное пуассоново многообразие. В координатах $\{x^\alpha\} = \{q^1, \ldots, q^{\mathrm{N}}, p_1, \ldots, p_{\mathrm{N}}\}$ структурные функции имеют вид

$$J^{\alpha\beta} = \varpi^{-1\alpha\beta} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \tag{18.31}$$

Эта пуассонова структура невырождена и называется *канонической*. Как матрица каноническая пуассонова структура с точностью до знака совпадает с канонической симплектической формой (18.1), однако ее индексы расположены сверху, что соответствует контравариантному тензору второго ранга.

Координаты q^i и сопряженные импульсы p_i являются системой локальных координат фазового пространства $\mathbb{T}^*(\mathbb{M})$, которое является кокасательным расслоением к конфигурационному пространству $\mathbb{M}$ с координатами q^i.

Канонические уравнения движения механической системы

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}$$

записываются в виде

$$\dot{x}^\alpha = [x^\alpha, H],$$

где $H = H(q, p)$ – гамильтониан системы. Гамильтоново векторное поле на фазовом пространстве $\mathbb{T}^*(\mathbb{M})$ имеет вид

$$X_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i}. \qquad \square$$

Замечание. С точки зрения динамики функции Казимира представляют собой первые интегралы движения механической системы, которые существуют для любых функций Гамильтона. В этом смысле они носят кинематический характер. $\square$

18.4 Структура Ли–Пуассона

Теперь рассмотрим важный класс пуассоновых структур, связанных с алгебрами Ли.

Определение. Пусть $\mathfrak{g}$ – алгебра Ли размерности N. Это линейное пространство, каждая точка которого $y = y^{\mathrm{A}} L_{\mathrm{A}} \in \mathfrak{g}$, $\mathrm{A} = 1, \ldots, \mathrm{N}$, разлагается по базису L_{A} с коммутационными соотношениями

$$[L_{\mathrm{A}}, L_{\mathrm{B}}] = f_{\mathrm{AB}}{}^{\mathrm{C}} L_{\mathrm{C}},$$

где $f_{\mathrm{AB}}{}^{\mathrm{C}}$ – структурные константы алгебры Ли $\mathfrak{g}$. Обозначим дуальное пространство к алгебре Ли через $\mathfrak{g}^*$ (множество линейных функционалов на $\mathfrak{g}$). Пусть ω^{A} – базис в $\mathfrak{g}^*$, который дуален к L_{A}, т.е. $\omega^{\mathrm{A}}(L_{\mathrm{B}}) = \delta^{\mathrm{A}}_{\mathrm{B}}$. Тогда точка дуального пространства представима в виде $x = \omega^{\mathrm{A}} x_{\mathrm{A}} \in \mathfrak{g}^*$. Значение функционала x на векторе y равно

$$x(y) = y^{\mathrm{A}} x_{\mathrm{A}}.$$

Рассмотрим функцию $f(x) \in \mathcal{C}^\infty(\mathfrak{g}^*)$ на дуальном пространстве $\mathfrak{g}^*$. Ее дифференциал равен

$$df = dx_{\mathrm{A}} \partial^{\mathrm{A}} f.$$

Мы пишем индекс у частной производной ∂^{A} вверху, т.к. координаты дуального пространства $\mathfrak{g}^*$ нумеруются нижним индексом. Каждый дифференциал естественным образом отождествляется с элементом алгебры Ли $\nabla f := \partial^{\mathrm{A}} f L_{\mathrm{A}} \in \mathfrak{g}$. Определим скобку Пуассона для функций на дуальном пространстве $f, g \in \mathcal{C}^\infty(\mathfrak{g}^*)$ в произвольной точке $x \in \mathfrak{g}^*$:

$$[f, g](x) := x\Big([\nabla f, \nabla g]\Big), \tag{18.32}$$

где $[\nabla f, \nabla g]$ – коммутатор в алгебре Ли $\mathfrak{g}$. В компонентах данная скобка Пуассона записывается в виде

$$[f, g] = J_{\mathrm{AB}} \partial^{\mathrm{A}} f \partial^{\mathrm{B}} g, \tag{18.33}$$

где структурные функции

$$J_{\mathrm{AB}} = f_{\mathrm{AB}}{}^{\mathrm{C}} x_{\mathrm{C}} \tag{18.34}$$

линейны по координатам x_{C}. Нетрудно проверить, что тождества Якоби для структурных функций (18.24) совпадают с тождествами Якоби для структурных констант алгебры Ли $\mathfrak{g}$:

$$f_{\mathrm{AB}}{}^{\mathrm{D}} f_{\mathrm{CD}}{}^{\mathrm{E}} + f_{\mathrm{BC}}{}^{\mathrm{D}} f_{\mathrm{AD}}{}^{\mathrm{E}} + f_{\mathrm{CA}}{}^{\mathrm{D}} f_{\mathrm{BD}}{}^{\mathrm{E}} = 0.$$

Скобка Пуассона (18.32) называется скобкой *Ли–Пуассона.* □

Таким образом, скобка Ли–Пуассона является частным случаем скобки Пуассона с линейными структурными функциями (18.34). Ранг этой скобки в нуле всегда равен нулю, $J_{\mathrm{AB}}|_{x=0} = 0$.

Пример 18.4.1. (Вращение твердого тела.) В трехмерном евклидовом пространстве с декартовыми координатами $x_i \in \mathbb{R}^3$, $i = 1, 2, 3$, которое мы отождествим с дуальной алгеброй Ли трехмерных вращений $\mathfrak{so}(3)^*$, скобка Ли–Пуассона равна

$$[x_i, x_j] = -\varepsilon_{ijk} x^k, \tag{18.35}$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга, а подъем и опускание индексов производится с помощью евклидовой метрики δ_{ij}. Эта структура вырождена, ее ранг равен двум всюду, кроме начала координат, где он равен нулю. Для пуассоновой структуры (18.35) существует единственная функция Казимира

$$c = x^i x_i.$$

Скобку Ли–Пуассона (18.35) можно сузить на сферу произвольного радиуса, соответствующую постоянному значению функции Казимира $c = \mathsf{const}$. Соответствующая пуассонова структура на сфере $x^i x_i = \mathsf{const} > 0$ невырождена. В качестве координат Дарбу, которые будут определены в следующем разделе, на сфере можно выбрать ось $z = x^3$ и полярный угол $\varphi = \mathsf{arctg}\,(x^2/x^1)$. Используя определение (18.35), нетрудно проверить, что

$$[z, \varphi] = 1, \quad [\varphi, \varphi] = 0, \quad [z, z] = 0.$$

Таким образом, симплектическими слоями, соответствующими постоянной функции Казимира, являются сферы, а координатами Дарбу – цилиндрические координаты.

Скобке Ли–Пуассона (18.35) соответствует хорошо известный пример из механики твердого тела. Рассмотрим вращающееся твердое тело с покоящимся центром масс в декартовой системе координат, оси которой направлены по главным осям инерции. Гамильтониан системы в этом случае имеет вид

$$H(p) = \frac{x_1^2}{2I_1} + \frac{x_2^2}{2I_2} + \frac{x_3^2}{2I_3},$$

где x_i, $i = 1, 2, 3$ – моменты количества движения и $I_{1,2,3}$ – моменты инерции твердого тела. Если для координат x_i определить скобку Пуассона (18.35), то гамильтоновы уравнения движения примут вид

$$\begin{aligned}
\dot{x}_1 &= \frac{I_2 - I_3}{I_2 I_3} x_2 x_3, \\
\dot{x}_2 &= \frac{I_3 - I_1}{I_1 I_3} x_1 x_3, \\
\dot{x}_3 &= \frac{I_1 - I_2}{I_1 I_2} x_1 x_2.
\end{aligned}$$

Это есть *уравнения Эйлера* вращения твердого тела (см., например, [?]). □

18.5 Отображения пуассоновых многообразий

Определение. Пусть $\varphi : \mathbb{M} \to \mathbb{N}$ – гладкое отображение двух пуассоновых многообразий. Это отображение называется *пуассоновым*, если оно сохраняет скобку Пуассона:

$$[f \circ \varphi, g \circ \varphi]_{\mathbb{M}} = [f, g]_{\mathbb{N}} \circ \varphi, \tag{18.36}$$

где f и g – произвольные функции на многообразии $\mathbb{N}$. Соответственно, подмногообразие $\mathbb{M} \subset \mathbb{N}$ называется *пуассоновым*, если вложение $\mathbb{M} \hookrightarrow \mathbb{N}$ является пуассоновым отображением. В этом случае мы говорим, что пуассонова структура на $\mathbb{N}$ *сужена* до пуассоновой структуры на подмногообразии $\mathbb{M} \subset \mathbb{N}$.

С другой стороны, если пуассонова структура задана на многообразии $\mathbb{M}$, то ее всегда можно отобразить на образ $\varphi(\mathbb{M}) \subseteq \mathbb{N}$ с помощью дифференциала отображения. Эту пуассонову структуру на образе $\varphi(\mathbb{M}) \subset \mathbb{N}$ будем называть *индуцированной*.

В классической механике пуассоново отображение фазового пространства на себя называется *каноническим преобразованием.* □

Выпишем структурные функции для индуцированной пуассоновой структуры. Пусть отображение $\varphi:\ \mathbb{M}\to\mathbb{N}$ в координатах задается функциями $y^{\mathrm{A}}(x)$, где y^{A}, $\mathrm{A}=1,\dots,\dim\mathbb{N}$ и x^α, $\alpha=1,\dots,\dim\mathbb{M}$ – координаты соответственно на $\mathbb{N}$ и $\mathbb{M}$. Тогда индуцированная пуассонова структура на образе $\varphi(\mathbb{M})\subseteq\mathbb{N}$ всегда определена и имеет следующие структурные функции:

$$J^{\mathrm{AB}}=J^{\alpha\beta}\partial_\alpha y^{\mathrm{A}}\partial_\beta y^{\mathrm{B}}.$$

Из свойств умножения матриц следует, что если отображение φ – вложение, то ранг индуцированной пуассоновой структуры J^{AB} равен рангу пуассоновой структуры на подмногообразии $\mathbb{M}$.

Рассмотренный ниже пример показывает, что пуассонову структуру на $\mathbb{N}$ в общем случае нельзя сузить на произвольное подмногообразие $\mathbb{M}\subset\mathbb{N}$.

Пример 18.5.1. Приведем простой пример непуассонова вложения, которое, на первый взгляд, должно быть таковым. Пусть многообразие $\mathbb{N}$ является прямым произведением двух многообразий, $\mathbb{N}=\mathbb{M}\times\mathbb{K}$. Обозначим координаты и пуассонову структуру на $\mathbb{M}$ через x^α, $\alpha=1,\dots,\dim\mathbb{M}$, и $J^{\alpha\beta}$. Пусть на $\mathbb{K}$ также заданы координаты y^μ, $\mu=1,\dots,\dim\mathbb{K}$, и пуассонова структура $J^{\mu\nu}$. Тогда матрица

$$J=\begin{pmatrix}J^{\alpha\beta}&0\\0&J^{\mu\nu}\end{pmatrix}$$

определяет пуассонову структуру на $\mathbb{N}$ в координатах $\{x,y\}$. Это пуассонова структура индуцирована двумя естественными вложениями $\mathbb{M}\hookrightarrow\mathbb{M}\times\mathbb{K}$ и $\mathbb{K}\hookrightarrow\mathbb{M}\times\mathbb{K}$. Рассмотрим вложение

$$\varphi:\quad \mathbb{M}\ni x\mapsto(x,0)\in\mathbb{N}=\mathbb{M}\times\mathbb{K}.$$

Пусть на $\mathbb{N}$ заданы две произвольные функции $f(x,y)$ и $g(x,y)$. Вычислим левую и правую части равенства (18.36):

$$\begin{aligned}[f\circ\varphi,g\circ\varphi]_{\mathbb{M}}&=J^{\alpha\beta}\partial_\alpha f\partial_\beta g|_{y=0},\\ [f,g]_{\mathbb{N}}\circ\varphi&=\left(J^{\alpha\beta}\partial_\alpha f\partial_\beta g+J^{\mu\nu}\partial_\mu f\partial_\nu g\right)|_{y=0}.\end{aligned}$$

Мы видим, что в общем случае равенство (18.36) не выполняется. Отсюда следует, что вложение φ не является пуассоновым. □

Предложение 18.5.1. *Пусть $(\mathbb{M},J)$ – пуассоново многообразие и X_H – гамильтоново векторное поле. Тогда поток векторного поля*

$$s_t:=\exp(tX_H):\quad \mathbb{M}\ni x(0)\mapsto x(t)\in\mathbb{M},$$

где $x(t)$ – интегральная кривая для X_H, определяет пуассоново отображение.

Доказательство. Пусть $f,g\in\mathcal{C}^\infty(\mathbb{M})$. При малых t поток векторного поля действует на x следующим образом $s_t:\ x^\alpha\mapsto x^\alpha+tX_H^\alpha$. Поэтому

$$f\circ s_t=f(x^\alpha+tX_H^\alpha)\simeq f(x)+tX_Hf.$$

Продифференцируем по t условие пуассоновости (18.36) и положим $t = 0$. Тогда условие пуассоновости примет вид

$$[X_H f, g] + [f, X_H g] = X_H[f, g].$$

Из формулы (18.30) следует, что полученное равенство совпадает с тождествами Якоби. При $t = 0$ экспоненциальное отображение является тождественным и, следовательно, пуассоново. Поэтому интегрирование условия (18.36) вдоль X_H доказывает его выполнение при всех t, для которых определены интегральные кривые. □

Пример 18.5.2. Пусть на фазовой плоскости $q, p \in \mathbb{R}^2$ задана каноническая пуассонова структура (18.31). Рассмотрим гармонический осциллятор с гамильтонианом

$$H = \frac{1}{2}(q^2 + p^2).$$

Соответствующее гамильтоново векторное поле имеет вид

$$X_H = q\partial_p - p\partial_q$$

и задает группу вращений плоскости. Таким образом, каждое вращение фазовой плоскости задает пуассоново преобразование для гармонического осциллятора. Оно является каноническим, т.к. сохраняет скобку Пуассона. □

Поскольку любой гамильтонов поток сохраняет скобку Пуассона, то он, в частности сохраняет ее ранг. Поэтому справедливо

Следствие. Для любого гамильтонова векторного поля X_H на пуассоновом многообразии $(\mathbb{M}, J)$ ранг пуассоновой структуры J постоянен вдоль произвольной интегральной кривой для X_H. □

Это следствие является, по существу, переформулировкой предложения 18.3.2.

Предложение 18.5.2. *Если ранг пуассоновой структуры в какой либо точке $x \in \mathbb{M}$ пуассонова многообразия $(\mathbb{M}, J)$ равен нулю, то эта точка является неподвижной для любой гамильтоновой системы H на $\mathbb{M}$.*

Доказательство. Из курса линейной алгебры известно, что если матрица антисимметрична и ее ранг равен нулю, то она сама равна нулю. Если в точке $x \in \mathbb{M}$ ранг пуассоновой структуры равен нулю, то гамильтоново векторное поле X_H в этой точке для произвольного гамильтониана H обращается в нуль. Следовательно, точка x является неподвижной. □

Пример 18.5.3. Пуассонова структура в начале координат в примере 18.4.1 имеет нулевой ранг. Она остается неподвижной для любого гамильтонова потока. □

Пусть $\mathbb{M}$ – пуассоново многообразие. Тогда для каждой точки $x \in \mathbb{M}$ существует единственное линейное отображение кокасательного пространства в соответствующее касательное

$$J : \quad \mathbb{T}^*_x(\mathbb{M}) \to \mathbb{T}_x(\mathbb{M}), \tag{18.37}$$

такое, что для любой функции $f(x)$ справедливо равенство

$$J(df) = [f, x^\alpha]\partial_\alpha \in \mathbb{T}_x(\mathbb{M}).$$

Это есть рассмотренное ранее отображение (18.27). Для произвольной 1-формы $A = dx^\alpha A_\alpha$ отображение J в компонентах задается матрицей структурных функций:

$$X^\alpha = -J^{\alpha\beta} A_\beta.$$

Для симплектических многообразий, для которых матрица $J^{\alpha\beta}$ невырождена, отображение J является взаимно однозначным.

Ясно, что ядром отображения (18.37) является линейная оболочка дифференциалов функций Казимира dc^{A}.

Обозначим образ отображения J в точке $x \in \mathbb{M}$ через

$$\mathbb{J}_x(\mathbb{M}) = \{J(A) \in \mathbb{T}_x(\mathbb{M}) : \quad \forall A \in \mathbb{T}^*_x(\mathbb{M})\}.$$

Размерность векторного подпространства $\mathbb{J}_x(\mathbb{M}) \subset \mathbb{T}_x(\mathbb{M})$ равна рангу пуассоновой структуры $\mathsf{dim}\, \mathbb{J}_x(\mathbb{M}) = \mathsf{rank}\, J^{\alpha\beta}(x)$. Если ранг пуассоновой структуры на $\mathbb{M}$ является постоянным и равен $2m$, то совокупность подпространств $\mathbb{J}_x(\mathbb{M})$ для всех точек x задает на $\mathbb{M}$ распределение векторных полей $\mathcal{J}_{2m}(\mathbb{M})$ размерности $2m$, которое было определено в разделе 3.11. При этом, если пуассонова структура дифференцируема, то соответствующее распределение векторных полей также дифференцируемо.

В общем случае ранг пуассоновой структуры может меняться от точки к точке. Из определения отображения J следует, что образ $\mathbb{J}_x(\mathbb{M})$ является линейной оболочкой всех гамильтоновых векторных полей на $\mathbb{M}$ в точке x.

Введенные понятия позволяют сформулировать следующее утверждение.

Теорема 18.5.1. *Подмногообразие $\mathbb{M}$ пуассонова многообразия $(\mathbb{N}, J)$ является пуассоновым тогда и только тогда, когда $\mathbb{J}_x(\mathbb{M}) := \mathbb{J}_x(\mathbb{N})|_{\mathbb{M}} \subset \mathbb{T}_x(\mathbb{M})$ для всех $x \in \mathbb{M}$, т.е. каждое гамильтоново векторное поле на $\mathbb{N}$ всюду касается $\mathbb{M}$. В частности, если $\mathbb{J}_x(\mathbb{M}) = \mathbb{T}_x(\mathbb{M})$ для всех $x \in \mathbb{M}$, то $\mathbb{M}$ является симплектическим подмногообразием в $\mathbb{N}$.*

Доказательство. См., например, [**?**]. □

Из этой теоремы следует, что, поскольку размерность пространства $\mathbb{J}_x(\mathbb{M})$ совпадает с рангом пуассоновой структуры в данной точке, то размерность пуассонова подмногообразия не может быть меньше ранга, $\mathsf{dim}\, \mathbb{M} \geq \mathsf{rank}\, J$.

Пусть $\mathbb{M} \subset \mathbb{N}$ – пуассоново подмногообразие. Как было отмечено в теореме 18.5.1, любое гамильтоново векторное поле X_H на $\mathbb{N}$ касается подмногообразия $\mathbb{M}$. Это означает, что его сужение на $\mathbb{M}$ может быть получено из сужения гамильтониана на $\mathbb{M}$:

$$X_H|_{\mathbb{M}} = X_{\widetilde{H}}, \quad \text{где} \quad \widetilde{H} = H|_{\mathbb{M}}.$$

Допустим, что нас интересуют траектории движения гамильтоновой системы с некоторым гамильтонианом H, которые начинаются в некоторой точке пуассонова подмногообразия $x \in \mathbb{M} \subset \mathbb{N}$. В этом случае можно ограничиться движением точки в подмногообразии $\mathbb{M}$, которое порождается суженным гамильтонианом $\widetilde{H}$, тем самым понизив порядок гамильтоновой системы. Можно поставить вопрос, каково минимальное пуассоново подмногообразие для данных начальных условий. Ответ дает следующая

Теорема 18.5.2. *Пусть $(\mathbb{N}, J)$ – пуассоново многообразие. Тогда соответствующее распределение гамильтоновых векторных полей $\mathcal{J}(\mathbb{N})$ на $\mathbb{N}$ интегрируемо, т.е. через каждую точку $x \in \mathbb{N}$ проходит интегральное подмногообразие $\mathbb{M}$ распределения*

$\mathcal{J}(\mathbb{N})$, для которого $\mathbb{T}_y(\mathbb{M}) = \mathbb{J}_y(\mathbb{M})$ в любой точке $y \in \mathbb{M}$. Всякое интегральное подмногообразие является симплектическим подмногообразием в $\mathbb{N}$, и в совокупности эти подмногообразия определяют симплектическое слоение пуассонова многообразия $\mathbb{N}$. Кроме того, если $H : \mathbb{N} \to \mathbb{R}$ – произвольный гамильтониан и $x(t)$ – соответствующая траектория системы, проходящая через точку $x_0 \in \mathbb{N}$, то $x(t)$ остается в одном и том же интегральном подмногообразии $\mathbb{M}$ при всех t.

Доказательство. Как уже отмечалось, распределение $\mathcal{J}(\mathbb{M})$ является линейной оболочкой всех гамильтоновых векторных полей на $\mathbb{M}$. Поскольку скобка Пуассона гамильтоновых векторных полей является гамильтоновым векторным полем (18.28), то распределение $\mathcal{J}(\mathbb{M})$ находится в инволюции. Отсюда по теореме Фробениуса следует существование интегрального подмногообразия. Остальные утверждения теоремы следуют из теоремы 18.5.1 и инвариантности ранга пуассоновой структуры вдоль гамильтоновых векторных полей. □

Проиллюстрируем данную теорему для пуассонова многообразия $(\mathbb{N}, J)$, $\dim \mathbb{N} = n$, с пуассоновой структурой постоянного ранга $\operatorname{rank} J^{\alpha\beta} = 2m < n$ путем построения специальной системы координат. В некоторой окрестности $\mathbb{U} \subset \mathbb{N}$ существует $(n - 2m)$ функционально независимых функций Казимира $c^{\text{A}}(x)$, $\text{A} = 1, \ldots, n - 2m$. Поверхности уровня функций Казимира $c^{\text{A}} = \operatorname{const}$ определяют $2m$-мерное подмногообразие $\mathbb{M} \subset \mathbb{U}$. Дополним функции Казимира $2m$ функционально независимыми скалярными полями g^{M}, $\text{M} = 1, \ldots, 2m$, (координаты на подмногообразии $\mathbb{M}$) таким образом, чтобы выполнялись условия

$$\operatorname{rank} [g^{\text{M}}, g^{\text{N}}]_{\mathbb{M}} = 2m.$$

Это всегда можно сделать, т.к. ранг пуассоновой структуры равен $2m$. Совокупность функций $\{c^{\text{A}}, g^{\text{M}}\}$ выберем в качестве новой системы координат на $\mathbb{U}$. Пуассонова структура в этой системе координат примет вид

$$J = \begin{pmatrix} J^{\text{AB}} = [c^{\text{A}}, c^{\text{B}}] & J^{\text{AN}} = [c^{\text{A}}, g^{\text{N}}] \\ J^{\text{MB}} = [g^{\text{M}}, c^{\text{B}}] & J^{\text{MN}} = [g^{\text{M}}, g^{\text{N}}] \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & J^{\text{MN}} \end{pmatrix},$$

т.к. скобка Пуассона функции Казимира с любой функцией на $\mathbb{N}$ равна нулю. Таким образом, сужение Пуассоновой структуры на подмногообразие $\mathbb{M}$ определено и невырождено. То есть поверхности уровней функций Казимира представляют собой симплектические многообразия. Пусть на $\mathbb{N}$ задан гамильтониан. Соответствующие гамильтоновы уравнения

$$\dot{c}^{\text{A}} = J^{\text{AB}} \partial_{\text{B}} H + J^{\text{AN}} \partial_{\text{N}} H = 0, \tag{18.38}$$

$$\dot{g}^{\text{M}} = J^{\text{MB}} \partial_{\text{B}} H + J^{\text{MN}} \partial_{\text{N}} H = J^{\text{MN}} \partial_{\text{N}} H \tag{18.39}$$

определяют траектории системы. Поскольку уравнение (18.38) имеет решение $c^{\text{A}} = \operatorname{const}$, то траектория гамильтоновой системы, проходящей через точку $x_0 \in \mathbb{N}$, принадлежит соответствующему симплектическому подмногообразию $\mathbb{M}$:

$$\dot{g}^{\text{M}} = \widetilde{J}^{\text{MN}} \partial_{\text{N}} \widetilde{H},$$

где

$$\widetilde{J}^{\text{MN}} = J^{\text{MN}}|_{\mathbb{M}} \quad \text{и} \quad \widetilde{H} = H|_{\mathbb{M}}.$$

Таким образом, всякое пуассоново многообразие $(\mathbb{N}, J)$ расщепляется на симплектические подмногообразия – слои симплектического слоения. Размерность любого

такого слоя $\mathbb{M}$ равна рангу пуассоновой структуры в произвольной точке $x \in \mathbb{M}$. Это означает, что, если ранг пуассоновой структуры на $\mathbb{N}$ непостоянен, то размерность симплектических слоев будет различной.

Теорема 18.5.3 (Дарбу). *Пусть $(\mathbb{M}, J)$ – пуассоново многообразие размерности n. Если пуассонова структура на многообразии имеет постоянный ранг, $\operatorname{rank} J = 2m \leq n$, то в некоторой окрестности $\mathbb{U}$ произвольной точки $x \in \mathbb{M}$ существует такая система координат*

$$\{x^\alpha\} = \{q^{\text{м}}, p_{\text{м}}, c^{\text{а}}\} = \{q^1, \dots, q^m, p_1, \dots, p_m, c^1, \dots, c^{n-2m}\},$$

в которой скобка Пуассона двух произвольных функций $f, g \in \mathcal{C}^\infty(\mathbb{U})$ имеет вид

$$[f, g] = \frac{\partial f}{\partial q^{\text{м}}} \frac{\partial g}{\partial p_{\text{м}}} - \frac{\partial f}{\partial p_{\text{м}}} \frac{\partial g}{\partial q^{\text{м}}}. \tag{18.40}$$

Координаты $c^{\text{а}}$ являются функциями Казимира пуассоновой структуры J, и их постоянное значение определяет симплектическое слоение пуассонова многообразия $(\mathbb{U}, J)$.

Доказательство. Если ранг пуассоновой структуры равен нулю, то $J^{\alpha\beta} = 0$ и доказывать нечего. Любая система координат на $\mathbb{U}$ удовлетворяет утверждению теоремы.

Допустим, что ранг отличен от нуля. Зафиксируем точку $x_0 \in \mathbb{M}$ и выберем в ее окрестности $\mathbb{U}$ две такие функции $Q, f \in \mathcal{C}^\infty(\mathbb{U})$, что

$$[Q, f] = X_Q f \neq 0.$$

В частности $X_Q|_{x_0} \neq 0$. Согласно теореме 3.6.3 возможно в меньшей окрестности существует такая система координат $Q, x^2, \dots, x^n$, что $X_Q = \partial_Q$. Тогда существует такая функция $P \in \mathcal{C}^\infty(\mathbb{U})$, что выполнено условие

$$X_Q P = [Q, P] = 1.$$

Из предложения 18.3.1 следует равенство

$$[X_Q, X_P] = X_{[Q,P]} = 0.$$

Кроме того в силу антисимметрии $[X_Q, X_Q] = 0$ и $[X_P, X_P] = 0$. Таким образом, X_Q, X_P образуют пару коммутирующих векторных полей. Теорема Фробениуса позволяет дополнить пару функций $q^1 = Q$, $p_1 = P$ до локальной системы координат $q^1, p_1, y^3, \dots, y^n$ в некоторой, возможно меньшей, окрестности точки x_0. Поскольку $J^{\alpha\beta} = [x^\alpha, x^\beta]$ и выполнены равенства $[q^1, y^i] = 0$, $[p_1, y^i] = 0$, $i = 3, \dots, n$, то структурные функции в данной системе координат принимают вид

$$J = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & J^{ij} \end{pmatrix},$$

где $J^{ij} = [y^i, y^j]$. Покажем, что матрица J^{ij} не зависит от координат q^1 и p_1. Действительно,

$$\frac{\partial J^{ij}}{\partial q^1} = [q^1, J^{ij}] = [q^1, [y^i, y^j]] = 0,$$

где мы воспользовались тождествами Якоби. Аналогично доказывается, что матрица J^{ij} не зависит от p_1. Таким образом, матрица J^{ij} задает пуассонову структуру на подмногообразии $q^1 = \mathsf{const}$, $p_1 = \mathsf{const}$. Ранг этой структуры на два меньше исходного, $\operatorname{rank} J^{ij} = 2m - 2$. Поэтому, если $m > 1$, то этот процесс можно продолжить. □

Глава 19

Принцип наименьшего действия

Можно с уверенностью сказать, что в основе построения моделей современной математической физики лежит принцип наименьшего действия. Этот принцип требует стационарности некоторого функционала – действия – относительно вариаций полей, описывающих данную модель. В результате мы получаем систему уравнений Эйлера–Лагранжа, которая принимается в качестве уравнений движения, уравнений равновесия и т.д. для данной модели. При этом инвариантность действия относительно некоторой группы преобразований приводит к ковариантным уравнениям движения и к законам сохранения, которые играют важнейшую роль в физике.

19.1 Постановка вариационных задач

Начнем с постановки вариационной задачи в евклидовом пространстве. Предположим, для простоты, что $\mathbb{M} \subset \mathbb{R}^n$ – ограниченная область евклидова пространства с кусочно гладкой границей $\partial\mathbb{M}$. Пусть в этой области задан некоторый набор дважды непрерывно дифференцируемых функций вплоть до границы, $\varphi = \{\varphi^a\} \in \mathcal{C}^2(\overline{\mathbb{M}})$, $a = 1, \dots, \mathrm{N}$. Это значит, что все функции и их производные до второго порядка непрерывны и ограничены в $\mathbb{M}$ и имеют конечный предел на границе. Если x^α, $\alpha = 1, \dots, n$, система координат на $\mathbb{M}$, то обозначим, для краткости, все первые производные полей через $\partial\varphi = \{\partial_\alpha \varphi^a\}$. Тогда на $\mathbb{M}$ определен *функционал действия* или, просто, *действие*

$$S[\varphi] = \int_{\mathbb{M}} dx\, L(x, \varphi, \partial\varphi), \tag{19.1}$$

где L – некоторая функция от n переменных x^α, N переменных φ^a и $n\mathrm{N}$ переменных $\partial_\alpha \varphi^a$. Она предполагается дважды непрерывно дифференцируемой функцией переменных $x \in \mathbb{M}$ и всех остальных переменных для всех конечных значений φ и $\partial\varphi$. Назовем функцию $L(x, \varphi, \partial\varphi)$ *лагранжевой плотностью* или *лагранжианом* данной модели, которая описывается набором полей φ.

Для определения функционала действия мы ограничили себя классом дважды непрерывно дифференцируемых функций $\mathcal{C}^2(\overline{\mathbb{M}})$. Тогда функционал действия задает отображение

$$S: \quad \left[\mathcal{C}^2(\overline{\mathbb{M}})\right]^{\mathrm{N}} \ni \varphi \mapsto S[\varphi] \in \mathbb{R} \tag{19.2}$$

бесконечномерного функционального пространства в поле вещественных чисел. Для постановки вариационной задачи нам важно, что функциональное пространство $\left[\mathcal{C}^2(\overline{\mathbb{M}})\right]^{\mathrm{N}}$ снабжено структурой линейного пространства с обычным поточечным сложением и умножением на вещественные числа. Для того, чтобы говорить о непрерывности и

вариационных производных отображения S, введем на $\left[\mathcal{C}^2(\overline{\mathbb{M}})\right]^{\mathrm{N}}$ норму:

$$\|\varphi\| := \sum_{a=1}^{\mathrm{N}} \left(\sup_{x\in\mathbb{M}} |\varphi^a| + \sum_{\alpha=1}^{n} \sup_{x\in\mathbb{M}} |\partial_\alpha \varphi^a| \right). \tag{19.3}$$

Тем самым класс рассматриваемых функций $\left[\mathcal{C}^2(\overline{\mathbb{M}})\right]^{\mathrm{N}}$ становится нормированным линейным функциональным пространством. По данной норме строится метрика и определяется топология, относительно которой отображение (19.2) непрерывно.

Назовем *вариацией* функции φ^a разность двух представителей из класса рассматриваемых функций: $\delta\varphi^a = \varphi'^a - \varphi^a$, где φ'^a – произвольная функция из $\mathcal{C}^2(\overline{\mathbb{M}})$. Ясно, что вариация функции принадлежит тому же классу, что и сама функция. Для малых вариаций функций $\epsilon\delta\varphi^a$, где $\epsilon > 0$ – малая величина, вариация (главная линейная часть) функционала действия, если она существует, равна

$$\begin{aligned} \delta S &= S[\varphi + \epsilon\delta\varphi] - S[\varphi] = \\ &= \int_{\mathbb{M}} dx \left[\frac{\partial L}{\partial \varphi^a} - \partial_\alpha \left(\frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \right) \right] \epsilon\delta\varphi^a + \int_{\partial\mathbb{M}} ds_\alpha \frac{\partial L}{\partial(\partial_\alpha\varphi^a)} \epsilon\delta\varphi^a + \mathsf{o}(\epsilon), \end{aligned} \tag{19.4}$$

где второе слагаемое возникло при интегрировании по частям, и ds_α обозначает ориентированный элемент объема границы $\partial\mathbb{M}$. Отметим, что при вычислении вариации действия (19.1) область интегрирования $\mathbb{M}$ считалась неизменной.

В рассматриваемом классе функций φ и лагранжианов L вариация функционала всегда существует.

Назовем набор функций φ *стационарной*, или *критической* точкой, или *экстремалью* функционала $S[\varphi]$, если в этой точке линейная часть вариация действия равна нулю, $\delta S = \mathsf{o}(\epsilon)$. Такие точки соответствуют либо локальному минимуму, либо локальному максимуму, либо седловой точке функционала S. Это можно проверить после нахождения экстремали функционала, рассмотрев члены более высокого порядка по ϵ.

Для действия (19.1) можно поставить различные вариационные задачи. Рассмотрим задачи, которые наиболее часто встречаются в физике.

19.1.1 Задача с заданными граничными условиями

Вариационная задача с заданными граничными условиями является наиболее распространенной и самой простой с точки зрения постановки. Рассмотрим класс функций в $\mathcal{C}^2(\overline{\mathbb{M}})$ с заданными граничными условиями

$$\varphi\Big|_{\partial\mathbb{M}} = \varphi_0. \tag{19.5}$$

Поскольку граничные условия для всех функций при фиксированном индексе a одни и те же, то вариации обращаются в нуль на границе

$$\delta\varphi\Big|_{\partial\mathbb{M}} = 0. \tag{19.6}$$

Тогда интеграл по границе области в вариации действия (19.4) обращается в нуль в силу граничных условий (19.6). В рассматриваемом случае существует предел

$$\lim_{\epsilon\to 0} \frac{S[\varphi + \epsilon\delta\varphi] - S[\varphi]}{\epsilon} = \int_{\mathbb{M}} dx \, \frac{\delta S}{\delta\varphi^a} \delta\varphi^a. \tag{19.7}$$

Функция $\delta S/\delta\varphi^a$, стоящая под знаком интеграла, называется *вариационной производной* функционала S по полю φ^a и обозначается также

$$S_{,\,a} = \frac{\delta S}{\delta\varphi^a}. \tag{19.8}$$

Из вида вариации (19.4) получаем явное выражение для вариационной производной

$$S_{,\,a} = \frac{\partial L}{\partial\varphi^a} - \partial_\alpha \frac{\partial L}{\partial(\partial_\alpha\varphi^a)}. \tag{19.9}$$

Вариационная производная $S_{,\,a}$ является ядром линейного оператора, который есть производная по Фреше и, следовательно, по Гато отображения (19.2).

Формула (19.7) позволяет дать другое определение вариационной производной. А именно, вариационной производной функционала S в точке $\varphi \in [\mathcal{C}^2(\mathbb{M})]^{\mathrm{N}}$ назовем производную

$$\frac{\delta S}{\delta\varphi^a} = \left.\frac{\partial S(\varphi+\lambda\psi)}{\partial\lambda}\right|_{\lambda=0}, \tag{19.10}$$

где $\psi \in [\mathcal{C}^2(\mathbb{M})]^{\mathrm{N}}$. Ясно, что если главная линейная часть приращения функционала существует, то она равна производной (19.10). Обратное утверждение в общем случае неверно. Существуют примеры функционалов (менее гладких, чем мы рассматриваем), для которых производная (19.10) определена, однако из их приращения нельзя выделить главную линейную часть. Поэтому определение вариационной производной (19.10) является более общим. В рассматриваемом нами классе функций и лагранжианов данные выше определения эквивалентны.

Определение вариационной производной (19.10) просто обобщается на случай вариационных производных более высокого порядка. Вторые вариационные производные (17.28) были рассмотрены для экстремалей функционала длины кривой.

Из условия стационарности действия $\delta S = \mathsf{o}(\epsilon)$ в силу произвольности вариации $\delta\varphi^a$ и основной леммы вариационного исчисления следует

Теорема 19.1.1. *Набор функций φ при заданных граничных условиях является стационарной точкой действия $S[\varphi]$ тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа*

$$S_{,\,a} = \frac{\partial L}{\partial\varphi^a} - \partial_\alpha \frac{\partial L}{\partial(\partial_\alpha\varphi^a)} = 0. \tag{19.11}$$

В общем случае уравнения Эйлера–Лагранжа представляют собой систему нелинейных дифференциальных уравнений в частных производных второго порядка, число которых N равно числу функций, от которых зависит функционал действия. Их надо решать при заданных граничных условиях (19.5). Решение поставленной вариационной задачи может не существовать, а если оно существует, то может быть неединственно. Это зависит от вида лагранжиана и области $\mathbb{M}$.

Замечание. В общем случае уравнения Эйлера–Лагранжа для заданного лагранжиана могут приводить к противоречию. Например, пусть лагранжиан зависит от одной функции φ и имеет вид $L = \varphi$. Тогда уравнения Эйлера–Лагранжа приводят к противоречию $1 = 0$. Таким образом, для того, чтобы уравнения Эйлера–Лагранжа имели решение, лагранжиан не может быть произвольной функцией полей и их производных. В дальнейшем мы предполагаем, что лагранжиан выбран таким образом, что соответствующие уравнения Эйлера–Лагранжа непротиворечивы. □

Действие (19.1) для заданных уравнений Эйлера–Лагранжа при постановке задачи с фиксированными граничными условиями определено неоднозначно. Действительно, рассмотрим новый лагранжиан, $\widetilde{L}_\alpha = L + \partial_\alpha F$, который отличается от исходного на частную производную от некоторой достаточно гладкой функции $F(x, \varphi, \partial\varphi)$. Тогда действие получит дополнительный вклад, сводящийся к интегралу по границе. Вариация дополнительного слагаемого равна нулю, т.к. вариации всех полей на границе равны нулю. Отсюда следует, что уравнения Эйлера–Лагранжа не изменятся при добавлении к лагранжиану частных производных $\partial_\alpha F$ от произвольной функции.

Рассмотренная вариационная задача наиболее часто рассматривается в моделях математической физики. При этом уравнения Эйлера–Лагранжа приводят к уравнениям движения, равновесия и т.д.

Замечание. Поверхностный интеграл в вариации действия (19.4) в задаче с заданными граничными условиями равен нулю, поскольку вариации полей обращаются в нуль на границе (19.6). Этого достаточно, если область $\mathbb{M}$ ограничена. Однако для действия (19.1), рассматриваемого во всем евклидовом пространстве $\mathbb{R}^n$, ситуация усложняется. В этом случае интеграл по границе является несобственным, т.к. площадь бесконечно удаленной границы стремится к бесконечности. Тогда важно не только граничное условие на вариации, но и их асимптотическое поведение. Так, в общей теории относительности в асимптотически плоском пространстве-времени поверхностный интеграл отличен от нуля, т.к. метрика недостаточно быстро стремится к метрике Минковского на пространственной бесконечности. Анализ граничного поведения полей важен и в общем случае сложен, поскольку зависит от рассматриваемой задачи. На данном этапе мы пренебрежем граничными слагаемыми. □

Замечание. Во многих моделях математической физики, например, в электродинамике, в качестве области $\mathbb{M}$ выбирается все пространство Минковского $\mathbb{R}^{1,3}$. При этом функционал действия не ограничен ни снизу, ни сверху. Кроме того, для многих решений уравнений Эйлера–Лагранжа он расходится. Поэтому говорить о принципе наименьшего действия в строгом смысле не приходится. Тем не менее уравнения Эйлера–Лагранжа имеют смысл, поскольку являются локальным объектом. Поэтому для действия часто пишут формальное выражение, не заботясь о сходимости интеграла. Этот способ очень удобен для получения уравнений с заданными свойствами симметрии. □

19.1.2 Задача со свободными граничными условиями

Для действия (19.1) можно поставить другую вариационную задачу, расширив класс рассматриваемых функций. Пусть по-прежнему все функции дважды непрерывно дифференцируемы $\varphi^a \in \mathcal{C}^2(\overline{\mathbb{M}})$ для всех $a = 1, \dots, \mathrm{N}$, но теперь снимем ограничения, накладываемые граничными условиями (19.5). В этом случае вариации полей $\delta\varphi^a$ также ничем не ограничены на границе $\partial\mathbb{M}$, и из явного вида вариации действия (19.4) следует

Теорема 19.1.2. *Набор функций φ является стационарной точкой действия $S[\varphi]$ тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа (19.11) и граничным условиям*

$$\left.\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\right|_{\partial\mathbb{M}} = 0. \tag{19.12}$$

Таким образом, в задаче со свободными граничными условиями граничные условия все таки возникают из условия стационарности действия.

Пример 19.1.1. Вариационная задача со свободными граничными условиями рассматривается в теории открытых бозонных струн, для которых из вариационного принципа вытекают граничные условия Неймана. □

Постановка вариационной задачи со свободными граничными условиями, зависит от добавления к действию граничных слагаемых. Иногда появления граничных условий (19.12) можно избежать, если из исходного действия вычесть все граничные вклады, которые возникают при интегрировании по частям. В этом случае условие (19.12) тождественно выполняется.

Замечание. В общем случае можно рассматривать смешанные вариационные задачи, когда граничные условия ставятся только для части полей. Именно такого типа задача естественным образом возникает в теории гравитации, где можно считать заданными на границе только физические поля. Нефизические поля находятся, как решение уравнений связей и калибровочных условий, и для них граничные значения не могут быть фиксированы произвольным образом. □

19.1.3 Задача с подвижной границей

Возможна также более общая постановка вариационной задачи для функционала (19.1), когда рассматриваются не только вариации полей, но и самой области $\mathbb{M}$. Сначала уточним постановку задачи. Предположим, что координаты и поля преобразуются следующим образом:

$$x^\alpha \mapsto x'^\alpha(x, \varphi, \partial\varphi, \epsilon), \tag{19.13}$$
$$\varphi^a \mapsto \varphi'^a(x, \varphi, \partial\varphi, \epsilon), \tag{19.14}$$

где ϵ – параметр преобразования. Если поля φ являются заданными функциями от координат, $\varphi = \varphi(x)$, то из уравнения (19.13) можно выразить старые координаты через новые: $x = x(x')$. Тогда подстановка найденных функций в формулу (19.14) позволяет рассматривать новые поля, как функции от новых координат, $\varphi' = \varphi'(x')$.

Мы считаем, что до и после преобразования координаты определены соответственно на ограниченных областях $\mathbb{M}$ и $\mathbb{M}'$ евклидова пространства $\mathbb{R}^n$, т.е. $x \in \mathbb{M}$ и $x' \in \mathbb{M}'$. Тогда под вариацией функционала подразумевается разность

$$\delta S := S[\varphi'(x')] - S[\varphi(x)] = \int_{\mathbb{M}'} dx' L(x', \varphi', \partial'\varphi') - \int_{\mathbb{M}} dx L(x, \varphi, \partial\varphi).$$

Нам требуется найти эту вариацию в линейном по ϵ приближении. При этом мы считаем, что при $\epsilon = 0$ преобразование (19.25) является тождественным. Разлагая формулы преобразования в ряд Тейлора при малых ϵ, получаем

$$\begin{aligned} x^\alpha \mapsto x'^\alpha &= x^\alpha + \delta x^\alpha + \mathsf{o}(\epsilon), \\ \varphi^a(x) \mapsto \varphi'^a(x') &= \varphi^a(x) + \bar\delta\varphi^a(x) + \mathsf{o}(\epsilon), \end{aligned} \tag{19.15}$$

где независимые вариации

$$\delta x^\alpha = \epsilon R^\alpha(x, \varphi, \partial\varphi), \tag{19.16}$$
$$\bar\delta\varphi^a = \varphi'^a(x') - \varphi^a(x) = \epsilon R^a(x, \varphi, \partial\varphi) \tag{19.17}$$

определяются некоторыми функциями $R^\alpha(x,\varphi,\partial\varphi)$ и $R^a(x,\varphi,\partial\varphi)$. Отметим, что вариация полей (19.17) определена, как разность значений полей в различных точках и не совпадает с вариацией формы поля, рассмотренной в разделе 3.13. Напомним, что под вариацией формы функции мы понимаем разность значений этой функции после и до преобразования в одной и той же точке:

$$\delta\varphi^a(x) = \varphi'^a(x) - \varphi^a(x).$$

При этом преобразования (19.15) рассматриваются, как активные (см. раздел 2.5). Тогда для вариации формы функции справедливо равенство

$$\delta\varphi^a(x) = \bar\delta\varphi^a - \delta x^\alpha\partial_\alpha\varphi^a = \epsilon(R^a - R^\alpha\partial_\alpha\varphi^a). \tag{19.18}$$

По-построению, для постоянного параметра преобразования вариация формы функции δ перестановочна с операцией частного дифференцирования ∂_α. Вариация действия относительно преобразований (19.15) имеет вид

$$\delta S = \int_{\mathbb{M}'} dx'\, L(x') - \int_{\mathbb{M}} dx\, L(x) = \int_{\mathbb{M}'} dx'\, \Big(L(x) + \bar\delta L(x)\Big) - \int_{\mathbb{M}} dx\, L(x),$$

где $L(x) = L\Big(x,\varphi(x),\partial_\alpha\varphi(x)\Big)$. Здесь и в дальнейшем мы будем рассматривать вариацию действия только с точностью до слагаемых, линейных по ϵ, и не будем это указывать. Вариации лагранжиана и элемента объема равны, соответственно,

$$\bar\delta L(x) = \frac{\partial L}{\partial x^\alpha}\delta x^\alpha + \frac{\partial L}{\partial\varphi^a}\bar\delta\varphi^a + \frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\bar\delta(\partial_\alpha\varphi^a) = \delta L + \frac{dL}{dx^\alpha}\delta x^\alpha, \tag{19.19}$$

$$dx' \simeq dx + dx\frac{\partial(\delta x^\alpha)}{\partial x^\alpha}, \tag{19.20}$$

где

$$\frac{dL}{dx^\alpha} = \frac{\partial L}{\partial x^\alpha} + \frac{\partial L}{\partial\varphi^a}\partial_\alpha\varphi^a + \frac{\partial L}{\partial(\partial_\beta\varphi^a)}\partial^2_{\alpha\beta}\varphi^a.$$

Используя преобразование (19.20), перейдем от интегрирования по $\mathbb{M}'$ к интегрированию по исходной области $\mathbb{M}$ и перепишем вариацию действия в виде

$$\delta S = \int_{\mathbb{M}} dx\, \left(\delta L + \frac{dL}{dx^\alpha}\delta x^\alpha + L\frac{\partial(\delta x^\alpha)}{\partial x^\alpha}\right) = \int_{\mathbb{M}} dx\, \Big(\delta L + \partial_\alpha(L\delta x^\alpha)\Big). \tag{19.21}$$

Теперь перепишем вариацию формы лагранжиана:

$$\begin{aligned}\delta L &= \frac{\partial L}{\partial\varphi^a}\delta\varphi^a + \frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\delta(\partial_\alpha\varphi^a) = \\ &= \left(\frac{\partial L}{\partial\varphi^a} - \partial_\alpha\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\right)\delta\varphi^a + \partial_\alpha\left(\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\delta\varphi^a\right).\end{aligned} \tag{19.22}$$

При этом существенно, что параметр преобразования постоянен, $\epsilon = \mathsf{const}$, и, следовательно,

$$\delta(\partial_\alpha\varphi^a) = \partial_\alpha(\delta\varphi^a).$$

Воспользовавшись формулой Стокса, дивергентные слагаемые можно переписать в виде поверхностного интеграла. Таким образом для вариации действия получаем окончательное выражение

$$\delta S = \int_{\mathbb{M}} dx\, \left(\frac{\partial L}{\partial\varphi^a} - \partial_\alpha\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\right)\delta\varphi^a + \int_{\partial\mathbb{M}} ds_\alpha\, \left(\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\delta\varphi^a + L\delta x^\alpha\right). \tag{19.23}$$

Поверхностный интеграл перепишем с учетом выражения для вариации формы функции (19.18):

$$\int_{\partial\mathbb{M}} ds_\alpha \left(\frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \bar{\delta}\varphi^a + \left(L\delta^\alpha_\beta - \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \partial_\beta \varphi^a \right) \delta x^\beta \right).$$

поскольку вариации координат δx^α и функций $\bar{\delta}\varphi^a$ произвольны и независимы, то из полученного выражения для вариации действия вытекает

Теорема 19.1.3. *Набор функций φ является стационарной точкой действия $S[\varphi]$ в вариационной задаче с подвижной границей тогда и только тогда, когда он удовлетворяет системе уравнений Эйлера–Лагранжа (19.11) и граничным условиям:*

$$\begin{aligned} \left. \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \right|_{\partial\mathbb{M}} &= 0, \\ \left. L\delta^\alpha_\beta - \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} \partial_\beta \varphi^a \right|_{\partial\mathbb{M}} &= 0. \end{aligned} \tag{19.24}$$

В последней вариационной задаче с подвижной границей мы имеем $n(n+\mathrm{N})$ граничных условий, число которых быстро растет с увеличением размерности пространства. Соответствующая вариационная задача далеко не всегда имеет решение. Чтобы уменьшить число независимых граничных условий, предположим, что в окрестности границы $\partial\mathbb{M}$ поля принимают наперед заданные значения:

$$\varphi^a(x) = \Phi^a(x),$$

где функции Φ^a заданы в некоторой окрестности $\partial\mathbb{M}$. Тогда вариации функций и координат связаны соотношением

$$\bar{\delta}\varphi^a\big|_{\partial\mathbb{M}} = \delta x^\alpha \partial_\alpha \Phi^a|_{\partial\mathbb{M}}\,.$$

В этом случае вместо $n(n+\mathrm{N})$ граничных условий (19.24) остаются n^2 условий:

$$\left[L\delta^\alpha_\beta + \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} (\partial_\beta \Phi^a - \partial_\beta \varphi^a) \right]_{\partial\mathbb{M}} = 0. \tag{19.25}$$

Эти граничные условия называются *условиями трансверсальности.*

Стационарные точки действия для задачи с подвижной границей являются также стационарными точками для двух задач с фиксированной областью $\mathbb{M}$. Выполнение уравнений Эйлера–Лагранжа является необходимым условием во всех трех рассмотренных вариационных задачах. В задачах со свободными граничными условиями и с подвижной границей к уравнениям Эйлера–Лагранжа добавляются граничные условия.

19.1.4 Задача на условную стационарную точку

В настоящем разделе мы ограничимся рассмотрением вариационных задач на конечном отрезке $[p,q] \subset \mathbb{R}$ с заданными граничными условиями. Пусть требуется найти стационарную точку действия (19.1) в классе функций $\varphi^a \in \mathcal{C}^2\big([p,q]\big)$, $a = 1, \ldots, \mathrm{N}$ при наличии $\mathrm{M} < \mathrm{N}$ независимых дополнительных условий, которые называются *связями*

$$G_{\mathrm{A}}(x, \varphi, \partial\varphi) = 0, \qquad \mathrm{A} = 1, \ldots, \mathrm{M} < \mathrm{N}, \tag{19.26}$$

где G_{A} – достаточно гладкие функции своих аргументов. Мы предполагаем, что связи G_{A} не противоречат граничным условиям и функционально независимы, т.е. ни одна из связей не является следствием остальных. В частности, ни одна из связей не выполняется тождественно для всех функций φ. Функциональная независимость связей означает, что матрица производных

$$\frac{\partial G_{\text{A}}}{\partial\Big(\varphi^a, \partial_x\varphi^a\Big)}$$

имеет постоянный ранг M. Отсюда следует, что локально связи можно разрешить относительно M функций или их первых производных, рассматривая остальные N−M функций как независимые.

В общем случае связи являются дифференциальными уравнениями, и их решения содержат произвольные постоянные. Мы предполагаем, что этот произвол устранен, например, наложением граничных условий, либо каким либо иным образом.

В частном случае связи могут быть алгебраическими уравнениями на неизвестные функции $G_{\text{A}}(x,\varphi) = 0$. В этом случае они называются *голономными*. В общем случае связи (19.26) называются *неголономными*.

При наличии связей вариации функций не являются независимыми, и выполнение уравнений Эйлера–Лагранжа для исходного действия (19.1) не является необходимым условием. Прямым способом решения задачи на условный экстремум является явное разрешение связей относительно M функций, подстановка полученного решения в исходное действие и исследование нового действия от N − M функций на безусловный экстремум.

Задачи на условную стационарную точку часто встречаются в математической физике. В частности, к ним приводят все калибровочные модели, инвариантные относительно локальных преобразований полей. В связи с этим введем удобную терминологию, которая часто используется в физике. А именно, назовем поля *нефизическими*, если связи разрешаются относительно этих полей. Остальные поля, относительно которых после исключения нефизических полей возникает задача на безусловную стационарную точку, называются *физическими*. Деление полей на физические и нефизические условно, т.к. связи можно разрешать относительно различных переменных. В то же время число физических (N − M) и нефизических (M) полей, по предположению, постоянно.

Прямой способ исключения нефизических полей неприменим, если связи не решаются явно. Кроме этого исключение части полей может нарушить симметрию задачи, например, лоренц-инвариантность, что часто приводит к существенному усложнению вычислений. Поэтому используют метод неопределенных множителей Лагранжа. А именно, строят полное (total) действие

$$S_{\text{T}} = \int_p^q dx(L - \lambda^{\text{A}} G_{\text{A}}), \tag{19.27}$$

где $\lambda(x) \in \mathcal{C}^1\Big([p,q]\Big)$ – новые функции, которые называются *множителями Лагранжа*. Это действие исследуется на безусловный экстремум. Вариация действия (19.27) по полям φ^a и множителям Лагранжа λ^{A} приводит к N + M уравнениям Эйлера–Лагранжа, M из которых, возникших при вариации по множителям Лагранжа, совпадают с уравнениями связей (19.26). При этом вариации множителей Лагранжа на границе не обязаны быть равными нулю, т.к. они входят в действие без производных и никаких дополнительных граничных условий не возникает. Решение новой задачи

на безусловный экстремум дает решение исходной задачи на условный экстремум, что является содержанием следующего утверждения.

Теорема 19.1.4. *Для функций φ, на которых функционал (19.1) имеет стационарное значение при выполнении уравнений связей (19.26), существует такой набор множителей Лагранжа, что они вместе с полями φ удовлетворяют уравнениям Эйлера–Лагранжа для действия (19.27):*

$$\frac{\delta S_{\text{т}}}{\delta \varphi^a} = 0, \qquad \frac{\delta S_{\text{т}}}{\delta \lambda^{\text{а}}} = G_{\text{а}} = 0.$$

Доказательство. См., например, [**?**]. □

Сформулированная теорема позволяет свести вариационную задачу на условный экстремум к вариационной задаче на безусловный экстремум, но для действия, зависящего от большего числа функций. В теории поля вариационные задачи рассматриваются не на прямой $\mathbb{R}$, а в евклидовом пространстве $\mathbb{R}^n$. Тогда связи представляют собой в общем случае дифференциальные уравнения в частных производных. Для того чтобы доказать аналог теоремы о множителях Лагранжа необходимо зафиксировать каким либо образом класс рассматриваемых связей, что является сложной задачей. На практике метод неопределенных множителей Лагранжа часто используют, не заботясь о его применимости. В таком случае применимость метода необходимо доказывать в каждом конкретном случае.

19.1.5 Другие задачи и терминология

В общем случае функционал действия может зависеть от частных производных полей φ^a любого порядка, вплоть до бесконечного:

$$S[\varphi] = \int_{\mathbb{M}} dx\, L(x, \varphi, \partial\varphi, \partial^2\varphi, \dots). \tag{19.28}$$

Тогда уравнения Эйлера–Лагранжа примут вид

$$S_{,\,a} = \frac{\partial L}{\partial(\varphi^a)} - \partial_\alpha \frac{\partial L}{\partial(\partial_\alpha \varphi^a)} + \partial^2_{\alpha\beta} \frac{\partial L}{\partial(\partial^2_{\alpha\beta}\varphi^a)} - \dots = 0. \tag{19.29}$$

Для простоты в настоящем разделе мы не будем обсуждать возможные граничные слагаемые.

Модели теории поля, лагранжиан которых зависит от производных бесконечного порядка, называются *нелокальными*. Для *локальных* моделей, порядок производных в которых ограничен, ряд (19.29) обрывается. Модели, для которых уравнения Эйлера–Лагранжа содержат производные третьего или более высокого, но конечного порядка называются *моделями с высшими производными*.

Хорошо известно, что порядок уравнений можно понизить, рассматривая частные производные в качестве новых независимых переменных. В этом смысле любую теорию с высшими производными можно свести к модели без высших производных.

Замечание. С физической точки зрения теории с высшими производными представляют определенные трудности, т.к. наличие векторных лоренцевых индексов у полей, как правило, приводит к каноническому гамильтониану, неограниченному снизу за счет вклада временны́х компонент. После квантования такие теории приводят к

гильбертову пространству с индефинитной метрикой, которая на допускает вероятностной интерпретации квантовой теории. Эти трудности можно избежать за счет выбора лагранжиана специального вида или налагая условие калибровочной инвариантности, которое позволяет исключить вклад временны́х компонент в канонический гамильтониан для физических степеней свободы. В квантовой теории поля модели с высшими производными принято считать неудовлетворительными до тех пор, пока не доказана положительная определенность канонического гамильтониана для физических степеней свободы.

Нелокальные теории поля представляют собой еще бóльшие трудности для физической интерпретации, т.к. помимо проблем с индефинитной метрикой гильбертова пространства, в общем случае они нарушают причинность. Это следует из того, что значение функции, разложимой в ряд Тейлора в точке x, в точке $y \neq x$ выражаются через ее значения и значения ее производных в точке x в виде бесконечного ряда. Следовательно, значение функции в некоторой точке может зависеть от ее значений в конусе будущего. □

Допустим, что задана система уравнений Эйлера–Лагранжа. Функционал действия, приводящий к этим или эквивалентным уравнениям Эйлера–Лагранжа, определен неоднозначно. Во-первых, как уже отмечалось при рассмотрении вариационной задачи с заданными граничными условиями, уравнения не изменятся, если к лагранжиану добавить частную производную от функции, зависящей произвольным образом от полей и их частных производных. В частности, можно добавлять члены, имеющие вид дивергенции. Отметим, что если сам лагранжиан равен полной дивергенции, то он не приводит ни к каким уравнениям Эйлера–Лагранжа (получается тождество $0 = 0$). Во-вторых, вместо одного набора полей φ^a можно выбрать другой: $\varphi'^a = \varphi'^a(\varphi)$. Если это преобразование полей невырождено, то новая система уравнений Эйлера–Лагранжа будет эквивалентна старой. Пример дают канонические преобразования в гамильтоновом формализме. Иногда можно изменить даже число новых переменных в лагранжиане, и доказать биективность пространств возникающих решений уравнений Эйлера–Лагранжа. В частности, лагранжев и гамильтонов способы описания динамики точечных частиц, рассмотренные в следующей главе, приводят к эквивалентным системам уравнений, но для разного числа переменных.

Замечание. Выбор независимых переменных в действии чрезвычайно важен, поскольку может привести к существенному упрощению возникающих уравнений движения, особенно в нелинейных теориях. С другой стороны, квантования моделей теории поля, основанные на различном выборе динамических переменных, может привести к различным квантовым теориям. В последнем случае теоретическим критерием выбора способа квантования является простота и самосогласованность конечной квантовой теории. Этот вопрос актуален для построения самосогласованной квантовой теории гравитации, которая в настоящее время отсутствует. □

В релятивистских моделях математической физики, т.е. в моделях, инвариантных относительно преобразований из группы Пуанкаре, действие записывается в виде интеграла по всему пространству Минковского $\mathbb{R}^{1,3}$. В этом случае уравнения Эйлера–Лагранжа (19.11) называются также *уравнениями движения*, поскольку описывают эволюцию системы во времени. При этом для уравнений движения часто ставится не краевая задача, а задача Коши.

Для корректной постановки вариационной задачи необходим глубокий анализ уравнений Эйлера–Лагранжа. Во многих важных случаях эти уравнения настолько сложны, что корректность постановки вариационной задачи доказать не удается. Поэтому в теоретической физике выбор действия означает, как правило, просто удобный способ задания уравнений движения для модели с заданными свойствами симметрии, что, конечно, чрезвычайно важно.

19.2 Первая теорема Нетер

В наиболее содержательных моделях математической физики функционал действия инвариантен относительно глобальных или локальных преобразований симметрии. С каждым преобразованием симметрии связан закон сохранения, что было установлено Эмми Нетер [**?**] в первой и второй теореме, соответственно, для глобальных и локальных преобразований.

Пусть функционал действия (19.1) инвариантен относительно бесконечно малых преобразований

$$x^\alpha \to x'^\alpha = x^\alpha + \delta x^\alpha, \tag{19.30}$$
$$\varphi^a \to \varphi'^a(x') = \varphi^a(x) + \bar\delta\varphi^a(x). \tag{19.31}$$

Рассмотрим независимые вариации координат и полей:

$$\delta x^\alpha = \epsilon^{\text{A}} R_{\text{A}}{}^\alpha(x,\varphi,\partial\varphi), \tag{19.32}$$
$$\bar\delta\varphi^a = \varphi'^a(x') - \varphi^a(x) = \epsilon^{\text{A}} R_{\text{A}}{}^a(x,\varphi,\partial\varphi), \tag{19.33}$$

где $R_{\text{A}}{}^\alpha(x,\varphi,\partial\varphi)$ и $R_{\text{A}}{}^a(x,\varphi,\partial\varphi)$ – некоторые достаточно гладкие функции своих аргументов, которые называются *генераторами* преобразований симметрии, а $\epsilon^{\text{A}}(x)$, $\text{A} = 1,2,\ldots,\text{K}$, – постоянные или локальные параметры преобразований, число которых зависит от рассматриваемой модели. Мы говорим, что каждому значению индекса A соответствует одно преобразование симметрии.

Преобразования (19.30), (19.31) уже рассматривались нами при обсуждении вариационной задачи с подвижной границей. Разница заключается в том, что сейчас у нас не один, а K параметров преобразования, которые, вдобавок, могут зависеть от точки $x \in \mathbb{M}$.

Начнем с доказательства первой теоремы Нетер, т.е. будем считать параметры преобразований постоянными, $\epsilon^{\text{A}} = \mathsf{const}$. Под инвариантностью функционала действия мы понимаем следующее равенство

$$S = \int_{\mathbb{M}} dx\, L(x,\varphi,\partial\varphi) = \int_{\mathbb{M}'} dx'\, L(x',\varphi',\partial'\varphi'), \tag{19.34}$$

где интегрирование производится по ограниченной области $\mathbb{M} \subset \mathbb{R}^n$, которая отображается в $\mathbb{M}'$ при преобразовании (19.30).

Преобразования (19.32), (19.33) нетривиально действуют как на поля, так и на координаты. В дальнейшем нам понадобится вариация формы функции в данной точке $x \in \mathbb{M}$ (см., раздел 3.13)

$$\delta\varphi^a(x) = \varphi'^a(x) - \varphi^a(x),$$

которая определяется разностью значений полей после и до преобразования в точке x. Она связана с вариацией (19.31) следующим соотношением

$$\delta\varphi^a(x) = \bar\delta\varphi^a - \delta x^\alpha \partial_\alpha\varphi^a = \epsilon^{\text{A}}(R_{\text{A}}{}^a - R_{\text{A}}{}^\alpha\partial_\alpha\varphi^a). \tag{19.35}$$

По-построению, для постоянных параметров преобразований вариация формы функции δ перестановочна с операцией частного дифференцирования ∂_α. Вариация действия относительно преобразований (19.30), (19.31) была вычислена ранее (19.23). Если выполнены уравнения Эйлера–Лагранжа, то вариацию действия при постоянных параметрах преобразований симметрии запишем в виде

$$\delta S = -\int_{\mathbb{M}} dx\, \epsilon^{\mathrm{A}} \partial_\alpha J_{\mathrm{A}}{}^{\alpha}, \tag{19.36}$$

где

$$J_{\mathrm{A}}{}^{\alpha} = -\frac{\partial L}{\partial(\partial_\alpha \varphi^a)}(R_{\mathrm{A}}{}^{a} - R_{\mathrm{A}}{}^{\beta}\partial_\beta \varphi^a) - L R_{\mathrm{A}}{}^{\alpha}. \tag{19.37}$$

Совокупность величин $J_{\mathrm{A}}{}^{\alpha}$, $\alpha = 1, \ldots, n$ можно рассматривать, как компоненты некоторого вектора (точнее, векторной плотности) J_{A}, который называется *сохраняющимся током* для каждого преобразования симметрии с параметром ϵ^{A}. Из полученного выражения следует

Теорема 19.2.1 (Первая теорема Нетер). *Если действие (19.1) инвариантно относительно преобразований (19.30)–(19.33) с постоянными параметрами, то для каждого преобразования симметрии и любого решения уравнений Эйлера–Лагранжа токи сохраняются:*

$$\partial_\alpha J_{\mathrm{A}}{}^{\alpha} = 0, \qquad \mathrm{A} = 1, \ldots, \mathrm{K}. \tag{19.38}$$

Заметим, что для сохранения тока достаточно глобальной инвариантности, когда параметр преобразования не зависит от точек пространства-времени.

Поскольку лагранжиан не содержит производных выше первого порядка, то компоненты токов в общем случае зависят только от координат, полей и их первых производных. Симметрия действия (19.1) относительно преобразований (19.30), (19.31) означает, что вариация действия равна нулю для любой области интегрирования даже если сам функционал действия расходится.

Закон сохранения (19.38) не нарушится, если к току (19.37) добавить слагаемое

$$J'_{\mathrm{A}}{}^{\alpha} = J_{\mathrm{A}}{}^{\alpha} + \partial_\beta f_{\mathrm{A}}{}^{\beta\alpha}, \tag{19.39}$$

где $f_{\mathrm{A}}{}^{\beta\alpha} = -f_{\mathrm{A}}{}^{\alpha\beta}$ – произвольная антисимметричная по индексам α, β функция. Чтобы не менять структуры тока (19.37), будем считать, что она зависит только от координат x^α, полей φ и их первых производных $\partial\varphi$. Это преобразование часто используется, чтобы упростить выражения для токов.

Перепишем закон сохранения (19.38) в интегральной форме

$$\int_{\mathbb{M}} dx\, \partial_\alpha J_{\mathrm{A}}{}^{\alpha} = \int_{\partial\mathbb{M}} ds_\alpha\, J_{\mathrm{A}}{}^{\alpha} = 0,$$

где интегрирование ведется по многообразию $\mathbb{M}$ и его границе $\partial\mathbb{M}$. Пусть на $\mathbb{M}$ задана (псевдо-)риманова геометрия, т.е. метрика $g_{\alpha\beta}$ и связность Леви–Чивиты $\widetilde{\Gamma}_{\alpha\beta}{}^{\gamma}$. Тогда, если индекс A не включает в себя пространственно временные индексы α, то первый интеграл по (псевдо-)риманову многообразию, можно переписать в ковариантной форме:

$$\int_{\mathbb{M}} dx\, \partial_\alpha J_{\mathrm{A}}{}^{\alpha} = \int_{\mathbb{M}} dx\, \sqrt{|g|}\widetilde{\nabla}_\alpha \left(\frac{1}{\sqrt{|g|}} J_{\mathrm{A}}{}^{\alpha}\right),$$

где

$$\widetilde{\nabla}_\alpha \left(\frac{1}{\sqrt{|g|}} J_{\mathrm{A}}{}^{\alpha}\right) = \partial_\alpha \left(\frac{1}{\sqrt{|g|}} J_{\mathrm{A}}{}^{\alpha}\right) + \widetilde{\Gamma}_{\alpha\beta}{}^{\alpha} \frac{1}{\sqrt{|g|}} J_{\mathrm{A}}{}^{\beta}$$

– ковариантная производная от вектора тока, и мы воспользовались формулой для дивергенции (7.54).

Рассмотрим действие (19.1) в пространстве Минковского. Пусть $\{x^\alpha\} = \{x^0, \boldsymbol{x}\}$ – декартова система координат, и все поля достаточно быстро убывают на пространственной бесконечности:

$$\lim_{|\boldsymbol{x}|\to\infty} \varphi^a = 0$$

для всех моментов времени. Тогда, интегрируя уравнение (19.38) по области пространства Минковского, которая ограниченна двумя пространственноподобными сечениями $x^0_1 = \mathsf{const}$ и $x^0_2 = \mathsf{const}$, получим закон сохранения

$$Q_{\text{A}} = \int_{\mathbb{S}} d\boldsymbol{x}\, J_{\text{A}}{}^0 = \mathsf{const}, \tag{19.40}$$

где $\mathbb{S}$ – произвольное сечение $x^0 = \mathsf{const}$. Это означает, что каждому преобразованию симметрии соответствует закон сохранения: для любого решения уравнений движения, достаточно быстро убывающего на пространственной бесконечности, интеграл (19.40) не зависит от времени. Интеграл (19.40) называется *сохраняющимся зарядом*, соответствующим току $J_{\text{A}}{}^\alpha$. Если для уравнений движения поставлена задача Коши, то значение заряда Q_{A} однозначно определяется начальными условиями.

19.2.1 Тензор энергии-импульса

Предположим, что некоторая модель описывается набором полей φ^a в пространстве Минковского $\mathbb{R}^{1,n-1}$ с декартовыми координатами x^α. При этом метрика $\eta_{\alpha\beta} = \mathsf{diag}\,(+-\ldots-)$ является заданной функцией в действии, по которой варьирование не проводится. Пусть действие инвариантно относительно трансляций

$$\delta x^\alpha = \epsilon^\alpha = \mathsf{const}, \tag{19.41}$$

$$\bar\delta \varphi^a = 0. \tag{19.42}$$

Для этого достаточно, чтобы лагранжиан модели $L(\varphi, \partial\varphi)$ не зависел явно от координат. Для трансляций индекс A в (19.32) пробегает те же значения, что и α, генератор трансляций совпадает с символом Кронекера, $R_{\text{A}}{}^\alpha \mapsto \delta^\alpha_\beta$, и $R_{\text{A}}{}^a = 0$. В этом случае выражение для тока (19.37) имеет вид

$$T_\alpha{}^\beta = \partial_\alpha\varphi^a \frac{\partial L}{\partial(\partial_\beta\varphi^a)} - \delta^\beta_\alpha L. \tag{19.43}$$

Это выражение называется *тензором энергии-импульса* полей φ^a. Он сохраняется:

$$\partial_\beta T_\alpha{}^\beta = 0. \tag{19.44}$$

Тензор энергии-импульса (19.43) будем называть *каноническим*.

Если лагранжиан модели является скалярным полем (функцией) относительно глобальных преобразований Лоренца $\mathbb{O}(1, n-1)$, то выражение (19.43) представляет собой тензор второго ранга типа $(1,1)$. Ясно, что выражение для $T_0{}^0$ всегда совпадает с плотностью гамильтониана для полей φ^a, и это оправдывает название “канонический”.

Замечание. В общей теории относительности, основанной на псевдоримановой геометрии, постулируется, что тензор Риччи пропорционален тензору энергии-импульса материи. При этом тензор энергии-импульса материи (21.10) определяется, как вариационная производная действия для полей материи по метрике. При таком определении тензор энергии-импульса всегда симметричен. Для скалярного поля вариационная производная действия по метрике является ковариантным обобщением тензора (19.43). В других случаях связь двух определений сложнее и будет обсуждаться в каждом конкретном случае. □

В общем случае тензор энергии-импульса с опущенным верхним индексом $T_{\alpha\beta}$ не является симметричным. Если это так, то в ряде случаев можно провести симметризацию, добавив соответствующую дивергенцию (19.39). Однако, это не всегда возможно. Действительно, после добавления дивергенции получим новый тензор энергии-импульса

$$T'_{\alpha\beta} = T_{\alpha\beta} + \partial^\gamma f_{\alpha\gamma\beta}.$$

Из условия симметрии $T'_{\alpha\beta} - T'_{\beta\alpha} = 0$ следуют уравнения на неизвестную функцию $f_{[\alpha\gamma\beta]} = 0$, в которые входят только полностью антисимметричные компоненты

$$\partial^\gamma f_{[\alpha\gamma\beta]} = -\frac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}).$$

Таким образом мы имеем $n(n-1)/2$ дифференциальных уравнений на $n(n-1)(n-2)/3!$ неизвестных компонент. При $n = 4$ возникает 6 уравнений на 4 неизвестные функции, которые не всегда имеют решения.

Введем стандартные 3-формы на координатных трехмерных гиперповерхностях в четырехмерном пространстве-времени $\mathbb{R}^{1,3}$:

$$ds_\alpha = \frac{1}{6}\varepsilon_{\alpha\beta\gamma\delta} dx^\beta \wedge dx^\gamma \wedge dx^\delta. \tag{19.45}$$

Определим сохраняющийся во времени *ковектор энергии-импульса* с помощью интеграла

$$P_\alpha = \int\limits_{x^0=\,\mathsf{const}} ds_\beta \, T_\alpha{}^\beta, \tag{19.46}$$

где по индексу β производится суммирование. Полученное выражение (19.46), по-построению, является ковектором относительно глобальных лоренцевых вращений. В предположении, что все поля достаточно быстро убывают на пространственной бесконечности ковектор энергии-импульса определяется одним интегралом по пространству,

$$P_\alpha = \int\limits_{x^0=\,\mathsf{const}} d\boldsymbol{x} \, T_\alpha{}^0. \tag{19.47}$$

Выражение для нулевой компоненты P_0 совпадает с гамильтонианом системы полей φ^a, т.е. равно сохраняющейся полной энергии. Это оправдывает название ковектора энергии-импульса. Компоненты тензора энергии-импульса

$$T_i^0 = \frac{\partial L}{\partial(\partial_0\varphi^a)}\partial_i\varphi^a \tag{19.48}$$

определяют сохраняющийся полный импульс системы полей φ^a

$$P_i = \int\limits_{x^0=\,\mathsf{const}} d\boldsymbol{x} \, T_i^0.$$

Полная энергия системы P_0 и каждая компонента полного импульса P_i относительно данной декартовой системы сохраняются во времени. В другой декартовой системе координат они тоже сохраняются, но имеют другие численные значения.

19.2.2 Тензор момента количества движения

Пусть действие $S[\varphi]$ в пространстве Минковского $\mathbb{R}^{1,n-1}$ инвариантно относительно лоренцевых вращений. Мы предполагаем, что набор полей φ^a преобразуется по некоторому, возможно, приводимому представлению группы Лоренца $\mathbb{SO}(1,n-1)$. Обозначим представление генераторов группы для полей через $L_{\gamma\delta b}{}^a = -L_{\delta\gamma b}{}^a$. Тогда в инфинитезимальной форме лоренцевы вращения примут вид

$$\delta x^\alpha = -x^\beta \omega_\beta{}^\alpha = \sum_{\gamma<\delta} \omega^{\gamma\delta}(x_\delta \delta^\alpha_\gamma - x_\gamma \delta^\alpha_\delta), \tag{19.49}$$

$$\delta\varphi^a = \sum_{\gamma<\delta} \omega^{\gamma\delta} L_{\gamma\delta b}{}^a \varphi^b, \tag{19.50}$$

$\omega^{\gamma\delta} = -\omega^{\delta\gamma}$ – параметры преобразований, которые предполагаются постоянными. Для инвариантности действия достаточно, чтобы лагранжиан был скалярным полем (функцией) от координат, полей и их производных. Для лоренцевых вращений индекс $\text{A} \mapsto (\alpha\beta) = -(\beta\alpha)$ представляет собой пару антисимметричных векторных индексов.

Выражение для тока (19.37) приводит к следующему тензору *момента количества движения*, который мы представим в виде суммы двух слагаемых:

$$\begin{aligned} J_{\gamma\delta}{}^\alpha &= -\frac{\partial L}{\partial(\partial_\alpha\varphi^a)}\left(L_{\gamma\delta b}{}^a\varphi^b - (x_\delta\delta^\beta_\gamma - x_\gamma\delta^\beta_\delta)\partial_\beta\varphi^a\right) - L(x_\delta\delta^\alpha_\gamma - x_\gamma\delta^\alpha_\delta) \\ &= M_{\gamma\delta}{}^\alpha + S_{\gamma\delta}{}^\alpha, \end{aligned} \tag{19.51}$$

где введен *орбитальный* и *спиновый* моменты, соответственно,

$$M_{\alpha\beta}{}^\gamma = x_\beta T_\alpha{}^\gamma - x_\alpha T_\beta{}^\gamma, \tag{19.52}$$

$$S_{\alpha\beta}{}^\gamma = -\frac{\partial L}{\partial(\partial_\gamma\varphi^a)} L_{\alpha\beta b}{}^a \varphi^b. \tag{19.53}$$

Оба объекта являются тензорами третьего ранга относительно преобразований Лоренца. Отметим, что орбитальный момент (19.52) не инвариантен относительно трансляций, т.к. явно зависит от координат. В противоположность этому спиновый момент инвариантен относительно трансляций.

Если все поля φ^a являются скалярами относительно лоренцевых вращений, то $L_{\alpha\beta b}{}^a = 0$ и спиновый момент равен нулю, $S_{\alpha\beta}{}^\gamma = 0$.

Допустим, что действие для некоторой системы полей инвариантно относительно трансляций и лоренцевых вращений (группы Пуанкаре), и спиновый момент равен нулю $S_{\alpha\beta}{}^\gamma = 0$, как для скалярных полей. Тогда закон сохранения момента количества движения принимает вид

$$\partial_\gamma M_{\alpha\beta}{}^\gamma = T_{\alpha\beta} - T_{\beta\alpha} + x_\beta \partial_\gamma T_\alpha{}^\gamma - x_\alpha \partial_\gamma T_\beta{}^\gamma.$$

С учетом закона сохранения тензора энергии-импульса (19.44) отсюда вытекает, что для такой системы ковариантный тензор энергии-импульса симметричен:

$$T_{\alpha\beta} = T_{\beta\alpha}. \tag{19.54}$$

Так же, как и для канонического тензора энергии-импульса, для тензора момента количества движения можно ввести полный момент системы. Для полей, достаточно быстро убывающих на пространственной бесконечности, он равен интегралу по пространству:

$$M_{\alpha\beta} = \int\limits_{x^0=\mathrm{const}} d\boldsymbol{x}\, M_{\alpha\beta}{}^{0}. \tag{19.55}$$

Полный момент количества движения является антисимметричным тензором второго ранга относительно преобразований Лоренца.

Замечание. Требование инвариантности моделей математической физики относительно преобразований группы Пуанкаре имеет глубокий физический смысл и составляет основное содержание специальной теории относительности. Инвариантность действия относительно трансляций означает однородность пространства-времени. То есть все точки пространства-времени равноправны, и законы природы имеют одинаковый вид в декартовых координатах с произвольно выбранным началом. Инвариантность относительно преобразований Лоренца означает изотропность пространства-времени. То есть равноправие всех направлений и одинаковый вид законов природы в декартовых координатах с произвольной ориентацией осей. Законы сохранения энергии-импульса и момента количества движения к настоящему времени нашли многочисленные экспериментальные подтверждения в различных областях физики. Поэтому инвариантность фундаментальных моделей математической физики относительно действия группы Пуанкаре следует считать экспериментально установленным фактом.

Помимо этого требование инвариантности функционала действия относительно преобразований группы Пуанкаре в квантовой теории поля означает, что все элементарные частицы должны описываться полями, принадлежащими одному из неприводимых представлений группы Пуанкаре, которые характеризуются массой и спином. Использование этих понятий в экспериментальной физике элементарных частиц также чрезвычайно плодотворно. Это также можно рассматривать, как экспериментальное подтверждение инвариантности законов природы относительно преобразований группы Пуанкаре. □

19.3 Вторая теорема Нетер

Рассмотрим действие (19.1), которое инвариантно относительно преобразований (19.30)–(19.33) с локальными параметрами $\epsilon^{\mathrm{A}}(x)$, зависящим от точек пространства-времени. Мы допускаем, что эти преобразования могут зависеть от частных производных $\partial_\alpha\epsilon^{\mathrm{A}}$ первого и более высокого порядка. Чтобы упростить формулы будем использовать обозначения Девитта [**?**], т.е. суммирование по индексу A в формулах (19.30), (19.31) подразумевает интегрирование, а генераторы локальных преобразований рассматриваются, как двухточечные функции, содержащие δ-функции и (или) их производные.

Пример 19.3.1. Калибровочное преобразование в электродинамике

$$\delta A_\alpha = \partial_\alpha\epsilon$$

будем записывать в виде

$$\delta A_\alpha = \epsilon R_\alpha = \partial_\alpha \int dx'\, \epsilon(x')\delta(x'-x), \tag{19.56}$$

где

$$R_\alpha(x',x)=\frac{\partial}{\partial x^\alpha}\delta(x'-x). \quad \square \tag{19.57}$$

Пример 19.3.2. Бесконечно малые общие преобразования координат для электромагнитного поля (3.99) можно записать в виде

$$\delta x^\alpha=\epsilon^\alpha=\epsilon^\beta R_\beta{}^\alpha, \tag{19.58}$$

$$\delta A_\alpha=-\partial_\alpha\epsilon^\beta A_\beta-\epsilon^\beta\partial_\beta A_\alpha=\epsilon^\beta N_{\beta\alpha}, \tag{19.59}$$

где

$$R_\beta{}^\alpha=\delta^\alpha_\beta\delta(x'-x), \tag{19.60}$$

$$N_{\beta\alpha}=F_{\alpha\beta}(x')\delta(x'-x)-A_\beta(x')\frac{\partial}{\partial x^\alpha}\delta(x'-x). \quad \square \tag{19.61}$$

Определение. Преобразования полей (19.33) с локальными параметрами $\epsilon^{\text{A}}(x)$ называются *калибровочными.* $\square$

Рассмотрим одну из вариационных задач. Будем считать, что параметры ϵ^{A} и их производные равны нулю на границе области. Тогда инвариантность действия относительно калибровочных преобразований можно записать в виде

$$\delta S=\int dx\,\delta\varphi^a S_{,\,a}=\int dx\,\epsilon^{\text{A}}(R_{\text{A}}{}^a-R_{\text{A}}{}^\alpha\partial_\alpha\varphi^a)S_{,\,a}=0. \tag{19.62}$$

При этом были отброшены все граничные слагаемые. Отсюда следует

Теорема 19.3.1 (Вторая теорема Нетер). *Если функционал действия (19.1) инвариантен относительно калибровочных преобразований, которые параметризуются* К *произвольными функциями* $\epsilon^{\text{A}}(x)$, $\text{A}=1,\dots,\text{K}$, *то уравнения Эйлера–Лагранжа удовлетворяют* К *тождествам:*

$$(R_{\text{A}}{}^a-R_{\text{A}}{}^\alpha\partial_\alpha\varphi^a)S_{,\,a}=0. \tag{19.63}$$

Замечание. В формулировке теоремы мы отбросили предположение о том, что параметры преобразований и их производные равны нулю на границе. Если это не так, то зависимость уравнений Эйлера–Лагранжа все равно сохранится. В этом случае из требования инвариантности действия появятся дополнительные следствия для граничных условий, которые мы не рассматриваем. $\square$

Напомним, что в линейном соотношении между уравнениями движения (19.63) суммирование по индексу a предполагает интегрирование. Отсюда следует, что если калибровочные преобразования зависят от частных производных l-того порядка от параметра преобразования, то соотношения (19.63) представляют собой систему K линейных дифференциальных уравнений в частных производных l-того порядка относительно вариационных производных $S_{,\,a}$.

Вторая теорема Нетер утверждает, что в калибровочных моделях, а также моделях, инвариантных относительно общих преобразований координат, не все уравнения движения являются линейно независимыми. Это указывает на то, что в решениях задачи Коши будет содержаться функциональный произвол, т.к. количества уравнений недостаточно для однозначного определения решений по начальным данным.

Для доказательства теоремы существенно, что параметры преобразований $\epsilon^{\text{A}}(x)$ является произвольными функциями, т.к. только в этом случае подынтегральное выражение в (19.62) согласно основной лемме вариационного исчисления должно обращаться в нуль.

Пример 19.3.3. Проведем аналогию с теорией функций многих переменных. Пусть $f = f(x)$ – функция n переменных $x = \{x^\alpha\}$. Аналогом вариационной производной действия в таком случае является обычная частная производная $\partial_\alpha f$. Допустим, что f инвариантна относительно калибровочных преобразований $\delta x^\alpha = \epsilon X^\alpha$, где $\epsilon = \epsilon(x)$ – параметр преобразования и X^α – векторное поле (генератор калибровочного преобразования), которое предполагается отличным от нуля. Тогда зависимость уравнений движения сводится к линейной зависимости частных производных $X^\alpha \partial_\alpha f = 0$. Поэтому функция f постоянна вдоль интегральной кривой $x(t)$ векторного поля X^α:

$$f\Big(x(t)\Big) = \mathsf{const}, \qquad \dot{x}^\alpha = X^\alpha.$$

Это значит, что локальный экстремум $\partial_\alpha f = 0$ достигается не в точке, а на интегральной кривой $x(t)$. □

Из второй теоремы Нетер следует, что функционал действия для калибровочных моделей достигает экстремального значения не на отдельных функциях, а на классах функций, которые связанных между собой калибровочными преобразованиями.

Если некоторая модель инвариантна относительно локальных преобразований, то она, в частности, инвариантна относительно тех же преобразований с постоянными параметрами. Это значит, что токи (19.37) приводят к законам сохранения (19.40) и для локальных преобразований. Поэтому в моделях, инвариантных относительно локальных преобразований можно применить обе теоремы Нетер. При этом первая теорема дает выражения для сохраняющихся токов, а вторая – зависимость уравнений движения. В общем случае это не одно и то же.

Пример 19.3.4. Рассмотрим модели математической физики, инвариантные относительно общих преобразований координат. Пусть действие $S = S(g, \Gamma)$ зависит только от метрики $g_{\alpha\beta}$ и аффинной связности $\Gamma_{\alpha\beta}{}^\gamma$. Обозначим вариационные производные действия следующим образом:

$$\sqrt{|g|} S,^{\alpha\beta} = \frac{\delta S}{\delta g_{\alpha\beta}}, \qquad \sqrt{|g|} S,^{\alpha\beta}{}_\gamma = \frac{\delta S}{\delta \Gamma_{\alpha\beta}{}^\gamma}. \tag{19.64}$$

Здесь мы явно ввели в качестве множителя определитель репера $\sqrt{|g|} = \mathsf{det}\, e_\alpha{}^a$, поскольку вариационные производные так же, как и лагранжиан, являются тензорными плотностями степени -1. Инвариантность действия относительно общих преобразований координат означает равенство нулю вариации

$$\delta S = \int dx \sqrt{|g|} (S,^{\alpha\beta} \delta g_{\alpha\beta} + S,^{\alpha\beta}{}_\gamma \delta \Gamma_{\alpha\beta}{}^\gamma) = 0.$$

Подставляя сюда вариации метрики (3.102) и связности (3.105) и интегрируя по частям (7.62), получим тождества

$$\begin{aligned} &2\widetilde{\nabla}_\alpha S,^\alpha{}_\gamma - \nabla_\beta \nabla_\alpha S,^{\alpha\beta}{}_\gamma + \\ &\quad + \nabla_\beta (T_\alpha + \tfrac{1}{2} Q_\alpha) S,^{\alpha\beta}{}_\gamma + (T_\alpha + \tfrac{1}{2} Q_\alpha) \nabla_\beta (S,^{\alpha\beta}{}_\gamma + S,^{\beta\alpha}{}_\gamma) - \nabla_\alpha S,^{\alpha\beta}{}_\delta T_{\beta\gamma}{}^\delta - \\ &\quad - (T_\alpha + \tfrac{1}{2} Q_\alpha)(T_\beta + \tfrac{1}{2} Q_\beta) S,^{\alpha\beta}{}_\gamma + (T_\alpha + \tfrac{1}{2} Q_\alpha) T_{\beta\gamma}{}^\delta S,^{\alpha\beta}{}_\delta + S,^{\alpha\beta}{}_\delta R_{\alpha\gamma\beta}{}^\delta = 0, \end{aligned} \tag{19.65}$$

где $\widetilde{\nabla}_\alpha$ и ∇_α – ковариантные производные соответственно со связностью Леви–Чивиты $\widetilde{\Gamma}_{\alpha\beta}{}^\gamma$ и аффинной связностью $\Gamma_{\alpha\beta}{}^\gamma$, а подъем и опускание индексов производится с

помощью метрики $g_{\alpha\beta}$. Таким образом, в моделях, инвариантных относительно общих преобразований координат, уравнения движения удовлетворяют n линейным дифференциальным тождествам.

В (псевдо-)римановой геометрии, когда гравитационная часть действия зависит только от метрики, эти тождества значительно упрощаются:

$$\widetilde{\nabla}_\alpha S,^\alpha{}_\beta = 0. \tag{19.66}$$

В общей теории относительности для действия Гильберта–Эйнштейна имеем

$$\widetilde{\nabla}_\alpha \left(\widetilde{R}^\alpha{}_\beta - \frac{1}{2}\delta^\alpha_\beta \widetilde{R} \right) = 0. \tag{19.67}$$

Это тождество совпадает со свернутыми тождествами Бианки (7.108). □

Пример 19.3.5. Модели гравитации, построенные в рамках геометрии Римана–Картана, в переменных Картана инвариантны относительно общих преобразований координат и локальных преобразований Лоренца. Как следствие не все уравнения движения являются независимыми, поскольку удовлетворяют тождествам. Пусть действие $S = S(e, \omega)$ зависит только от репера $e_\alpha{}^a$ и лоренцевой связности $\omega_\alpha{}^{ab}$. Обозначим вариационные производные следующим образом:

$$\sqrt{|g|} S,^\alpha{}_a = \frac{\delta S}{\delta e_\alpha{}^a}, \qquad \sqrt{|g|} S,^\alpha{}_{ab} = \frac{\delta S}{\delta \omega_\alpha{}^{ab}}. \tag{19.68}$$

Используя формулы для вариаций репера (**??**) и лоренцевой связности (**??**) при локальных лоренцевых вращениях, получим следующую зависимость уравнений движения

$$\widetilde{\nabla}_\alpha S,^\alpha{}_{ab} + \frac{1}{2}(S,_{ab} - S,_{ba}) = 0, \tag{19.69}$$

где переход от греческих индексов к латинским осуществляется с помощью репера. Полученная зависимость соответствует инвариантности действия относительно локальных лоренцевых вращений.

Вариации полей e и ω при общих преобразованиях координат имеют вид (3.111) и (3.113). Поэтому инвариантность действия относительно общих преобразований координат приводит к тождествам:

$$\nabla_\alpha S,^\alpha{}_\beta + T_\alpha S,^\alpha{}_\beta + S,^\alpha{}_a T_{\alpha\beta}{}^a + S,^\alpha{}_{ab} R_{\alpha\beta}{}^{ab} = 0, \tag{19.70}$$

где мы учли полученное ранее тождество (19.69).

При добавлении к гравитационному действию слагаемых, зависящих от других полей (полей материи), тождества, которым удовлетворяют уравнения движения меняются, т.к. необходимо учитывать вариации всех полей. □

19.4 Эффективное действие

При исследовании моделей математической физики, действие которых зависит от нескольких полей, иногда удается решить часть уравнений Эйлера–Лагранжа явно. В этом случае вариационную задачу можно свести к новому эффективному действию, зависящему от меньшего числа переменных. В настоящем разделе мы докажем простую теорему, позволяющую строить эффективное действие в случае вариационной

задачи с фиксированными граничными условиями. То есть будем пренебрегать всеми граничными слагаемыми. Обобщение на более сложные случаи будет ясно из дальнейшего рассмотрения.

Начнем с простейшего случая. Пусть на ограниченной области $\mathbb{M} \subset \mathbb{R}^n$ заданы два скалярных поля $\varphi(x)$ и $\psi(x)$, и пусть функция $\psi = \psi(x, \varphi, \partial\varphi, \partial^2\varphi)$ в каждой точке $x \in \mathbb{M}$ задана как функция φ, ее первых и вторых частных производных: $\partial_\alpha\varphi$ и $\partial_\alpha\partial_\beta\varphi$. Представим значение функции $\psi(x) := \psi\big[x, \varphi(x), \partial\varphi(x), \partial^2\varphi(x)\big]$ в точке $x \in \mathbb{M}$ в виде функционала, используя δ-функцию,

$$\psi(x) = \int_{\mathbb{M}} dy\, \psi(y)\delta(y - x).$$

Вариация функционала $\psi(x)$, вызванная вариацией $\delta\varphi$, имеет вид

$$\delta\psi(x) = \int_{\mathbb{M}} dy \left(\frac{\partial\psi}{\partial\varphi}\delta\varphi + \frac{\partial\psi}{\partial(\partial_\alpha\varphi)}\delta(\partial_\alpha\varphi) + \frac{\partial\psi}{\partial(\partial_\alpha\partial_\beta\varphi)}\delta(\partial_\alpha\partial_\beta\varphi) \right) \delta(y - x).$$

Проинтегрировав второе и третье слагаемые по частям, получим выражение для вариационной производной

$$\frac{\delta\psi(x)}{\delta\varphi(y)} = \frac{\partial\psi}{\partial\varphi}\delta(y - x) - \frac{\partial}{\partial y^\alpha}\left(\frac{\partial\psi}{\partial(\partial_\alpha\varphi)}\delta(y - x) \right) + \partial_\alpha\partial_\beta\left(\frac{\partial\psi}{\partial(\partial_\alpha\partial_\beta\psi)}\delta(y - x) \right), \quad (19.71)$$

где в правой части $\psi = \psi(y)$ и $\varphi = \varphi(y)$.

Теперь обсудим вариационную задачу. Пусть действие $S[\varphi, \psi]$ зависит от двух функций φ и ψ. Тогда из принципа наименьшего действия следуют два уравнения Эйлера–Лагранжа:

$$\frac{\delta S}{\delta\varphi} = 0, \quad (19.72)$$

$$\frac{\delta S}{\delta\psi} = 0. \quad (19.73)$$

Допустим, что второе уравнение Эйлера–Лагранжа допускает общее решение для ψ, как функции от φ и ее производных:

$$\psi = \psi(x, \varphi, \partial\varphi, \partial^2\varphi). \quad (19.74)$$

При этом мы предполагаем, что общее решение не имеет особенностей. Если действие зависит только от самих функций и их первых производных, то в общее решение будут входить производные от φ не выше второго порядка. Поскольку уравнение Эйлера–Лагранжа (19.73) является дифференциальным уравнением в частных производных, то общее решение зависит также от некоторого набора произвольных функций и постоянных. Часть этих произвольных функций и постоянных фиксируется, если это возможно, граничными условиями $\psi|_{\partial\mathbb{M}} = \psi_0$ и $\varphi|_{\partial\mathbb{M}} = \varphi_0$. Используем полученное решение для построения нового *эффективного* действия

$$S_{\text{eff}}[\varphi] := S[\varphi, \psi(\varphi)], \quad (19.75)$$

которое зависит только от одной функции φ. Тогда уравнение Эйлера–Лагранжа для φ связано со старыми уравнениями (19.72), (19.73) простым соотношением

$$\frac{\delta S_{\text{eff}}}{\delta\varphi(x)} = \frac{\delta S}{\delta\varphi(x)} + \frac{\delta S}{\delta\psi(y)}\frac{\delta\psi(y)}{\delta\varphi(x)}\bigg|_{\psi=\psi(\varphi)} = 0, \quad (19.76)$$

где во втором слагаемом подразумевается интегрирование по аргументу поля $\psi(y)$, которое снимается δ-функцией в вариационной производной. Ясно, что второе слагаемое равно нулю, если выполнено уравнение Эйлера–Лагранжа для ψ (19.73).

Проведенные вычисления остаются в силе и в том случае, когда мы рассматриваем наборы полей $\varphi = \{\varphi^a\}$, $a = 1, \ldots, \mathrm{N}$ и ψ^{A}, $\mathrm{A} = 1, \ldots, m$. Отсюда следует

Теорема 19.4.1. *Пусть дано действие $S[\varphi, \psi]$, зависящее от двух наборов полей φ и ψ. Тогда множество экстремалей этого действия для вариационной задачи с заданными граничными условиями совпадает с множеством экстремалей эффективного действия (19.75), дополненным выражением ψ через φ (19.74).*

Замечание. Если рассматриваются другие вариационные задачи, то слова "множество экстремалей" нужно заменить на "множество решений уравнений Эйлера–Лагранжа". □

Эта теорема важна, поскольку позволяет строить эффективное действие, которое зависит от меньшего числа полей, подставляя решение части системы уравнений Эйлера–Лагранжа непосредственно в исходное действие.

При доказательстве теоремы 19.4.1 мы предположили, что всеми граничными слагаемыми можно пренебречь. В полевых моделях математической физики это не всегда так. В разделе 20.2.6 будет построен пример, где подстановка решения части уравнений Эйлера–Лагранжа в действие не воспроизводит оставшиеся уравнения движения. Это связано с нетривиальной ролью граничных слагаемых в действие для полевых моделей.

19.5 Редуцированное действие

В настоящем разделе мы рассмотрим еще один способ сведения сложной вариационной задачи к более простой. Пусть задано действие $S[\varphi]$. Рассмотрим для него вариационную задачу с фиксированными граничными условиями. Как правило, уравнения Эйлера–Лагранжа (уравнения движения) настолько сложны, что не позволяют найти все решения. В таких случаях для нахождения частных решений делают упрощающие предположения: решение уравнений движения ищется в определенном классе функций. Например, ищутся статические или сферически симметричные решения. Чтобы найти стационарную точку действия, упрощающую подстановку, которую часто называют *анзац* (от немецкого ansatz $\simeq$ исходное математическое выражение), следует производить в уравнения Эйлера–Лагранжа (19.11), а не в действие. В этом случае найденное точное решение уравнений движения действительно будет стационарной точкой исходного действия.

Однако в ряде случаев подстановки можно производить непосредственно в действие. Это означает следующее. Пусть в результате упрощающих предположений исходный набор функций φ^a, $a = 1, \ldots, \mathrm{N}$, будет выражен через меньшее число независимых функций ψ^{A}, $\mathrm{A} = 1, \ldots, m$, и координаты

$$\varphi^a = \varphi^a(x, \psi). \tag{19.77}$$

При этом функции ψ могут зависеть от меньшего числа координат. В результате подстановки будет получено новое *редуцированное действие*

$$S_{\mathrm{red}}[\psi] := S[\varphi(\psi)], \tag{19.78}$$

зависящее от меньшего числа независимых полей.

Допустим, что найдено решение уравнений движения для редуцированного действия

$$\frac{\delta S_{\rm red}}{\delta \psi^{\rm A}} = 0.$$

Тогда функции (19.77), как правило, не будут удовлетворять исходным уравнениям (19.11). Тем не менее в ряде случаев исходные уравнения все же будут удовлетворены. Это замечательные случаи, которые позволяют существенно упростить вычисления. Вместе с этим наличие редуцированного действия помогает в анализе свойств рассматриваемой модели.

Опишем достаточные условия для возможности подстановки (19.77) непосредственно в действие. Пусть на многообразии $\mathbb{M}$ действует группа Ли преобразований $\mathbb{G}$ справа:

$$\mathbb{M} \times \mathbb{G} \ni x, g \mapsto xg \in \mathbb{M}.$$

Предположим, что набор полей $\varphi^a(x)$ при этом преобразуются по некоторому представлению $T(g)_a{}^b$ группы Ли $\mathbb{G}$:

$$\varphi'^a(xg) = \varphi^b(x) T(g)_b{}^a, \tag{19.79}$$

где штрихом обозначены новые поля на $\mathbb{M}$, полученные в результате действия группы $\mathbb{G}$. Поля $\varphi^a(x)$ называются $\mathbb{G}$-*инвариантными*, если в равенстве (19.79) можно убрать штрих. То есть выполнено условие

$$\varphi^a(xg) = \varphi^b(x) T(g)_b{}^a,$$

или

$$\varphi^a(x) = \varphi^b(xg^{-1}) T(g)_b{}^a. \tag{19.80}$$

Эти условия представляют собой уравнения, определяющие $\mathbb{G}$-инвариантные поля на многообразии $\mathbb{M}$.

Обозначим множество всех $\mathbb{G}$-инвариантных полей на $\mathbb{M}$ через Φ. Пусть это множество параметризуется некоторым набором функций $\psi^{\rm A}$, которые могут зависеть от меньшего числа координат и не определяются из уравнений (19.80). В результате $\mathbb{G}$-инвариантные функции будут представлены в виде (19.77).

Теорема 19.5.1 (**Принцип Коулмена**). *Пусть исходный функционал действия $S[\varphi]$ инвариантен относительно действия группы Ли преобразований (19.79),*

$$S[\varphi'] = S[\varphi].$$

Тогда поля (19.77), построенные для стационарных точек редуцированного действия (19.78), будут удовлетворять уравнениям Эйлера–Лагранжа исходного действия.

Это утверждение известно, как *принцип Коулмена*. Оно было высказано Коулманом и проиллюстрировано на нескольких примерах [**?**] (см. также [**?**]). Строгое доказательство вместе с ограничениями на его применимость было дано в статье [**?**]. Этот результат затем был обобщен в работах [**?**, **?**]. Другая его формулировка, на языке теории стратов была дана еще до Коулмена [**?**]. В настоящее время доказано, что принцип Коулмана справедлив для всех компактных групп преобразований, полупростых групп, а также для унитарных представлений некомпактных групп.

Глава 20

Канонический формализм

Трудно переоценить роль канонического (гамильтонова) формализма в классической и квантовой механике, а также в теории поля. Он предоставляет наиболее мощные методы интегрирования уравнений движения и является основой для канонического квантования различных моделей математической физики. К его недостаткам относится явное нарушение Лоренц-инвариантности моделей теории поля, поскольку время в гамильтоновом формализме играет выделенную роль. Это усложняет вычисления, проводимые в рамках теории возмущений. Однако принципиальные вопросы, связанные с физической интерпретацией математических моделей, невозможно решить без обращения к гамильтоновой формулировке. В настоящей главе рассматривается канонический формализм для системы точечных частиц, формализм Дирака для систем со связями и его обобщение на теорию поля.

20.1 Канонический формализм в механике точечных частиц

20.1.1 Преобразование Лежандра

Рассмотрим *выпуклую функцию* $y = f(x)$ на интервале $-\infty \le a < x < b \le \infty$, т.е. функцию, у которой $f''(x) > 0$ при $x \in (a, b)$. Преобразованием Лежандра функции f на интервале (a, b) называется новая функция $g(p)$ нового переменного p, которая строится следующим образом (см. рис. 20.1). Нарисуем на плоскости x, y график функции f. Рассмотрим прямую $y = px$, где p – фиксированное число. Найдем точку $x(p)$, в которой кривая дальше всего от прямой по вертикали, т.е. функция $F(x, p) := px - f(x)$ имеет максимум по x при фиксированном p. Тогда, по-определению, $g(p) := F\Big(x(p), p\Big)$. Функция $g(p)$ определена на некотором интервале $-\infty \le c < p < d \le \infty$.

Точка $x(p)$ определяется из условия экстремума $\partial F/\partial x = 0$ или

$$p = f'(x)$$

Ввиду выпуклости f, если это уравнение имеет решение, то оно единственно.

Пример 20.1.1. Нетрудно проверить, что функция $f(x) = \frac{mx^2}{2}$, $m = \mathsf{const} > 0$ выпукла на всей вещественной оси $x \in \mathbb{R}$ и $p = mx$. Ее преобразование Лежандра определено для всех $p \in \mathbb{R}$ и имеет вид

$$g(p) = (xp - f)\big|_{x=p/m} = \frac{p^2}{2m} \qquad \square$$

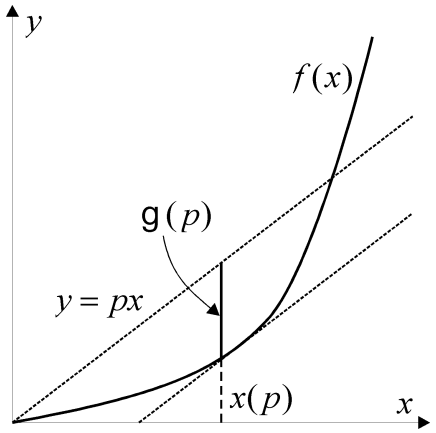


Рис. 20.1: Преобразование Лежандра $g(p)$ выпуклой функции $f(x)$.

Пусть функция f достаточно гладкая. Тогда преобразование Лежандра переводит выпуклые функции в выпуклые. Это значит, что преобразование Лежандра можно применить дважды. Можно доказать [?], что преобразование Лежандра *инволютивно*, т.е. его квадрат равен тождественному преобразованию.

По-определению $F(x,p) = px - f(x) \leq g(p)$ для всех x и p. Отсюда вытекает *неравенство Юнга*

$$px \leq f(x) + g(p)$$

Преобразование Лежандра без труда обобщается на функции нескольких переменных. Пусть $f(\boldsymbol{x})$ – выпуклая функция нескольких переменных $\boldsymbol{x} = (x^1, \ldots, x^n) \in \mathbb{R}^n$, т.е. квадратичная форма $dx^\alpha dx^\beta \partial^2_{\alpha\beta} f$, $\alpha = 1, \ldots, n$, положительно определена в некоторой окрестности $\mathbb{U} \subset \mathbb{R}^n$. Тогда преобразованием Лежандра называется функция $g(\boldsymbol{p})$ того же числа переменных $\boldsymbol{p} = (p_1, \ldots, p_n)$, которая строится аналогично случаю одного переменного:

$$g(\boldsymbol{p}) = F\Big(\boldsymbol{x}(\boldsymbol{p}), \boldsymbol{p}\Big) = \max_{\boldsymbol{x}} F(\boldsymbol{x}, \boldsymbol{p}),$$

где

$$F(\boldsymbol{x}, \boldsymbol{p}) = x^\alpha p_\alpha - f(\boldsymbol{x}), \quad \text{и} \quad p_\alpha = \frac{\partial f}{\partial x^\alpha}.$$

Преобразование Лежандра $g(\boldsymbol{p})$ определено в некоторой окрестности $\boldsymbol{p} \in \mathbb{V} \subset \mathbb{R}^n$, которая определяется исходной функцией $f(\boldsymbol{x})$.

В этом определении мы различаем верхние и нижние индексы по следующей причине. Поскольку $\boldsymbol{x}$ – точка многообразия $\mathbb{R}^n$ и f – функция на $\mathbb{R}^n$, то набор переменных p_α определяет компоненты некоторого ковектора (1-формы). Отметим также, что в этом определении все координаты x^α равноправны.

Пример 20.1.2. Преобразованием Лежандра положительно определенной квадратичной формы

$$f(\boldsymbol{x}) = \frac{1}{2} g_{\alpha\beta} x^\alpha x^\beta,$$

где $g_{\alpha\beta} = g_{\beta\alpha}$ – постоянная матрица, снова является положительно определенная квадратичная форма

$$g = \Big(x^\alpha p_\alpha - f(\boldsymbol{x})\Big)\Big|_{x^\alpha = g^{\alpha\beta} p_\beta} = \frac{1}{2} g^{\alpha\beta} p_\alpha p_\beta.$$

где $g^{\alpha\beta}$ – матрица, обратная к $g_{\alpha\beta}$. При этом значения обеих форм в соответствующих точках совпадают:

$$f\Big(\boldsymbol{x}(\boldsymbol{p})\Big) = g(\boldsymbol{p}), \qquad g\Big(\boldsymbol{p}(\boldsymbol{x})\Big) = f(\boldsymbol{x}).$$

Обе функции $f(\boldsymbol{x})$ и $g(\boldsymbol{p})$ определены на всем евклидовом пространстве $\mathbb{R}^n$. □

20.1.2 Гамильтонова динамика точечных частиц

Механика точечных частиц может быть описана на двух эквивалентных языках: лагранжевом и гамильтоновом. Каждый подход имеет свои преимущества и недостатки. В лагранжевом подходе совокупность N частиц описывается обобщенными координатами $q^i(t)$, $i = 1, \ldots, \mathrm{N}$, зависящих от времени t. Если каждая частица движется в трехмерном пространстве, то q^i представляет собой трехмерный вектор в евклидовом пространстве $\mathbb{R}^3$ для каждого значения индекса i. Для определенности мы считаем, что каждая частица движется в одномерном пространстве. Тогда совокупность всех координат q^i можно рассматривать, как координаты одной частицы в *конфигурационном* пространстве $\mathbb{R}^{\mathrm{N}}$. Размерность конфигурационного пространства называется *числом степеней свободы* механической системы. Говорят, что механическая система имеет N степеней свободы.

Для простоты обозначений условимся считать, что символ с индексом q^i обозначает i-ю координату частицы, а по повторяющимся индексам производится суммирование. Если индекс отсутствует, то символ q обозначает весь набор координат $q := (q^1, \ldots, q^{\mathrm{N}})$. В лагранжевом подходе уравнения движения имеют второй порядок, поэтому будем рассматривать дважды непрерывно дифференцируемые функции $q^i \in \mathcal{C}^2\Big([t_1, t_2]\Big)$ для всех i на конечном отрезке $[t_1, t_2]$.

Предположим, что механическая система описывается некоторым действием

$$S[q] = \int_{t_1}^{t_2} dt\, L(q, \dot{q}, t), \tag{20.1}$$

где функция Лагранжа или лагранжиан $L(q, \dot{q})$ зависит только от обобщенных координат q и их первых производных по времени $\dot{q}$, которые называются обобщенными *скоростями*. Рассмотрим вариационную задачу с фиксированными граничными условиями

$$q(t_1) = q_1, \qquad q(t_2) = q_2, \tag{20.2}$$

т.е. траектория механической системы представляет собой кривую в конфигурационном пространстве, которая соединяет две фиксированные точки q_1 и q_2. Тем самым вариации координат на концах интервала обращаются в нуль.

Обычно предполагают, что конфигурационное пространство представляет собой евклидово пространство $\mathbb{R}^{\mathrm{N}}$ с заданной метрикой δ_{ij}, а обобщенные координаты – это декартовы координаты в $\mathbb{R}^{\mathrm{N}}$. Кроме того, мы предполагаем, что лагранжиан в действии (20.1) представляет собой функцию (скалярное поле) на $\mathbb{R}^{\mathrm{N}}$. Поэтому метрика необходима для построения инвариантов из координат q^i и скоростей $\dot{q}^i$, и она входит в действие. В этом действии евклидова метрика рассматривается как внешнее поле, и по ней варьирование не проводится. Конечно, после того, как задача поставлена, действие можно переписать в произвольной криволинейной системе координат в $\mathbb{R}^{\mathrm{N}}$. Тогда обобщенные координаты q^i станут криволинейными координатами в $\mathbb{R}^{\mathrm{N}}$, а компоненты евклидовой метрики $g_{ij}(x)$ будут нетривиальными функциями от x. Действие по этим компонентам не варьируется.

Замечание. При рассмотрении движения механической системы, естественная линейная структура в $\mathbb{R}^N$ не играет никакой роли и нигде не используется. Важно только наличие метрики. В общем случае можно считать, что конфигурационное пространство – это произвольное риманово многообразие $(\mathbb{M}, g)$ с заданной метрикой g, которая необходима для построения инвариантов. Действие в этом случае будет зависеть от метрики, которая рассматривается как внешнее поле, и по ней варьирование не проводится. Заметим также, что задание лагранжиана ничего не говорит о топологии конфигурационного пространства. Для того, чтобы задать топологию $\mathbb{M}$ необходимо сделать какие либо дополнительные предположения. В дальнейшем, для простоты, мы будем рассматривать в основном топологически тривиальные конфигурационные пространства, $\mathbb{M} \approx \mathbb{R}^N$. □

Уравнения Эйлера–Лагранжа (уравнения движения) для действия (20.1),

$$\frac{\delta S}{\delta q^i} = \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0, \tag{20.3}$$

представляют собой систему уравнений не выше второго порядка. Функции $q^i(t)$, удовлетворяющие уравнениям Эйлера–Лагранжа, соответствуют стационарным точкам действия (20.1). Они определяют *траекторию* механической системы. Траектории частиц не зависят от выбора координат в конфигурационном пространстве $\mathbb{R}^N$, которые выбираются из соображений удобства.

Как правило, решение системы уравнений движения (20.3) содержит $2N$ произвольных постоянных, которые находятся из граничных условий (20.2). У краевой задачи решения может не существовать, а если оно существует, то может не быть единственным. Вместо краевой задачи можно поставить задачу Коши, которая имеет единственное решение при корректной постановке. В этом случае произвольные постоянные, возникающие в решении уравнений движения, находятся из начальных данных для q и $\dot{q}$ в начальный момент времени $t = 0$. Именно эта задача наиболее часто ставится в физических приложениях для уравнений движения (20.3).

Перейдем к гамильтонову или каноническому формализму. В этом случае механическая система, состоящая из N частиц, описывается N обобщенными координатами q^i и N обобщенными импульсами

$$p_i := \frac{\partial L}{\partial \dot{q}^i}, \tag{20.4}$$

которые рассматриваются, как независимые переменные. Обобщенные координаты и импульсы являются координатами механической системы в $2N$-мерном *фазовом* пространстве. Мы говорим, что координаты и импульсы являются канонически сопряженными переменными.

Предположим, что конфигурационное пространство с координатами q^i является N-мерным многообразием $\mathbb{M}$. Тогда производные $\dot{q}^i$ являются компонентами касательного вектора вдоль траектории частицы. Это значит, что лагранжиан $L(q, \dot{q})$ является функцией (скалярным полем) на касательном расслоении $\mathbb{T}(\mathbb{M})$. Тогда формула (20.4) определяет компоненты ковариантного вектора (1-формы). Это значит, что фазовое пространство есть ни что иное, как кокасательное расслоение $\mathbb{T}^*(\mathbb{M})$ к конфигурационному пространству $\mathbb{M}$.

Поскольку q^i и p_i являются координатами в базе и в кокасательном слое, то мы пишем координатные индексы, соответственно, вверху и внизу.

Предположим, что функция Лагранжа $L(q, \dot{q})$ выпукла по скоростям, т.е. квадратичная форма

$$\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j, \tag{20.5}$$

которая называется *гессианом*, положительно определена. Это условие не является ковариантным относительно произвольной замены координат на касательном расслоении $\mathbb{T}(\mathbb{M})$. Однако оно ковариантно относительно общих преобразований координат в конфигурационном пространстве $\mathbb{M}$. Действительно, при замене координат $q \mapsto q'(q)$ скорости преобразуются как векторы:

$$\dot{q}^i \mapsto \dot{q}'^i = \frac{\partial q'^i}{\partial q^j}\dot{q}^j.$$

Поскольку якобиан преобразования $\partial q'^i/\partial q^j$ не зависит от скоростей, то гессиан преобразуется как ковариантный тензор второго ранга:

$$\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \mapsto \frac{\partial^2 L}{\partial \dot{q}'^i \partial \dot{q}'^j} = \frac{\partial q^k}{\partial q'^i}\frac{\partial}{\partial \dot{q}^k}\left(\frac{\partial q^l}{\partial q'^j}\frac{\partial L}{\partial \dot{q}^l}\right) = \frac{\partial q^k}{\partial q'^i}\frac{\partial q^l}{\partial q'^j}\frac{\partial^2 L}{\partial \dot{q}^k \partial \dot{q}^l}.$$

Таким образом, понятие выпуклости лагранжиана по скоростям не зависит от выбора координат в конфигурационном пространстве $\mathbb{M}$.

Замечание. Рассмотрим координаты q и время t в определении обобщенных импульсов (20.4) как параметры. Тогда формулы (20.4) задают скорости $\dot{q}$ как неявные функции обобщенных импульсов p. Из курса математического анализа известно, что уравнения (20.4) локально разрешимы относительно скоростей тогда и только тогда, когда гессиан (20.5), который в данном случае совпадает с якобианом преобразования координат $\dot{q} \mapsto p$, является невырожденной матрицей. □

Если функция Лагранжа выпукла по скоростям, то можно построить ее преобразование Лежандра (по скоростям):

$$H(q,p) := p_i\dot{q}^i - L(q,\dot{q}), \tag{20.6}$$

которое называется *функцией Гамильтона* или *гамильтонианом* системы. В силу определения преобразования Лежандра гамильтониан системы зависит только от обобщенных координат и импульсов. Это можно проверить и непосредственно

$$\frac{\partial H}{\partial \dot{q}^i} = p_i - \frac{\partial L}{\partial \dot{q}^i} = 0,$$

что следует из определения обобщенных импульсов (20.4).

Используя связь между гамильтонианом и лагранжианом (20.6), рассмотрим действие

$$S[q,p] = \int_{t_1}^{t_2} dt\,\Big(p_i\dot{q}^i - H(q,p)\Big) \tag{20.7}$$

как функционал от канонически сопряженных координат и импульсов. Соответствующая система уравнений Эйлера–Лагранжа имеет вид

$$\frac{\delta S}{\delta q^i} = -\dot{p}_i - \frac{\partial H}{\partial q^i} = 0, \tag{20.8}$$

$$\frac{\delta S}{\delta p_i} = \quad \dot{q}^i - \frac{\partial H}{\partial p_i} = 0. \tag{20.9}$$

Нетрудно проверить, что N уравнений Эйлера–Лагранжа второго порядка (20.3) эквивалентны 2N уравнениям первого порядка (20.8) (предложение **??**. Понижение порядка уравнений движения произошло за счет введения новых независимых переменных – импульсов.

Поскольку конфигурационное и фазовое пространства имеют разную размерность, то уточним понятие эквивалентности. Для однозначного определения траектории частицы $q(t)$ в конфигурационном пространстве необходимо решить уравнения Эйлера–Лагранжа второго порядка (20.3), например, с начальными условиями $q(0) = q_0$, $\dot{q}(0) = q_1$. Траектория движения в фазовом пространстве $q(t), p(t)$ однозначно находится решением уравнений движения первого порядка (20.8), (20.9) с начальными условиями $q(0) = q_0$, $p(0) = p_0$. Каждая траектория в фазовом пространстве естественным образом проектируется на траекторию в конфигурационном пространстве $\{q(t), p(t)\} \mapsto \{q(t)\}$, при этом начальное условие для скорости q_1 определяется из уравнения (20.9). Обратно. Для любой траектории в конфигурационном пространстве $q(t)$ уравнение (20.9) определяет импульсы $p(t)$ и начальные условия p_0 так, что пара $q(t), p(t)$ является траекторией в фазовом пространстве. Здесь мы предполагаем, что уравнение (20.9) имеет единственное решение для p (в противном случае преобразование Лежандра не определено).

При постановке граничной задачи для действия (20.7) мы не можем рассматривать траектории, соединяющие две произвольные точки (q_1, p_1) и (q_2, p_2) фазового пространства, т.к. возникает $4\mathrm{N}$ условий для $2\mathrm{N}$ обыкновенных дифференциальных уравнений первого порядка (20.8), (20.9). В этом случае количество граничных условий превышает количество уравнений, и задача может не иметь решения. Поэтому можно рассматривать, например, те траектории, которые имеют начало и конец на N-мерных подмногообразиях фазового пространства, определяемых условиями (20.2), как и в лагранжевом подходе. При этом никаких граничных условий на импульсы не возникает, поскольку они входят в подынтегральное выражение без производных, и при вариации импульсов интегрирования по частям не происходит.

Замечание. Как правило для уравнений Эйлера–Лагранжа в фазовом пространстве ставится задача Коши, которая имеет единственное решение. При этом начальной точкой траектории может быть произвольная точка фазового пространства $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) = \mathbb{R}^{2\mathrm{N}}$. Тем самым мы предполагаем, что фазовое пространство так же как и конфигурационное пространство топологически тривиально. Обсуждение области определения различных функций в конфигурационном и фазовом пространствах требует знания явного вида лагранжиана, гамильтониана и анализа уравнений движения. Поскольку в общем случае учесть все возможности нельзя, то в дальнейшем при обсуждении общей схемы все рассматриваемые функции будут считаться определенными на всем фазовом пространстве и достаточно гладкими. В каждом конкретном случае области определения должны быть проанализированы, а постановка задачи уточнена. □

Посмотрим на уравнения (20.8), (20.9) с другой точки зрения. Пусть заданы две функции от канонических переменных $f(q,p)$ и $g(q,p)$, т.е. два скалярных поля на фазовом пространстве. Определим для них *скобку Пуассона*

$$[f,g] = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}. \tag{20.10}$$

Легко проверить, что она обладает следующими свойствами:

1) $[af + bg, h] = a[f,g] + b[g,h], \quad a,b \in \mathbb{R}$ – линейность,
2) $[f,g] = -[g,f]$ – антисимметрия,
3) $[f,gh] = [f,g]h + g[f,h]$ – правило Лейбница,
4) $[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0$ – тождество Якоби,

и, следовательно, определяет пуассонову структуру на фазовом пространстве (см. раздел 18.3.

Скобку Пуассона (20.10) можно переписать в эквивалентном виде. Обозначим координаты фазового пространства через $\{x^\alpha\} = \{q^1 \dots q^{\mathrm{N}}, p_1 \dots p_{\mathrm{N}}\}$, $\alpha = 1, \dots, 2\mathrm{N}$, или, короче, $x = (q, p)$. Тогда

$$[f, g] = \varpi^{-1\alpha\beta} \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\beta}, \tag{20.11}$$

где

$$\varpi^{-1\alpha\beta} = [x^\alpha, x^\beta] = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

– каноническая пуассонова структура. Каноническая форма ϖ невырождена и точна $d\varpi = 0$, поскольку ее компоненты постоянны. Поэтому фазовое пространство представляет собой также симплектическое многообразие (см. раздел 18.2).

Рассмотрим простейшие свойства скобки Пуассона. Очевидно, скобки Пуассона произвольной функции f с самой собой и константой c равны нулю:

$$[f, f] = 0, \qquad [f, c] = 0.$$

Из определения (20.10) следует, что скобка Пуассона обобщенной координаты с импульсом равна символу Кронекера

$$[q^i, p_j] = \delta^i_j. \tag{20.12}$$

Рассмотрим траекторию в фазовом пространстве $q(t)$, $p(t)$. Тогда, используя скобку Пуассона, уравнения движения (20.8), (20.9) можно переписать в виде

$$\begin{aligned} \dot{q}^i &= [q^i, H] = \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= [p_i, H] = -\frac{\partial H}{\partial q^i}. \end{aligned} \tag{20.13}$$

При этом скобка Пуассона для координат фазового пространства (20.12) рассматривается как одновременная:

$$[q^i(t), p_j(t)] = \delta^i_j, \qquad \forall t \in \mathbb{R}. \tag{20.14}$$

Скобка Пуассона для различных моментов времени $[q^i(t_1), p_j(t_2)]$ при $t_1 \neq t_2$ не определена. Вообще, эволюция во времени любой функции $f(t, q, p)$, зависящей от времени и канонических переменных определяется уравнением

$$\dot{f} = \frac{df}{dt} = \frac{\partial f}{\partial t} + [f, H]. \tag{20.15}$$

В частности, если функция Гамильтона не зависит от времени явно, то

$$\frac{dH}{dt} = [H, H] = 0,$$

ввиду антисимметрии скобки Пуассона. Это значит, что для заданной траектории механической системы в фазовом пространстве гамильтониан является интегралом движения и его численное значение сохраняется. Это значение называется *энергией* механической системы и определяется начальными данными. Системы, у которых гамильтониан не зависит от времени, называются *консервативными*.

Пример 20.1.3 (Гармонический осциллятор с затуханием). Рассмотрим функцию Лагранжа для одной точечной частицы (осциллятора), которая явно зависит от времени

$$L = \frac{1}{2}\,\mathsf{e}^{2\mu t}(\dot{q}^2 - \omega^2 q^2), \qquad \mu, \omega = \mathsf{const}.$$

Постоянные ω и μ называются соответственно собственной частотой и коэффициентом затухания осциллятора. Обобщенный импульс частицы равен

$$p := \frac{\partial L}{\partial \dot{q}} = \mathsf{e}^{2\mu t}\dot{q}.$$

Гамильтониан осциллятора также зависит от времени явно:

$$H = \frac{1}{2}\,\mathsf{e}^{-2\mu t}p^2 + \frac{1}{2}\,\mathsf{e}^{2\mu t}\omega^2 q^2.$$

Гамильтоновы и лагранжевы уравнения движения для осциллятора с затуханием имеют вид

$$\left.\begin{aligned} \dot{q} &= \;\;\mathsf{e}^{-2\mu t}p, \\ \dot{p} &= -\,\mathsf{e}^{2\mu t}\omega^2 q, \end{aligned}\right\} \quad \Leftrightarrow \quad \mathsf{e}^{2\mu t}(\ddot{q} + 2\mu\dot{q} + \omega^2 q) = 0.$$

Общее решение этих уравнений параметризуется двумя постоянными: амплитудой A_0 и фазой φ,

$$q = A_0\,\mathsf{e}^{-\mu t}\,\mathsf{cos}\,(\tilde{\omega}t + \varphi), \qquad \tilde{\omega}^2 = \omega^2 - \mu^2.$$

Произвольная фаза φ соответствует произволу в выборе начала отсчета времени, и в дальнейшем мы положим $\varphi = 0$. При малых коэффициентах затухания $\mu^2 < \omega^2$ амплитуда колебаний экспоненциально затухает, если $\mu > 0$. Затухающие колебания происходят с частотой, меньшей собственной частоты осциллятора $\tilde{\omega} < \omega$. Для произвольной траектории численное значение гамильтониана зависит от времени явно

$$H = \frac{1}{2}A_0^2\omega^2\left[1 - \frac{1}{2}\,\mathsf{cos}\,2\tilde{\omega}t - \frac{1}{2}\,\mathsf{cos}\,(2\tilde{\omega}t + 2\psi)\right],$$

где $\mathsf{cos}\,\psi = \mu/\omega$. □

Функция канонических переменных $f(q,p)$ называется *интегралом движения*, если $f(q,p) = \mathsf{const}$ для любого решения канонических уравнений движения, т.е. является интегралом уравнений движения.

Теорема 20.1.1. *Если известны два интеграла движения f и g, то их скобка Пуассона $[f,g]$ также является интегралом движения.*

Доказательство. Прямая проверка равенства $d[f,g]/dt = 0$. □

Замечание. Теорема не гарантирует того, что вычисление скобки Пуассона двух интегралов движения дает новый интеграл движения. Часто она равна нулю, или интегралы движения f, g и $[f,g]$ – функционально зависимы. □

Если гамильтониан не зависит от какой-либо из координат, например, от q^1, т.е. $\partial H/\partial q^1 = 0$, то эта координата называется *циклической*, а соответствующий обобщенный импульс сохраняется в силу второго уравнения (20.13)

$$p_1 = c = \mathsf{const}.$$

При этом изменение остальных координат и импульсов во времени такое же, как в системе с координатами $q^2, \dots, q^{\mathrm{N}}$, импульсами $p_2, \dots, p_{\mathrm{N}}$ и функцией Гамильтона $H(q^2, \dots, q^{\mathrm{N}}, c, p_2, \dots, p_{\mathrm{N}})$.

В большинстве физических приложений для системы уравнений Гамильтона (20.13) решается задача Коши, т.е. ищется решение системы уравнений (20.13) с заданными начальными условиями

$$q(0) = q_0, \qquad p(0) = p_0. \tag{20.16}$$

Канонические уравнения движения (20.13) можно записать в виде

$$\dot{x} = X,$$

где $x = (q, p)$, а правая часть уравнений определяются векторным полем

$$X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i}$$

на фазовом пространстве. С геометрической точки зрения решение уравнений Гамильтона с заданными начальными условиями задает интегральную кривую этого векторного поля, проходящую через точку (20.16). Как было показано в разделе 3.6.5, интегральные кривые векторного поля задают абелеву однопараметрическую группу преобразований многообразия, которым в данном случае является фазовое пространство. Эта группа преобразований в гамильтоновой динамике имеет специальное название.

Определение. *Фазовым потоком* называется однопараметрическая группа преобразований фазового пространства

$$s_t : \quad \mathbb{T}^*(\mathbb{M}) \ni q(0), p(0) \;\mapsto\; q(t), p(t) \in \mathbb{T}^*(\mathbb{M}),$$

где $q(t), p(t)$ – решение системы уравнений Гамильтона (20.13). □

20.1.3 Потенциальное движение точечной частицы

Потенциальное движение точечной частицы массы m в евклидовом пространстве $\mathbb{R}^{\mathrm{N}}$ с декартовыми координатами q^i, $i = 1, \dots, \mathrm{N}$ описывается следующей функцией Лагранжа

$$L = m\frac{\dot{q}^i\dot{q}_i}{2} - U(q), \tag{20.17}$$

где $U(q) \geq 0$ – некоторая положительно определенная функция координат. В настоящем разделе подъем и опускание индексов производится с помощью евклидовой метрики $\delta_{ij} = \mathsf{diag}\,(+\dots+)$. Уравнения движения (*уравнения Ньютона*) при этом имеют вид

$$m\ddot{q}^i = -\frac{\partial U}{\partial q_i}. \tag{20.18}$$

Это – система обыкновенных дифференциальных уравнений второго порядка и их решение зависит от $2\mathrm{N}$ произвольных постоянных. Для их определения решают, как правило, задачу Коши. То есть ищется решение уравнений (20.18) с заданными начальными условиями:

$$q(0) = q_0, \qquad \dot{q}(0) = v_0.$$

В физических приложениях функции $U(q)$ обычно таковы, что решение этой задачи существует, единственно и определено при всех $t \in (-\infty, \infty)$. Заметим, что потенциальная энергия определена с точностью до постоянной, которая не влияет на уравнения движения.

Если потенциальная энергия имеет локальный экстремум в точке q_0, то

$$\left.\frac{\partial U}{\partial q^i}\right|_{q=q_0} = 0.$$

Следовательно, постоянная траектория $q = q_0$ удовлетворяет системе уравнений (20.18). Таким образом, локальные экстремумы потенциальной энергии определяют положения равновесия частицы. Эти положения могут быть устойчивы или не устойчивы по отношению к малым возмущениям в зависимости от того является ли локальный экстремум минимумом или максимумом потенциальной энергии. Для положительно определенной потенциальной энергии существует по крайней мере одна точка равновесия.

Посмотрим на эту задачу с точки зрения принципа наименьшего действия. Во-первых, действие

$$S[q] = \int dt \left(m\frac{\dot{q}^i \dot{q}_i}{2} - U(q) \right)$$

не является положительно определенным. Поэтому нужно говорить не о минимуме действия, а только о стационарных точках. Во-вторых, для свободного движения, $U = 0$, траектории представляют собой прямые линии

$$q = v_0 t + q_0,$$

по которым частица движется с постоянной скоростью v_0. Соответствующее действие при интегрировании по бесконечному интервалу расходится и говорить даже о стационарности действия в этом случае не имеет смысла. Подытожить сделанные замечания можно следующим образом. Потенциальное движение точечной частицы происходит таким образом, что для каждого конечного интервала времени (t_1, t_2) действие стационарно среди всех возможных траекторий, соединяющих точки q_1 и q_2. При этом первую граничную точку можно выбрать произвольным образом, а вторая должна быть такой, чтобы задача Коши, поставленная в первой точке, имела решение для некоторой скорости $\dot{q}_1 = v_0$.

Переформулируем потенциальное движение точечной частицы на гамильтоновом языке. Компоненты импульса и гамильтониан точечной частицы равны

$$\begin{aligned} p_i &= m\dot{q}_i, \\ H &= \frac{p^i p_i}{2m} + U(q). \end{aligned} \tag{20.19}$$

Гамильтоновы уравнения движения принимают вид

$$\begin{aligned} \dot{q}^i &= \frac{p^i}{m}, \\ \dot{p}_i &= -\frac{\partial U}{\partial q^i}, \end{aligned} \tag{20.20}$$

которые эквивалентны уравнениям Ньютона. Заметим, что в данном случае первое уравнение совпадает с определением импульса частицы.

Гамильтониан точечной частицы (20.19) положительно определен и для заданной траектории его значение (энергия) постоянно. Первое и второе слагаемые в (20.19) называются соответственно *кинетической* и *потенциально* энергией точечной частицы. Следующий пример показывает, что функция Гамильтона зависит от выбора системы координат в фазовом пространстве, а каноническая пуассонова структура – нет.

Пример 20.1.4. Рассмотрим частицу массы m, которая движется в трехмерном конфигурационном пространстве $q \in \mathbb{R}^3$ с заданным потенциалом U. Гамильтониан является функцией на фазовом пространстве и в произвольной криволинейной системе координат имеет вид

$$H = \frac{1}{2m} g^{ij} p_i p_j + U(q),$$

где $g^{ij} = g^{ij}(q)$ – обратная метрика в выбранной системе отсчета. Например, в декартовых координатах x, y, z

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + U(x, y, z).$$

В цилиндрических координатах r, φ, z

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2} + p_z^2\right) + U(r, \varphi, z).$$

В сферических координатах r, θ, φ

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right) + U(r, \theta, \varphi). \qquad \square$$

Каноническая пуассонова структура была определена в произвольной, в общем случае криволинейной, системе координат. Проверим корректность этого определения относительно преобразования координат, поскольку при переходе от одной системы координат к другой структурные функции могли бы измениться. Допустим, что мы определили каноническую пуассонову структуру в декартовой системе координат. В рассматриваемом случае преобразование координат имеет вид $q, p \mapsto Q(q), P(q, p)$, где большие буквы обозначают криволинейные (сферические или цилиндрические) координаты в конфигурационном пространстве. Скобки Пуассона новых координат и импульсов равны:

$$\begin{aligned}
[Q^i, Q^j] &= 0, \\
[Q^i, P_j] &= \frac{\partial Q^i}{\partial q^k}\frac{\partial P_j}{\partial p_k}, \\
[P_i, P_j] &= \frac{\partial P_i}{\partial q^k}\frac{\partial P_j}{\partial p_k} - \frac{\partial P_i}{\partial p_k}\frac{\partial P_j}{\partial q^k}.
\end{aligned}$$

Поскольку компоненты обобщенных импульсов являются компонентами ковектора, то они преобразуются по правилу

$$P_i = \frac{\partial q^k}{\partial Q^i} p_k.$$

Теперь нетрудно проверить выполнение скобок Пуассона:

$$[Q^i, P_j] = \delta^i_j, \quad [Q^i, Q^j] = 0, \quad [P_i, P_j] = 0.$$

Таким образом, структурные функции имеют канонический вид и не зависят от выбора системы координат в конфигурационном пространстве $\mathbb{R}^3$. Рассматриваемые преобразования координат $q, p \mapsto Q, P$ в фазовом пространстве относятся к классу канонических преобразований, которые будут рассмотрены в разделе 20.1.10.

Определение. Преобразование координат $q = q(Q)$ конфигурационного пространства $\mathbb{R}^{\mathrm{N}}$ называется *точечным*. □

20.1.4 Лемма Стокса

В этом и следующих разделах механика частиц будет рассмотрена с более общей точки зрения. Начнем с геометрического рассмотрения, которое затем применим к гамильтоновой динамике точечных частиц. Пусть на многообразии $\mathbb{M}$ нечетной размерности $\dim \mathbb{M} = 2\mathrm{N} + 1$ задана 2-форма

$$B = \frac{1}{2} dx^\alpha \wedge dx^\beta B_{\alpha\beta}.$$

Поскольку матрица, задающая 2-форму в локальной системе координат антисимметрична, $B_{\alpha\beta} = -B_{\beta\alpha}$, а многообразие нечетномерно, то ее определитель равен нулю. Это значит, что в каждой точке $x \in \mathbb{M}$ у нее существует по крайней мере один нетривиальный собственный вектор $X = X^\alpha \partial_\alpha$ с нулевым собственным значением. Отсюда вытекает, что для 2-формы B существует *нулевое* векторное поле X со свойством

$$B(X, Y) = X^\alpha Y^\beta B_{\alpha\beta} = 0, \quad \forall\, Y = Y^\alpha \partial_\alpha.$$

Пространство нулевых векторных полей линейно.

Определение. 2-форма называется *неособой*, если размерность пространства нулевых векторов минимальна, т.е. равна нулю или единице на многообразиях четной и нечетной размерности соответственно. □

Пример 20.1.5. Рассмотрим 2-форму ϖ в фазовом пространстве $\mathbb{R}^{2\mathrm{N}}$ с координатами $x = (x^1 \dots x^{2\mathrm{N}}) = (q^1 \dots q^{\mathrm{N}}, p_1 \dots p_{\mathrm{N}})$,

$$\varpi = dp_i \wedge dq^i = dx^{\mathrm{N}+1} \wedge dx^1 + dx^{\mathrm{N}+2} \wedge dx^2 + \ldots + dx^{2\mathrm{N}} \wedge dx^{\mathrm{N}} = \frac{1}{2} dx^\alpha \wedge dx^\beta \varpi_{\alpha\beta}, \quad (20.21)$$

где ϖ – каноническая симплектическая форма (18.1). Эта форма неособа, т.к. ее определитель равен единице $\det \varpi_{\alpha\beta} = 1$. □

Пусть на многообразии $\mathbb{M}$, $\dim \mathbb{M} = 2\mathrm{N} + 1$, задана 1-форма $A = dx^\alpha A_\alpha$. Предположим, что внешний дифференциал этой формы является неособым. Это значит, что в каждой точке многообразия $x \in \mathbb{M}$ существует единственный, с точностью до умножения на постоянную, нулевой вектор X такой, что

$$dA(X, Y) = 0, \quad \forall\, Y. \quad (20.22)$$

Тем самым 1-форма A определяет нулевое векторное поле X на $\mathbb{M}$ с точностью до умножения на отличную от нуля функцию. Назовем интегральные кривые $x^\alpha(t)$,

$$\frac{dx^\alpha}{dt} = X^\alpha,$$

нулевого векторного поля *характеристиками* формы A. Далее, пусть γ_1 – замкнутая кривая на $\mathbb{M}$, которая ни в одной своей точке не касается характеристик. Тогда множество характеристик, выходящих из точек кривой γ_1, образует *трубку характеристик*. Трубка характеристик определена по крайней мере в некоторой окрестности кривой γ_1.

Лемма 20.1.1 (Стокс). *Интеграл от 1-формы A с неособым внешним дифференциалом dA по любой из двух замкнутых кривых γ_1 и γ_2, охватывающих одну и ту же трубку характеристик, одинаков:*

$$\oint_{\gamma_1} A = \oint_{\gamma_2} A,$$

если $\gamma_1 - \gamma_2 = \partial S$, где S – часть трубки характеристик.

Доказательство. По формуле Стокса имеем

$$\oint_{\gamma_1} A - \oint_{\gamma_2} A = \oint_{\partial S} A = \int_S dA.$$

Этот интеграл обращается в нуль, т.к. значение формы dA на любой паре векторов, касательных к трубке характеристик, равен нулю в силу (20.22). □

20.1.5 Канонические уравнения Гамильтона

Из безобидной, на первый взгляд, леммы Стокса непосредственно вытекают все основные положения гамильтоновой динамики. Предположим, что механической системе соответствует топологически тривиальное фазовое пространство $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \approx \mathbb{R}^{2\mathrm{N}}$. Рассмотрим *расширенное фазовое пространство* $\mathbb{R}^{2\mathrm{N}+1}$ с координатами $q^1, \ldots, q^{\mathrm{N}}, p_1, \ldots, p_{\mathrm{N}}, t$, в котором время t рассматривается, как независимая дополнительная координата.

Замечание. Предположение о том, что частица движется в тривиальном конфигурационном пространстве $\mathbb{R}^{\mathrm{N}}$ и, следовательно, в тривиальном фазовом пространстве $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \approx \mathbb{R}^{2\mathrm{N}}$ не означает, что приведенные ниже формулы справедливы только для движения в топологически тривиальных пространствах. Данное предположение просто означает, что мы рассматриваем механическую систему в определенной карте. При движении частицы по нетривиальному многообразию необходимо дополнительно проследить за склейкой карт. В дальнейшем мы не будем обсуждать этот вопрос. □

Пусть на расширенном фазовом пространстве задана некоторая функция Гамильтона $H(q, p, t)$. Определим 1-форму на расширенном фазовом пространстве

$$A = dq^i p_i - dtH. \tag{20.23}$$

Эта форма называется *интегральным инвариантом Пуанкаре–Картана*. Точнее, интегральный инвариант получается после интегрирования формы A по замкнутому контуру в расширенном фазовом пространстве. Смысл названия будет ясен из дальнейшего рассмотрения. Внешний дифференциал 1-формы (20.23) является неособым, т.к. каноническая форма $\varpi = dp_i \wedge dq^i$ неособа.

Теорема 20.1.2. *Характеристики формы A в расширенном фазовом пространстве $\mathbb{R}^{2N+1}$ однозначно проектируются на фазовое пространство q, p, т.е. задаются функциями $q(t)$, $p(t)$. Эти функции удовлетворяют уравнениям Гамильтона:*

$$\begin{aligned} \frac{dq^i}{dt} &= \frac{\partial H}{\partial p_i}, \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q^i}, \end{aligned} \tag{20.24}$$

Другими словами, характеристики формы A представляют собой траектории фазового потока в расширенном фазовом пространстве, т.е. интегральные кривые канонических уравнений (20.24).

Доказательство. Дифференциал формы (20.23) равен

$$dA = dp_i \wedge dq^i - \frac{\partial H}{\partial q^i} dq^i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt$$

Эта 2-форма неособа, т.к. ранг матрицы, составленной из ее координат, равен $2N$. Прямая подстановка показывает, что вектор

$$X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial}{\partial t} \tag{20.25}$$

является нулевым вектором формы A

$$dA(X,Y) = -\frac{\partial H}{\partial q^i}Y^i - Y_i\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}Y^0\frac{\partial H}{\partial q^i} + Y^i\frac{\partial H}{\partial q^i} + \frac{\partial H}{\partial q^i}Y^0\frac{\partial H}{\partial p_i} + Y_i\frac{\partial H}{\partial p_i} = 0.$$

Это значит, что векторное поле (20.25) задает направление характеристик 1-формы (20.23). С другой стороны, вектор (20.25) является вектором скорости фазового потока (20.24). Действительно, интегральные кривые $\{q(\tau), p(\tau), t(\tau)\}$ в расширенном фазовом пространстве задаются уравнениями

$$\frac{dq^i}{d\tau} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{d\tau} = -\frac{\partial H}{\partial q^i}, \qquad \frac{dt}{d\tau} = 1. \tag{20.26}$$

В силу последнего уравнения они однозначно проектируются на фазовое пространство. Таким образом, интегральные кривые (20.24) представляют собой проекции характеристик формы (20.23) на фазовое пространство. □

Применим теперь лемму Стокса и интегральному инварианту Пуанкаре–Картана.

Теорема 20.1.3. *Пусть две замкнутые кривые γ_1 и γ_2 охватывают одну и ту же трубку характеристик формы (20.23). Тогда интегралы по ним от интегрального инварианта Пуанкаре–Картана одинаковы:*

$$\oint_{\gamma_1}\bigl(dqp - dtH\bigr) = \oint_{\gamma_2}\bigl(dqp - dtH\bigr).$$

Именно поэтому 1-форма $dqp - dtH$ называется интегральным инвариантом.

В частном случае, когда замкнутые кривые лежат в фазовом подпространстве, соответствующем постоянному значению времени, $t = \mathsf{const} \Leftrightarrow dt = 0$, получаем

Следствие. Фазовый поток сохраняет интеграл

$$\oint_\gamma dqp$$

вдоль произвольной замкнутой кривой в фазовом пространстве.

Форма $dqp := dq^i p_i$ в фазовом пространстве называется *относительным интегральным инвариантом Пуанкаре* или *формой Лиувилля*. Точнее, интегральный инвариант получается после интегрирования формы Лиувилля по замкнутому контуру γ. Он имеет простой геометрический смысл. Пусть S – двумерная ограниченная ориентированная поверхность с кусочно гладкой границей такая, что $\partial S = \gamma$, тогда по формуле Стокса

$$\oint_\gamma dqp = \int_S dp \wedge dq,$$

где $dp \wedge dq = dp_i \wedge dq^i$. Отсюда вытекает

Следствие. Фазовый поток $s_t := \exp(tX_H)$ сохраняет сумму ориентированных площадей проекций поверхности на N координатных плоскостей q^i, p_i:

$$\int_S dp \wedge dq = \int_{s_t S} dp \wedge dq.$$

В этом смысле каноническая 2-форма $\varpi = dp \wedge dq$ является абсолютным интегральным инвариантом фазового потока s_t.

Замечание. Эпитеты "относительный" и "абсолютный" интегральный инвариант не несут глубокого смысла. Исторически сложилась так, что относительным называют инвариант, полученный после интегрирования по замкнутому (компактному и без края) многообразию. Абсолютным называют инвариант, возникающий после интегрирования некоторой формы по компактному многообразию с краем. □

20.1.6 Принцип Мопертюи

В приложениях часто встречаются гамильтонианы, не зависящие от времени явно $H = H(q,p)$. В этом случае численное значение гамильтониана (энергия) на фазовой траектории сохраняется

$$\frac{dH}{dt} = [H,H] = 0.$$

Покажем, что наличие интеграла энергии позволяет понизить размерность расширенного фазового пространства $\mathbb{R}^{2N+1}$ на две единицы и свести задачу к интегрированию некоторой системы канонических уравнений в $(2N-1)$-мерном пространстве. Предположим, что в некоторой области фазового пространства уравнение

$$H(q,p) = E$$

можно решить относительно p_1:

$$p_1 = K(Q,P,T;E),$$

где $Q = (q^2,\dots,q^N)$, $P = (p_2,\dots,p_N)$ и $T = -q^1$. Тогда

$$pdq - Hdt = PdQ - KdT - d(Ht) + tdH. \tag{20.27}$$

Пусть γ – интегральная кривая канонических уравнений (20.24). Она лежит на $2\mathrm{N}$-мерном подмногообразии $H(q,p)=E$ в расширенном фазовом пространстве $\mathbb{R}^{2\mathrm{N}+1}$. Спроектируем расширенное фазовое пространство на фазовое пространство $\mathbb{R}^{2\mathrm{N}}$. При этом поверхность $H=E$ проектируется на $(2\mathrm{N}-1)$-мерное подмногообразие $\mathbb{M}^{2\mathrm{N}-1}$ фазового пространства, которое определяется тем же уравнением $H=E$, а кривая γ – на кривую $\bar\gamma$, лежащую на этом подмногообразии. Переменные Q,P,T образуют локальную систему координат на $\mathbb{M}^{2\mathrm{N}-1}$. Но характеристики формы $PdQ-KdT$ удовлетворяют уравнениям Гамильтона, поскольку $dH=0$ на $\mathbb{M}^{2\mathrm{N}-1}$, а полный дифференциал $d(Ht)$ в (20.27) не влияет на уравнения движения. Это доказывает следующее утверждение.

Теорема 20.1.4. *Если гамильтониан не зависит от времени явно, то фазовые траектории канонических уравнений (20.24) на подмногообразии $\mathbb{M}^{2\mathrm{N}-1}$ фазового пространства, определяемом уравнением $H(q,p)=E$, удовлетворяют каноническим уравнениям*

$$\frac{dq^i}{dq^1}=-\frac{\partial K}{\partial p_i},\qquad \frac{dp_i}{dq^1}=\frac{\partial K}{\partial q^i},\qquad i=2,\dots,\mathrm{N}, \tag{20.28}$$

где функция $K(q^1,\dots,q^{\mathrm{N}},p_2,\dots,p_{\mathrm{N}},E)$ определяется уравнением

$$H(q^1,\dots,q^{\mathrm{N}},K,p_2,\dots,p_{\mathrm{N}})=E.$$

Замечание. В этой теореме роль времени играет первая координата q^1, и канонические уравнения определяют не динамику системы, а форму траектории. □

Покажем, каким образом канонические уравнения Гамильтона связаны с принципом наименьшего действия. Рассмотрим интегральную кривую уравнений (20.26) в расширенном фазовом пространстве (q,p,t), соединяющую две точки (q_0,p_0,t_0) и (q_1,p_1,t_1).

Теорема 20.1.5. *Кривая γ является стационарной точкой интеграла*

$$\int_\gamma (dq\,p-dtH) \tag{20.29}$$

при таких вариациях γ, когда концы кривой остаются на N-мерных подмногообразиях $(q=q_0,t=t_0)$ и $(q=q_1,t=t_1)$.

Доказательство. В силу третьего уравнения (20.26) вариацию интеграла (20.29) можно записать как вариацию функционала действия

$$\delta\int_\gamma dt(p\dot q-H)=p\delta q\Big|_{q_0}^{q_1}+\int_\gamma dt\left[\left(\dot q^i-\frac{\partial H}{\partial p_i}\right)\delta p_i-\left(\dot p_i+\frac{\partial H}{\partial q^i}\right)\delta q^i\right].$$

Отсюда вытекает, что для исчезновения граничных вкладов достаточно зафиксировать только значения обобщенных координат при $t=t_{0,1}$. □

Понижение размерности расширенного фазового пространства при гамильтониане, не зависящем от времени, позволяет по новому взглянуть и в определенной степени оправдать употребление термина “принцип наименьшего действия” в механике. Фазовые траектории исходных канонических уравнений (20.24) целиком лежат на подмногообразии $\mathbb{M}\subset\mathbb{R}^{2\mathrm{N}+1}$, $\dim\mathbb{M}=2\mathrm{N}-1$, соответствующем фиксированному значению энергии, и являются характеристиками формы $dq\,p=dQP-dTK$. Отсюда следует

Теорема 20.1.6. *Если функция Гамильтона не зависит от времени явно, то фазовые траектории канонических уравнений (20.24), лежащие на подмногообразии* $\mathbb{M}^{2n-1}$*, соответствующем фиксированному значению энергии* $H(q,p)=E$*, являются стационарными точками интеграла*

$$\int dq\, p$$

в классе кривых, лежащих на $\mathbb{M}^{2N-1}$ *и соединяющих подпространства* $q=q_0$ *и* $q=q_1$*.*

Рассмотрим теперь проекцию экстремали, лежащую на подмногообразии постоянной энергии $\mathbb{M}$ на конфигурационное пространство. Эта кривая соединяет точки с координатами q_0 и q_1. Пусть $q=q(\tau)$, $\tau\in[a,b]$ – некоторая кривая, соединяющая те же точки $q(a)=q_0$ и $q(b)=q_1$ в конфигурационном пространстве. Она является проекцией некоторой кривой γ на подмногообразии $\mathbb{M}$. Эту кривую нетрудно построить. Для этого достаточно найти обобщенный импульс $p=\partial L/\partial\dot q$, где $\dot q=dq/d\tau$ – вектор скорости кривой в конфигурационном пространстве. Если параметр τ подобран так, что $H(q,p)=E$, то мы получаем кривую $q=q(\tau)$, $p=\partial L/\partial\dot q$ на поверхности $\mathbb{M}$. Применяя предыдущую теорему, получаем

Следствие. Среди всех кривых $q=q(\tau)$, $\tau\in[a,b]$ соединяющих точки q_0 и q_1 в конфигурационном пространстве и параметризованных так, что функция Гамильтона имеет фиксированное значение

$$H(q,\partial L/\partial\dot q)=E$$

траекторией движения механической системы является экстремаль *укороченного действия*

$$\int_a^b d\tau\, p_i\dot q^i=\int_a^b d\tau\frac{\partial L}{\partial\dot q^i}\dot q^i. \tag{20.30}$$

Это следствие называется *принципом наименьшего действия Мопертюи*. Важно отметить, что отрезок $\tau\in[a,b]$ не фиксирован и может быть разным у сравниваемых кривых. Зато одинаковой должна быть энергия. Принцип Мопертюи определяет только форму траектории, а для определения зависимости координат от времени необходимо воспользоваться условием постоянства энергии. Заметим также, что во многих задачах подынтегральное выражение в (20.30) положительно определено. В таких случаях движение происходит действительно по экстремалям укороченного действия, и можно говорить о принципе наименьшего действия.

Пример 20.1.6. Покажем, что если материальная точка движется по риманову многообразию $(\mathbb{M},g)$ только под действием сил инерции, то движение происходит по экстремалям. Пусть $ds^2=dq^idq^jg_{ij}$ – интервал риманова многообразия. Тогда функция Лагранжа для инерциального движения частицы определяется только кинетической энергией:

$$L=T=H=\frac12\left(\frac{ds}{d\tau}\right)^2=\frac12 g_{ij}\frac{dq^i}{d\tau}\frac{dq^j}{d\tau},\qquad \frac{\partial L}{\partial\dot q^i}\dot q^i=2T=\left(\frac{ds}{d\tau}\right)^2.$$

Чтобы обеспечить фиксированное значение энергии вдоль траектории, параметр τ необходимо выбрать пропорциональным длине траектории $d\tau=ds/\sqrt{2E}$. Тогда укороченное действие принимает вид

$$\int_a^b d\tau\frac{\partial L}{\partial\dot q^i}\dot q^i=\sqrt{2E}\int_a^b ds.$$

Это означает, что инерциальное движение происходит по экстремалям римановой метрики g_{ij}. □

Пример 20.1.7. Пусть движение материальной точки происходит по риманову многообразию $(\mathbb{M}, g)$ с интервалом $ds^2 = dq^i dq^j g_{ij}$ в потенциальном поле $U(q) \in \mathcal{C}^k(\mathbb{M})$. Функция Лагранжа и гамильтониан имеют вид

$$L = T - U, \quad H = T + U, \quad T = \frac{1}{2}\left(\frac{ds}{d\tau}\right)^2.$$

Чтобы обеспечить фиксированное значение энергии $H = E$, параметр τ вдоль траектории необходимо выбрать пропорциональным длине:

$$d\tau = \frac{ds}{\sqrt{2(E-U)}}.$$

Тогда укороченное действие примет вид

$$\int_\gamma d\tau \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i = \int_\gamma ds\sqrt{2(E-U)}.$$

Это значит, что движение материальной частицы по риманову многообразию $(\mathbb{M}, g)$ в потенциальном поле U происходит вдоль экстремалей метрики $\rho = g\,2(E-U)$, которая связана с исходной метрикой множителем $2(E-U)$. Эта метрика имеет особенность на границе $U(q) = E$. □

Если начальная и конечная точки экстремали в рассмотренных примерах достаточно близки, то экстремум длины является минимумом. Это оправдывает название "принцип наименьшего действия".

Пример 20.1.8 (**Метод факторизации**). Для нахождения траектории частицы иногда удобно выделить из гамильтониана $H(q,p)$ отличный от нуля множитель. Пусть $f(q,p) > 0$ – некоторая положительная функция на фазовом пространстве. Рассмотрим новый гамильтониан

$$\widetilde{H} = f(q,p)\Big(H(q,p) - E\Big), \quad E = \mathsf{const}.$$

Соответствующие уравнения движения имеют вид

$$\begin{aligned} \frac{dq}{ds} &= \quad f\frac{\partial H}{\partial p} + \frac{\partial f}{\partial p}(H-E), \\ \frac{dp}{ds} &= -f\frac{\partial H}{\partial q} - \frac{\partial f}{\partial q}(H-E). \end{aligned} \tag{20.31}$$

Эти уравнения на инвариантной поверхности фазового пространства, определяемой уравнением $\widetilde{H} = 0$ или $H = E$, эквивалентны гамильтоновым уравнениям (20.13) для исходного гамильтониана H. Уравнения (20.31) позволяют определить только форму траектории. Для определения эволюции частицы во времени достаточно проинтегрировать уравнение

$$\frac{dt}{ds} = f\Big(q(s), p(s)\Big),$$

где $q(s), p(s)$ – соответствующая траектория. □

20.1.7 Уравнение Гамильтона–Якоби

Рассмотрим *расширенное конфигурационное* пространство $\mathbb{R}^{N+1}$ с координатами $\{q,t\} = \{q^1,\dots,q^N,t\}$, которое получается из конфигурационного пространства $\mathbb{R}^N$ добавлением еще одного измерения – времени.

Определение. Функцией действия $S(q,t)$ называется интеграл

$$S(q,t) = \int_\gamma d\tau L, \tag{20.32}$$

где $L(q,\dot q,\tau)$ – функция Лагранжа, и интегрирование ведется вдоль экстремали $\gamma := \{q(\tau),t(\tau)\}$, с фиксированным началом q_0,t_0 и переменным концом q,t расширенного конфигурационного пространства. □

Это определение корректно по крайней мере в малой окрестности начальной точки q_0,t_0. Точнее, функция действия (20.32) определена в некоторой окрестности $\mathbb{U}$ точки q_0,t_0, если любую точку этой окрестности можно соединить с точкой q_0,t_0 экстремалью, целиком лежащей в $\mathbb{U}$, и эта экстремаль единственна. Если точка q,t лежит далеко от q_0,t_0, то интеграл (20.32) может не определить функцию действия $S(q,t)$ по двум причинам. Во-первых, может существовать несколько экстремалей, соединяющих точки q_0,t_0 и q,t. Во-вторых, возможно, что точки q_0,t_0 и q,t вообще нельзя соединить экстремалью. С другой стороны, для любой точки q,t расширенного конфигурационного пространства можно так подобрать начальную точку q_0,t_0, что функция действия будет определена в некоторой окрестности точки q,t. В дальнейшем мы будем считать, что точка q_0,t_0 фиксирована, и рассматривать те области расширенного конфигурационного пространства, где функция действия определена.

Аргументом функции действия является верхний предел интеграла (20.32), поэтому ее дифференциал равен

$$dS(q,t) = dq^i p_i - dtH, \tag{20.33}$$

где $p_i = \partial L/\partial\dot q^i$ и $H = \dot q^i p_i - L$ определяются в конечной точке кривой γ. Таким образом, дифференциал функции действия dS является интегральным инвариантом Пуанкаре–Картана (20.23).

Теорема 20.1.7. *Функция действия удовлетворяет уравнению Гамильтона–Якоби*

$$\frac{\partial S}{\partial t} + H\left(q,\frac{\partial S}{\partial q},t\right) = 0. \tag{20.34}$$

Доказательство. Достаточно заметить, что из выражения (20.33) следуют равенства

$$\frac{\partial S}{\partial t} = -H(q,p,t), \qquad p_i = \frac{\partial S}{\partial q^i}. \qquad \square$$

Очевидно, что любое решение уравнения Гамильтона–Якоби определено с точностью до аддитивной постоянной, которая соответствует произволу в выборе начальной точки q_0,t_0 экстремали γ в определении функции действия (20.32).

Уравнение Гамильтона–Якоби представляет собой уравнение в частных производных первого порядка. Покажем, что его интегрирование сводится к интегрированию обыкновенных дифференциальных уравнений Гамильтона. Поставим для уравнения

Гамильтона–Якоби задачу Коши, т.е. будем искать решение уравнения (20.34) с начальным условием

$$S(q,t_0)=S_0(q). \tag{20.35}$$

Чтобы построить решение этой задачи, рассмотрим задачу Коши для уравнений Гамильтона (20.24) с начальными условиями:

$$q^i(t_0)=q^i_0,\qquad p_i(t_0)=\left.\frac{\partial S_0}{\partial q^i}\right|_{q_0}.$$

Соответствующее этим начальным условиям решение изображается в расширенном конфигурационном пространстве кривой $q(t)$, которая является стационарной точкой действия $\int dtL$, где $L(q,\dot q,t)$ есть преобразование Лежандра по импульсу p от гамильтониана $H(q,p,t)$. Эта траектория называется *характеристикой* задачи Коши для уравнения Гамильтона–Якоби, выходящей из точки q_0,t_0. При временах, достаточно близких к t_0, значения $q(t),t$ можно принять за координаты в окрестности точки q_0,t_0 расширенного конфигурационного пространства. Построим теперь функцию действия

$$S(q,t)=S_0(q_0)+\int_{q_0,t_0}^{q,t}d\tau L(q,\dot q,\tau),$$

где интегрирование ведется вдоль экстремали, соединяющей точки q_0,t_0 и q,t. После этого проверяется, что эта функция действия удовлетворяет уравнению Гамильтона–Якоби (20.34) и начальному условию (20.35). Можно также доказать, что это решение задачи Коши для уравнения Гамильтона–Якоби единственно.

Если гамильтониан не зависит от времени явно, то зависимость функции действия от времени легко находится:

$$S(q,t)=-Et+W(q), \tag{20.36}$$

где $W(q)$ – *укороченная функция действия*, зависящая только от координат, и $E=$ const. Из уравнения Гамильтона–Якоби (20.34) следует, что укороченная функция действия должна удовлетворять уравнению в частных производных

$$H\left(q,\frac{\partial W}{\partial q}\right)=E. \tag{20.37}$$

Отсюда следует, что постоянная E равна энергии системы. Это уравнение называется *укороченным* уравнением Гамильтона–Якоби.

Пример 20.1.9. Поясним определение функции действия на простом примере свободной точечной частицы единичной массы, движущейся в евклидовом пространстве $\mathbb{R}^n$. В этом случае лагранжиан и гамильтониан частицы имеют вид

$$L=\frac{1}{2}\dot q^i\dot q_i,\qquad H=\frac{1}{2}p^ip_i,$$

где подъем и опускание индексов осуществляется с помощью символов Кронекера. Траекториями свободной частицы являются прямые линии и только они. Экстремаль $q^i(\tau)$, соединяющая точки q^i_0,t_0 и q^i,t, задается линейной функцией

$$q^i(\tau)=q^i_0+\frac{q^i-q^i_0}{t-t_0}(\tau-t_0),$$

описывающей равномерное прямолинейное движение частицы со скоростью $\dot{q}^i = (q^i - q^i_0)/(t - t_0)$. Нетрудно вычислить соответствующую функцию действия

$$S(q,t) = \int_{t_0}^{t} d\tau \frac{1}{2}\left(\frac{q-q_0}{t-t_0}\right)^2 = \frac{1}{2}\frac{(q-q_0)^2}{t-t_0},$$

где

$$(q-q_0)^2 = (q^i - q^i_0)(q_i - q_{0i}).$$

Функция действия определена для всех $q \in \mathbb{R}^{\mathrm{N}}$, $t > t_0$ и удовлетворяет уравнению Гамильтона–Якоби:

$$\frac{\partial S}{\partial t} + \frac{1}{2}\frac{\partial S}{\partial q^i}\frac{\partial S}{\partial q_i} = 0.$$

Поскольку гамильтониан частицы не зависит от времени явно, то для любой траектории частицы энергия сохраняется:

$$E = \frac{1}{2}\left(\frac{q-q_0}{t-t_0}\right)^2 = \mathsf{const}.$$

Отсюда можно выразить время через координаты

$$t - t_0 = \sqrt{\frac{(q-q_0)^2}{2E}}.$$

Поскольку $S = E(t - t_0)$, то укороченное действие равно

$$W = Et + S = 2E(t-t_0) + Et_0 = \sqrt{2E(q-q_0)^2} + Et_0.$$

Нетрудно проверить, что укороченное действие удовлетворяет укороченному уравнению Гамильтона–Якоби

$$E = \frac{1}{2}\frac{\partial W}{\partial q^i}\frac{\partial W}{\partial q_i}. \qquad \square$$

Уравнение Гамильтона–Якоби предоставляет мощный метод решения задач механики точечных частиц. При обобщении этого метода на теорию поля возникают существенные усложнения. В настоящее время, насколько известно автору, это интересное обобщение не развито.

20.1.8 Принцип Гюйгенса

Многие понятия гамильтоновой механики возникли при перенесении на общие вариационные принципы весьма простых и наглядных понятий геометрической оптики. В настоящем разделе мы обсудим некоторые аспекты геометрической оптики.

Рассмотрим свет, распространяющийся в среде (конфигурационном пространстве), которую мы отождествим с евклидовым пространством $\mathbb{R}^3$ с декартовыми координатами q^i, $i = 1, 2, 3$. В общем случае скорость света зависит от точки $q \in \mathbb{R}^3$ (неоднородная среда) и направления луча света (неизотропная среда). Неизотропность среды можно описать, задав в каждой точке q поверхность в касательном пространстве $\mathbb{T}_q(\mathbb{R}^3)$. Для этого отложим в начале координат касательного пространства $\mathbb{T}_q(\mathbb{R}^3)$ вектор скорости распространения света в точке q. Эта поверхность называется *индикатрисой*. Согласно *принципу Ферма*, свет распространяется в среде из точки q_0 в точку q за кратчайшее время.

Если среда изотропна, то принцип Ферма приобретает простую математическую форму. Пусть $q(t)$ – траектория луча. Тогда для скорости света $v^i := dq^i/dt$ справедливо равенство

$$v^2 dt^2 = ds^2,$$

где $v^2 = v^i v^j \delta_{ij}$ и $ds^2 := dq^i dq^j \delta_{ij}$. Отсюда следует, что время прохождения луча между точками q_0 и q вдоль траектории $q(t)$ равно

$$t = \int_{q_0}^{q} \frac{ds}{v},$$

где $v := \sqrt{v^2}$ и $ds := \sqrt{ds^2}$. Для изотропной среды показатель преломления n равен отношению скорости света c в вакууме к скорости света v в среде, $n = c/v$. Поэтому время распространения луча света дается интегралом

$$t = \frac{1}{c} \int_{q_0}^{q} ds\, n. \tag{20.38}$$

Таким образом, принцип Ферма для распространения лучей света в изотропной среде сводится к вариационной задаче для действия (20.38), в котором показатель преломления $n(q)$ является заданной функцией на конфигурационном пространстве.

Распространение света в среде можно также описать на языке волновых фронтов. Допустим, что в момент времени $t = 0$ в точке q_0 произошла вспышка света. Рассмотрим множество точек, до которых свет дойдет за время, меньшее или равное $t > 0$. Граница этого множества $\Phi_0(t)$, которая представляет собой некоторую поверхность в $\mathbb{R}^3$, называется *волновым фронтом* точки q_0 через время t и состоит из точек, до которых свет дойдет за время t и не может дойти быстрее. При этом мы предполагаем, что среда такова, что волновые фронты представляют собой гладкие поверхности в $\mathbb{R}^3$. Тогда между волновыми фронтами для разных моментов времени t имеется замечательное соотношение.

Теорема 20.1.8 (Принцип Гюйгенса). *Рассмотрим волновой фронт $\Phi_0(t)$ точки q_0 в момент времени t. Для каждой точки этого фронта $q \in \Phi_0(t)$ построим волновой фронт $\Phi_q(s)$ через время $s > 0$. Тогда волновой фронт $\Phi_0(t+s)$ точки q_0 через время $t+s$ будет огибающей поверхностью всех фронтов $\Phi_q(s)$ для $q \in \Phi_0(t)$.*

Доказательство. Пусть $q_{t+s} \in \Phi_0(t+s)$. Тогда существует путь из начальной точки q_0 в точку q_{t+s}, по которому свет распространяется за время $t+s$ и нет более короткого. Рассмотрим точку q_t на этом пути, до которой свет идет время t. Никакого более короткого пути из q_0 в q_t не существует, иначе путь не был бы кратчайшим. Поэтому точка q_t лежит на фронте $\Phi_0(t)$. Точно так же

путь из q_t в q_{t+s} свет проходит за время s и нет более короткого пути между этими точками. Поэтому точка q_{t+s} лежит на фронте $\Phi_q(s)$ точки $q(t)$ через время s. Осталось показать, что фронты $\Phi_q(s)$ и $\Phi_0(t+s)$ в точке $q(t+s)$ касаются. Действительно, предположим, что фронты пересекаются, как показано на рис.**??**. Тогда в некоторые точки фронта $\Phi_0(t+s)$ из точки q_t можно было бы добраться за время, меньшее s. В свою очередь это значит, что в эти точки из q_0 можно было бы добраться за время, меньшее $t+s$, что противоречит определению фронта $\Phi_0(t+s)$. □

Разумеется, точку q_0 в принципе Гюйгенса можно заменить на кривую, поверхность или вообще на произвольной замкнутое множество, а трехмерное евклидово пространство на произвольное дифференцируемое многообразие $\mathbb{M}$. При этом свет

можно заменить на распространение любого возмущения, передающегося локально, т.е. вдоль линии $\gamma \in \mathbb{M}$, на которой некоторый функционал принимает наименьшее значение.

Принцип Гюйгенса приводит к двум способам описания распространения возмущений. Во-первых, можно следить за *лучами*, т.е. кратчайшими путями, вдоль которых распространяется свет. В этом случае распространение света задается вектором скорости $\dot{q}$ в каждой точке конфигурационного пространства, т.е. некоторой точкой на индикатрисе. Во-вторых, мы можем следить за распространением волнового фронта.

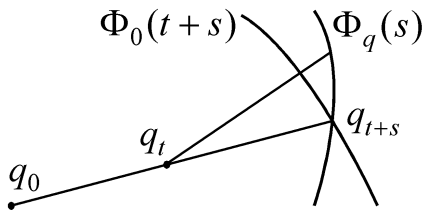


Рис. 20.2: Невозможность пересечения фронтов.

Определим скорость движения p волнового фронта следующим образом. Для каждой точки q_0 определим функцию $S_0(q)$, как наименьшее время распространения света из точки q_0 в точку q. По-построению, поверхность уровня функции $S_0(q)$ является волновым фронтом:

$$\Phi_0(t) := \{q \in \mathbb{R}^3 : \quad S_0(q) = t\}.$$

Определение. Ковектор с компонентами

$$p_i := \frac{\partial S_0(q)}{\partial q^i}$$

называется *ковектором нормальной медлительности фронта.* □

Такое название связано с тем, что чем больше
градиент функции $S_0(q)$, тем медленнее движется фронт. Действительно, в линейном приближении справедливо равенство

$$S_0(q + dq) = S_0(q) + p_i dq^i = t + dt.$$

Отсюда следует, что чем больше расстояние dq, пройденное светом за время dt, тем меньше ковектор нормальной медлительности фронта p. Векторы X, касательные к волновому фронту, определяются, как ортогональное дополнение к ковектору нормальной медлительности фронта:

$$X^i p_i = 0.$$

Если на конфигурационном пространстве определена риманова метрика g_{ij}, то можно определить вектор нормальной медлительности фронта с компонентами $p^i = g^{ij} p_j$. По-построению, этот вектор ортогонален волновому фронту.

Если среда анизотропна, то направление лучей света $\dot{q}$ и направление движения фронта p не совпадают. Однако они связаны между собой простым соотношением, которое легко выводится из принципа Гюйгенса. Напомним, что оптические свойства среды в точке q характеризуются поверхностью в касательном пространстве $\mathbb{T}_q(\mathbb{R}^3)$ – индикатрисой.

Определение. Направление гиперплоскости (т.е. вектор, перпендикулярный к данной гиперплоскости), касающейся индикатрисы в точке $\dot{q}$, называется *сопряженным* к направлению $\dot{q}$. □

Предложение 20.1.1. *Направление p волнового фронта $\Phi_0(t)$ в точке q_t сопряжено направлению луча $\dot{q}$ в данной точке.*

Доказательство. Рассмотрим точки $q_{t-\epsilon}$, лежащие на луче с началом в точке q_0 и концом в q_t, при малых $\epsilon > 0$. Фронт волны $\Phi_{q_{t-\epsilon}}(\epsilon)$ в точке $q_{t-\epsilon}$ в момент времени ϵ можно представить в виде

$$\Phi_{q_{t-\epsilon}}(\epsilon) = \{q^i_{t-\epsilon} + \dot{q}^i_t \epsilon\},$$

где вектор скорости $\dot{q}$ пробегает всю индикатрису, с точностью порядка $\mathsf{o}(\epsilon)$. Из принципа Гюйгенса следует, что фронт $\Phi_{q_{t-\epsilon}}(\epsilon)$ касается фронта $\Phi_0(t)$ в точке q_t. В пределе $\epsilon \to 0$ получаем утверждение предложения. □

Теперь сравним геометрическую оптику с построениями предыдущего раздела. Лучом является траектория частицы $q(t)$ в конфигурационном пространстве $\mathbb{R}^3$. Ковектор нормальной медлительности фронта – это импульс частицы p. Принцип Ферма соответствует принципу наименьшего действия. Функция $S_0(q)$ есть ни что иное, как функция действия (20.32), в которой время t соответствует концу траектории $q = q(t)$. Поверхности уровня функции действия $S(q,t)$ соответствуют волновым фронтам.

20.1.9 Переменные действие-угол

Знание интегралов движения позволяет упросить задачу. При этом интерес представляют функционально независимые интегралы движения. Из определения функциональной независимости следует, что на фазовом пространстве размерности 2N может существовать не более 2N функционально независимых интегралов движения.

Чтобы проинтегрировать систему из 2N обыкновенных дифференциальных уравнений первого порядка, достаточно знать 2N независимых первых интегралов. Оказывается, что если задана каноническая система уравнений движения, то ситуация существенно проще: достаточно знать только N независимых первых интегралов. Это происходит потому что каждый интеграл движения позволяет понизить порядок системы уравнений не на одну, а на две единицы.

Определение. Функция $F \in \mathcal{C}^1(\mathbb{M})$ на пуассоновом многообразии $\mathbb{M}$ называется *первым интегралом* механической системы с гамильтонианом H, если ее скобка Пуассона с гамильтонианом равна нулю, $[F,H] = 0$. Две функции $F_1, F_2 \in \mathcal{C}^1(\mathbb{M})$ находятся в *инволюции*, если их скобка Пуассона равна нулю, $[F_1, F_2] = 0$. □

Теорема 20.1.9 (Лиувилль). *Предположим, что на симплектическом многообразии $\mathbb{M}$ размерности 2N заданы N функционально независимых дифференцируемых функций $\{F_i\}$, $i = 1, \dots \mathrm{N}$, которые находятся в инволюции, $[F_i, F_j] = 0$. Рассмотрим N-мерное подмногообразие $\mathbb{M}_f \hookrightarrow \mathbb{M}$, которое является множеством уровня функций F_i:*

$$\mathbb{M}_f := \{x \in \mathbb{M} : \quad F_i = f_i = \mathsf{const},\ i = 1, \dots, \mathrm{N}\}.$$

Тогда справедливы следующие утверждения:

1) *Подмногообразие $\mathbb{M}_f$ инвариантно относительно фазового потока с функцией Гамильтона $H = F_i$ при любом фиксированном i.*
2) *Если подмногообразие $\mathbb{M}_f$ компактно и связно, то оно диффеоморфно N-мерному тору $\mathbb{T}^{\mathrm{N}}$.*

3) Фазовый поток с функцией Гамильтона H определяет на $\mathbb{M}_f$ условно-периодическое движение, т.е. в угловых координатах $\{\varphi^i,\ \mathsf{mod}\, 2\pi\}$ на торе уравнения движения имеют вид

$$\frac{d\varphi^i}{dt} = \omega^i, \qquad \omega^i = \omega^i(f) = \mathsf{const}, \quad \forall i. \tag{20.39}$$

4) Канонические уравнения движения с функцией Гамильтона H интегрируются в квадратурах.

Доказательство. См., например, [**?**].

Прокомментируем первые два утверждения теоремы.

Поскольку фазовое пространство $\mathbb{M}$ является симплектическим и, следовательно, пуассоновым многообразием, то каждой функции F_i ставится в соответствие векторное поле X_i по правилу (18.27). При этом инволютивность первых интегралов движения влечет за собой инволютивность распределения векторных полей $\{X_i\}$. Согласно теореме Фробениуса у этого распределения существует интегральное подмногообразие, которым является $\mathbb{M}_f \hookrightarrow \mathbb{M}$. Отсюда следует инвариантность $\mathbb{M}_f$ относительно фазового потока для любой функции F_i.

Тор возникает из-за того, что векторные поля X_i, которые являются генераторами группы однопараметрических преобразований $\mathbb{M}_f$, не просто находятся в инволюции, а коммутируют между собой. □

Замечание. Если гамильтониан механической системы H не зависит от времени, то его можно выбрать в качестве одного из первых интегралов F_i. □

Определение. Гамильтонова механическая система, называется *интегрируемой*, если она имеет N или более функционально независимых интегралов движения. □

Если фазовое пространство механической системы таково, что выполнены условия теоремы Лиувилля, то на фазовом пространстве $\mathbb{M}$ существует выделенная система координат (действие-угол), в которой уравнения движения выглядят особенно просто. Для определенности будем считать, что гамильтониан системы совпадает с первым интегралом движения: $H = F_1$. В теореме Лиувилля утверждается, что подмногообразие $\mathbb{M}_f \hookrightarrow \mathbb{M}$ является N-мерным тором, инвариантно относительно фазового потока и на нем существуют угловые координаты φ^i, для которых уравнения движения имеют вид (20.39). Общее решение этих уравнений имеет простой вид

$$\varphi = \varphi_0 + \omega t, \qquad \varphi_0 = \mathsf{const}.$$

Поскольку интегралы движения функционально независимы, то в некоторой окрестности подмногообразия $\mathbb{M}_f$ в качестве координат можно выбрать совокупность функций $\{F, \varphi\}$. В этой системе координат уравнения движения принимают простой вид

$$\frac{dF}{dt} = 0, \qquad \frac{d\varphi}{dt} = \omega(F),$$

и легко интегрируются:

$$F(t) = F(0), \qquad \varphi(t) = \varphi_0 + \omega\Big(F(0)\Big)t.$$

В общем случае в координатах $\{F, \varphi\}$ симплектическая форма не будет иметь канонического вида. Однако существует существует такой набор функционально независимых функций $\{I = I(F)\}$, $i = 1, \ldots, \mathrm{N}$, что переменные $\{I, \varphi\}$ образуют такую

систему координат в окрестности $\mathbb{M}_f$, в которой симплектическая форма имеет канонический вид

$$\varpi = dI_i \wedge d\varphi^i. \tag{20.40}$$

Определение. Координаты $\{\varphi^1, \dots, \varphi^{\mathrm{N}}, I_1, \dots, I_{\mathrm{N}}\}$ в некоторой окрестности подмногообразия фазового пространства $\mathbb{M}_f \subset \mathbb{R}^{2\mathrm{N}}$, в которых канонические уравнения движения имеют вид

$$\frac{dI}{dt} = 0, \qquad \frac{d\varphi}{dt} = \omega(I),$$

и симплектическая форма является канонической (20.40) называются *переменными действие-угол.* □

Переменные действие-угол являются координатами на фазовом пространстве. При этом угловые переменные удобно рассматривать как обобщенные координаты, а переменные действия – как сопряженные импульсы. Из определения сразу следует, что преобразование координат $q, p \mapsto \varphi, I$, если оно существует, является каноническим.

Переменные I так же как и функции F являются первыми интегралами движения. В переменных действие-угол гамильтониан механической системы имеет вид $H = F_1 = H(I)$ и не зависит от угловых переменных φ. То есть каждая из угловых координат φ является циклической.

Теорема 20.1.10. *Координаты действие-угол $\{\varphi, I\}$ существуют в некоторой окрестности подмногообразия $\mathbb{M}_f \hookrightarrow \mathbb{R}^{2\mathrm{N}}$.*

Доказательство. Рассмотрим множество торов $\mathbb{M}_F$, соответствующих различным значениям интегралов движения F_i. Пусть γ_i, $i = 1, \dots, \mathrm{N}$, – базисные одномерные циклы на торах $\mathbb{M}_F$, т.е. приращение координаты φ^i на цикле γ_j равно 2π, если $i = j$ и 0, если $i \neq j$. Каждое подмногообразие $\mathbb{M}_F \hookrightarrow \mathbb{R}^{2\mathrm{N}}$ является нулевым. Действительно, векторные поля X_i, соответствующие интегралам движения F_i, образуют базис касательных пространств $\mathbb{T}_x(\mathbb{M}_F)$ для всех $x \in \mathbb{M}_F$ и попарно коммутируют. Поэтому значение канонической симплектической формы $\varpi = dp_i \wedge dq^i$ на двух произвольных базисных полях равно нулю:

$$\varpi(X_i, X_j) = \varpi_{kl} X_i^k X_j^l = \varpi^{-1kl} \partial_k F_i \partial_l F_j = [F_i, F_j] = 0.$$

Следовательно, подмногообразие $\mathbb{M}_F$ является нулевым. Отсюда вытекает, что 1-форма dqp замкнута на $\mathbb{M}_F$, т.е. ее внешняя производная $dp \wedge dq$ обращается в нуль на подмногообразии $\mathbb{M}_F$. Положим

$$I_i(F) := \oint_{\gamma_i} dqp.$$

Эти функции зависят от значений интегралов движения F_i, определяющих тор $\mathbb{M}_F$, но не зависят от выбора базисных циклов γ_i, т.к. форма dqp замкнута. Это означает, что функции $I_i(F)$ не зависят от координат φ на торе.

Теперь совершим два канонических преобразования. Поскольку скобка Пуассона интегралов движения F_i между собой равна нулю, то их можно выбрать в качестве новых импульсов: $P_i = F_i$. Совершим первое каноническое преобразование $q, p \mapsto Q, F$ с производящей функцией $S_4(p, F)$, зависящей от новых и старых импульсов F и p. Тогда старые и новые координаты определены соотношениями (20.68):

$$q = -\frac{\partial S_4}{\partial p}, \qquad Q = \frac{\partial S_4}{\partial F}.$$

Подмногообразие $\mathbb{M}_F$ в новой системе координат задано соотношениями $P = 0$. Следовательно, новые координаты Q образуют некоторую систему координат на $\mathbb{M}_F$.

Зафиксируем точку $Q_0 \in \mathbb{M}_F$. В некоторой ее окрестности можно выбрать систему координат Q, I, т.к. функции $I = I(P)$ зависят только от импульсов $P = F$. При этом импульсы P можно выразить, как функции от Q и I, т.е. $P = P(Q, I)$. Поскольку преобразование $q, p \mapsto Q, P$ каноническое, то 1-форма dQP замкнута. Поэтому в односвязной окрестности $\mathbb{U}$ точки $Q_0 \in \mathbb{U} \subset \mathbb{M}_F$ определена функция

$$S_2(Q, I) = \int_{Q_0}^{Q} dQ' P(Q', I).$$

Этот интеграл не зависит от кривой, целиком лежащей в $\mathbb{U}$ и соединяющей точки Q_0 и Q. Следовательно функцию $S_2(Q, I)$ можно выбрать в качестве производящей функции второго канонического преобразования $Q, P \mapsto \varphi, I$, зависящей от старых координат Q и новых импульсов I. Формулы преобразования имеют вид (20.66):

$$P = \frac{\partial S_2}{\partial Q}, \qquad \varphi = \frac{\partial S_2}{\partial I}.$$

Таким образом, преобразование координат $q, p \mapsto \varphi, I$ является произведением двух канонических преобразований $q, p \mapsto Q, P$ и $Q, P \mapsto \varphi, I$ и, следовательно, само является каноническим. По-построению, оно определено в окрестности тора $\mathbb{M}_f$. □

Замечание. В доказательстве теоремы переход от координат q, p к переменным действие-угол φ, I содержит только алгебраические операции и интегрирование. Это доказывает утверждение 4 теоремы Лиувилля 20.1.9. □

Переменные действие-угол определены неоднозначно. Очевидно, что координаты можно сдвигать:

$$I' = I + \mathsf{const}, \qquad \varphi' = \varphi + \mathsf{const}(I).$$

20.1.10 Канонические преобразования

Наиболее мощный и гибкий метод нахождения точных решений уравнений движения дают канонические преобразования. По-прежнему, обозначим конфигурационное и фазовое пространства соответственно через $\mathbb{R}^{\mathrm{N}}$, и $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \approx \mathbb{R}^{2\mathrm{N}}$. Будем считать, что на фазовом пространстве заданы координаты $\{x^\alpha\} = \{q^i, p_i\} := (q^1, \ldots, q^{\mathrm{N}}, p_1, \ldots, p_{\mathrm{N}})$ и каноническая пуассонова структура (18.31).

Определение. Дифференцируемое биективное отображение g фазового пространства $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}})$ на себя называется каноническим, если оно сохраняет каноническую симплектическую форму:

$$g^* \varpi = \varpi,$$

где $\varpi := dp \wedge dq = dp_i \wedge dq^i$. □

В координатах это определение записывается следующим образом. Отображение g фазового пространства задается $2\mathrm{N}$ функциями, которые в общем случае могут зависеть от времени,

$$g: \quad \mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \ni (q^i, p_i) \mapsto \Big(Q^i(q, p, t), P_i(q, p, t)\Big) \in \mathbb{T}^*(\mathbb{R}^{\mathrm{N}}), \tag{20.41}$$

где время t входит в качестве параметра. Поскольку ранг 2-формы ϖ равен $2\mathrm{N}$, то канонические преобразования задают преобразование координат (диффеоморфизм) фазового пространства с отличным от нуля якобианом:

$$\frac{\partial(Q,P)}{\partial(q,p)} \neq 0.$$

Условие сохранения канонической симплектической формы ϖ принимает вид

$$dp_i \wedge dq^i = dP_i \wedge dQ^i. \tag{20.42}$$

Это равенство приводит к дифференциальным уравнениям на функции $Q^i(q,p,t)$, $P_i(q,p,t)$, которые мы рассмотрим несколько позже.

Определение канонических преобразований можно записать в эквивалентном интегральном виде

$$\int_{\mathbb{S}} dp \wedge dq = \int_{g\mathbb{S}} dp \wedge dq, \tag{20.43}$$

где интегрирование ведется по произвольной двумерной поверхности $\mathbb{S} \subset \mathbb{T}^*(\mathbb{R}^{\mathrm{N}})$ с кусочно гладкой границей. Поскольку кокасательное расслоение $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}})$ односвязно, то к интегралу (20.43) можно применить формулу Стокса:

$$\int_{\mathbb{S}} dp \wedge dq = \oint_{\partial\mathbb{S}} dqp,$$

где мы предположили, что граница $\partial\mathbb{S}$ такова, что dq на ней может обратиться в нуль только в изолированных точках. Поэтому для односвязных фазовых пространств интегральное определение канонического преобразования (20.43) можно переписать в виде

$$\oint_{\gamma} dqp = \oint_{g\gamma} dqp, \tag{20.44}$$

где γ – произвольная кусочно гладкая замкнутая кривая в фазовом пространстве. В общем случае неодносвязных многообразий из (20.44) следует (20.43), но не наоборот. То есть сохранение 1-формы dqp является достаточным условием того, что преобразование будет каноническим.

Пример 20.1.10. Каноническая пуассонова структура определена на произвольном кокасательном расслоении $\mathbb{T}^*(\mathbb{M})$. При этом вид структурных функций не зависит от выбора координат на конфигурационном пространстве $\mathbb{M}$. Поэтому преобразование $q,p \mapsto Q(q), P(q,P)$, где Q зависит только от q, является каноническим (см., пример 20.1.4). □

Пример 20.1.11. Пусть фазовое пространство топологически тривиально, $\mathbb{T}^*(\mathbb{M}) \approx \mathbb{R}^{2\mathrm{N}}$. Поскольку линейное преобразование координат фазового пространства, порождаемое симплектической группой $\mathbb{SP}(\mathrm{N},\mathbb{R})$ (см. раздел 18.1), сохраняет каноническую симплектическую форму, то оно является каноническим. □

На практике наиболее полезными и содержательными являются нелинейные канонические преобразования, связанные со спецификой задачи: симметриями и наличием интегралов движения.

Пример 20.1.12. Фазовый поток является каноническим отображением в силу следствия из теоремы 20.1.3. □

Пример 20.1.13 (Теория возмущений). Допустим, что гамильтониан механической системы состоит из двух слагаемых,

$$H(q,p) = H_0(q,p) + H_1(q,p). \tag{20.45}$$

При этом гамильтониан H_0 настолько прост, что для него известно общее решение

$$q = q(t; q_0, p_0), \qquad p = p(t; q_0, p_0), \tag{20.46}$$

задачи Коши с начальными условиями:

$$q(0; q_0, p_0) = q_0, \qquad p(0; q_0, p_0) = p_0.$$

Как было отмечено, фазовый поток определяет каноническое преобразование, которое задается функциями (20.46). Обратное преобразование $q, p \mapsto q_0, p_0$ также является каноническим. При этом преобразовании гамильтониан H_0 обращается в нуль и уравнения движения исходной гамильтоновой системы упрощаются:

$$\dot q_0 = \frac{\partial \tilde H_1}{\partial p_0}, \qquad \dot p_0 = -\frac{\partial \tilde H_1}{\partial q_0},$$

где

$$\tilde H_1(q_0, p_0, t) = H_1\Big(q(t; q_0, p_0), p(t; q_0, p_0)\Big). \tag{20.47}$$

Обозначим через $\tilde q_0(t; q_0.p_0)$ и $\tilde p_0(t; q_0.p_0)$ решение задачи Коши для полученной системы уравнений движения с начальными данными:

$$\tilde q_0(0; q_0, p_0) = q_0, \qquad \tilde p_0(0; q_0, p_0) = p_0.$$

Это решение обладает замечательным свойством: для невозмущенной системы оно постоянно и совпадает с начальными данными. Таким образом, задача Коши для гамильтоновой системы (20.45) сведена к задаче Коши для возмущения, которое определяется гамильтонианом (20.47). Формулы для решения исходной задачи в старых координатах q, p получаются после подстановки решения $\tilde q_0, \tilde p_0$ в общее решение (20.46) невозмущенной задачи. □

Из определения канонических преобразований фазового пространства следует, что фазовое пространство вместе с каноническими отображениями образуют группу преобразований. Эта группа является подгруппой группы общих преобразований координат фазового пространства $\mathsf{diff}\,\mathbb{T}^*(\mathbb{M})$. В общем случае канонические преобразования $q^i, p_i \mapsto Q^i, P_i$ не сохраняют структуру кокасательного расслоения на фазовом пространстве, соответствующую координатам q^i, p_i, т.к. могут перемешивать канонические координаты и импульсы.

По-определению, каноническое преобразование сохраняет 2-форму ϖ (20.21) и, следовательно, все ее внешние степени. Отсюда вытекает

Теорема 20.1.11. *Канонические преобразования сохраняют интегральные инварианты* $\varpi^2, \dots, \varpi^{\mathrm{N}}$.

Поскольку $2\mathrm{N}$-форма ϖ^{N} существует и невырождена, то на фазовом пространстве существует также форма объема υ, которая ей пропорциональна (18.12). Отсюда вытекает

Следствие. Канонические преобразования сохраняют форму объема υ фазового пространства. Другими словами, якобиан канонического преобразования координат равен единице,

$$\det \frac{\partial(Q,P)}{\partial(q,p)} = 1.$$

Это значит, что объем

$$V = \int_{\mathbb{U}} \upsilon$$

произвольной области $\mathbb{U}$ фазового пространства сохраняется при каноническом преобразовании: $gV = V$.

Поскольку фазовый поток для произвольной гамильтоновой системы является каноническим преобразованием, то он также имеет интегральные инварианты $\varpi, \varpi^2, \dots, \varpi^{\mathrm{N}}$. Последний из этих инвариантов есть фазовый объем. Отсюда следует

Теорема 20.1.12 (Лиувилль). *Фазовый поток сохраняет объем произвольной области фазового пространства.*

Верно также обратное утверждение. А именно, пусть задан функционал

$$I_k = \int_{\mathbb{U}} d^{2k}x F(t,x), \qquad k = 1, \dots, \mathrm{N},$$

где $\mathbb{U} \subset \mathbb{R}^{2\mathrm{N}}$ – произвольное четномерное подмногообразие фазового пространства размерности $\dim \mathbb{U} = 2k$ с кусочно гладкой границей и $F(t,x)$ – некоторая функция времени и канонических переменных. Если этот функционал инвариантен относительно фазового потока с произвольным гамильтонианом, то он называется интегральным инвариантом.

Теорема 20.1.13. *Любой интегральный инвариант I_k фазового потока с произвольным гамильтонианом имеет вид*

$$I_k = c \int_{\mathbb{U}} \varpi^k,$$

где c – некоторая постоянная, для всех $k = 1, \dots, \mathrm{N}$.

Доказательство. Дано сравнительно недавно [**?**]. □

Важность канонических преобразований заключается в том, что они позволяют менять координаты в фазовом пространстве не меняя вида канонических уравнений движения. Чтобы показать это, рассмотрим преобразование координат в расширенном фазовом пространстве $\mathbb{R}^{2\mathrm{N}+1}$, с координатами q, p, t. Как было показано в разделе 20.1.5, фазовый поток взаимно однозначно связан с характеристиками интегрального инварианта Пуанкаре–Картана в расширенном фазовом пространстве. Поскольку характеристики 1-формы являются инвариантным объектом, то замену координат $q, p, t \mapsto Q, P, T$ удобно провести в расширенном фазовом пространстве.

Теорема 20.1.14. *Пусть Q, P, T – новая система координат в расширенном фазовом пространстве $\mathbb{R}^{2\mathrm{N}+1}$ и $K(Q,P,T)$, $S(Q,P,T)$ – такие функции, что выполнено условие*

$$dqp - dtH = dQP - dTK + dS. \tag{20.48}$$

Тогда траектории фазового потока (20.24) изображаются в координатах Q, P, T интегральными кривыми канонических уравнений

$$\frac{dQ^i}{dT} = \frac{\partial K}{\partial P_i}, \qquad \frac{dP_i}{dT} = -\frac{\partial K}{\partial Q^i}. \tag{20.49}$$

Доказательство. Поскольку $ddS = 0$, то 1-форма dS не влияет на характеристики формы (20.48). Это значит, что фазовый поток определяется также характеристиками формы $dQP - dTK$. Уравнения (20.49) следуют из теоремы 20.1.2. □

Каноническим преобразованиям координат в фазовом пространстве $q, p \mapsto Q, P$ соответствует преобразование координат $q, p, t \mapsto Q, P, t$ в расширенном фазовом пространстве, которое не меняет времени.

Теорема 20.1.15. *В новых координатах Q, P канонические уравнения движения (20.24) имеют канонический вид (20.49) со старой функцией Гамильтона, выраженной через новые координаты $K(Q, P, t) = H\Big(q(Q, P), p(Q, P), t\Big)$.*

Доказательство. Рассмотрим 1-форму $dqp - dQP$ в фазовом пространстве. Тогда для любой замкнутой кривой γ справедливо равенство

$$\oint_\gamma (dqp - dQP) = 0,$$

поскольку преобразование является каноническим. Это равенство можно записать в дифференциальной форме $dqp - dQP = dS$, где $S(Q, P)$ – некоторая функция на фазовом пространстве. Следовательно, в расширенном фазовом пространстве имеем

$$dqp - dtH = dQP - dtH + dS,$$

и справедлива предыдущая теорема. □

Таким образом, мы показали, что при канонических преобразованиях вид гамильтоновых уравнений движения не меняется. Часто это свойство принимается за определение канонических преобразований. Следующий пример показывает, что такое определение является более общим, чем принятое в настоящей монографии.

Пример 20.1.14. Преобразование координат

$$\varphi_c : \quad \mathbb{R}^{2N} \ni (q, p) \mapsto (Q = q, P = cp) \in \mathbb{R}^{2N}, \tag{20.50}$$

где $c \neq 0$ – некоторая постоянная, сохраняет гамильтонову форму уравнений движения с гамильтонианом $K = cH$. Тем не менее оно не является каноническим, т.к. каноническая форма

$$\varpi = dp \wedge dq = \frac{1}{c} dP \wedge dQ$$

меняет свой вид.

В более общем случае преобразование $Q = bq, P = cp$, где b и c – произвольные отличные от нуля постоянные, также сохраняет вид гамильтоновых уравнений движения. При этом новый гамильтониан равен $K = bcH$. □

По сути дела этот пример описывает все отличия в разных определениях канонического преобразования.

Предложение 20.1.2. *Пусть при преобразовании координат $\varphi : \ q, p \mapsto Q, P$ фазового пространства $\mathbb{R}^{2N}$ гамильтоновы уравнения движения (20.13) сохраняют свой вид:*

$$\dot{Q}^i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q^i},$$

где $K(Q,P)$ – некоторый новый гамильтониан. Тогда

$$\oint_{\varphi(\gamma)} dQP = c \oint_{\gamma} dqp, \tag{20.51}$$

где $c \neq 0$ – некоторая постоянная.

Доказательство. Рассмотрим расширенное фазовое пространство $\mathbb{R}^{2N+1}$. Пусть γ – замкнутый контур в фазовом пространстве, соответствующий фиксированному моменту времени $t = \mathsf{const}$. Тогда в силу следствия из теоремы 20.1.3 интегралы

$$\oint_{\gamma} dqp, \qquad \oint_{\varphi(\gamma)} dQP$$

являются интегральными инвариантами фазовых потоков. По теореме Стокса интеграл по контуру γ преобразуется в интеграл по поверхности, натянутой на данный контур. Используя теорему 20.1.13, заключаем, что

$$\oint_{\varphi(\gamma)} dQP = c \oint_{\gamma} dqp,$$

где $c \neq 0$ – некоторая постоянная. □

Ясно, что множество преобразований, сохраняющих вид гамильтоновых уравнений, образует группу преобразований фазового пространства.

Следствие. Любое преобразование координат $q, p \mapsto Q, P$ фазового пространства, которое сохраняет форму гамильтоновых уравнений движения, является композицией преобразования φ_c (20.50) и некоторого канонического преобразования. □

Доказательство. При преобразовании координат $q, p \mapsto Q, P$ относительный интегральный инвариант Пуанкаре–Картана умножается на некоторую постоянную c (20.51). Представим это преобразование в виде двух последовательных преобразования координат:

$$\varphi_c : \ q, p \mapsto \tilde{q}, \tilde{p} \quad \text{и} \quad g : \ \tilde{q}, \tilde{p} \mapsto Q, P.$$

Поскольку при первом преобразовании инвариант умножается на ту же постоянную c, то при втором преобразовании координат он сохраняется,

$$g : \ \oint_{\varphi_c(\gamma)} d\tilde{q}\tilde{p} = \oint_{g\circ\varphi_c(\gamma)} dQP.$$

Ввиду произвольности контура γ это значит, что преобразование g является каноническим. □

Поскольку преобразование φ_c (20.50) является простой растяжкой координат, то канонические преобразования составляют нетривиальную и наиболее содержательную часть преобразований координат, сохраняющих вид гамильтоновых уравнений движения.

Сформулируем три критерия для канонических преобразований. С этой целью введем новое понятие. Обозначим координаты фазового пространства через $\{x^\alpha\} = \{q^i, p_i\}$, $\alpha = 1, \ldots, 2N$.

Определение. Пусть на фазовом пространстве $\mathbb{R}^{2N}$ с координатами $\{x^\alpha\} = \{q^i, p_i\}$ и канонической симплектической формой ϖ задано преобразование координат $\{x^\alpha\} \mapsto \{y^\alpha\} = \{Q^i(x), P_i(x)\}$. Назовем *скобкой Лагранжа* двух новых координат y^α и y^β для заданного набора функций $q^i(y^\alpha, y^\beta)$, $p_i(y^\alpha, y^\beta)$, определяющих преобразование координат, следующее выражение

$$\begin{aligned}[y^\alpha, y^\beta]_{\text{L}} &:= \varpi_{\delta\gamma}\frac{\partial x^\gamma}{\partial y^\alpha}\frac{\partial x^\delta}{\partial y^\beta} \\ &= \frac{\partial q^i}{\partial y^\alpha}\frac{\partial p_i}{\partial y^\beta} - \frac{\partial q^i}{\partial y^\beta}\frac{\partial p_i}{\partial y^\alpha}. \qquad \square\end{aligned} \tag{20.52}$$

$\square$

Скобка Лагранжа антисимметрична и определена для любой пары новых координат y^α и y^β.

Замечание. Если на фазовом пространстве просто заданы две дифференцируемые функции $F, G \in \mathcal{C}^1(\mathbb{R}^{2N})$, то скобка Лагранжа для них неопределена, т.к. не определены частные производные $\partial x^\alpha/\partial F$ и $\partial x^\alpha/\partial G$. $\square$

Теорема 20.1.16. *Для того, чтобы преобразование координат фазового пространства $q, p \mapsto Q, P$ было каноническим, необходимо и достаточно выполнения следующих условий для скобок Лагранжа новых координат:*

$$[Q^i, Q^j]_{\text{L}} = 0, \quad [P_i, P_j]_{\text{L}} = 0, \quad [Q^i, P_j]_{\text{L}} = \delta^i_j. \tag{20.53}$$

Доказательство. Подставляя в определение канонических преобразований (20.42) функции $q^i(Q, P)$, $p_i(Q, P)$, определяющие канонические преобразования, получим

$$\begin{aligned}dP_i \wedge dQ^i = &\frac{1}{2}dQ^l \wedge dQ^k \left(\frac{\partial q^i}{\partial Q^k}\frac{\partial p_i}{\partial Q^l} - \frac{\partial q^i}{\partial Q^l}\frac{\partial p_i}{\partial Q^k}\right) \\ &+ dP_l \wedge dQ^k \left(\frac{\partial q^i}{\partial Q^k}\frac{\partial p_i}{\partial P_l} - \frac{\partial q^i}{\partial P_l}\frac{\partial p_i}{\partial Q^k}\right) \\ &+ \frac{1}{2}dP_l \wedge dP_k \left(\frac{\partial q^i}{\partial P_k}\frac{\partial p_i}{\partial P_l} - \frac{\partial q^i}{\partial P_l}\frac{\partial p_i}{\partial P_k}\right).\end{aligned} \tag{20.54}$$

Отсюда следует первый критерий для канонических преобразований координат. $\square$

Таким образом, при каноническом преобразовании скобки Лагранжа для новых координат обязаны совпадать со скобкой Пуассона.

Матрицы Якоби для канонических преобразований обладают замечательным свойством, которое можно сформулировать в виде второго критерия канонических преобразований.

Теорема 20.1.17. *Преобразование координат фазового пространства является каноническим тогда и только тогда, когда матрица Якоби этого преобразования*

$$J_\alpha{}^\beta = \frac{\partial y^\beta}{\partial x^\alpha} = \begin{pmatrix} \dfrac{\partial Q}{\partial q} & \dfrac{\partial P}{\partial q} \\ \dfrac{\partial Q}{\partial p} & \dfrac{\partial P}{\partial p} \end{pmatrix} \tag{20.55}$$

является симплектической, $J_\alpha{}^\beta \in \mathbb{SP}(n, \mathbb{R})$.

Доказательство. Доказательство проводится прямой проверкой равенства

$$\varpi_{\alpha\beta} = J_\alpha{}^\gamma J_\beta{}^\delta \varpi_{\gamma\delta} \tag{20.56}$$

с учетом свойств (20.53). Вычисления проще проводить в блочно диагональном виде (20.55). □

Поскольку определитель симплектической матрицы равен единице (18.5), то форма объема фазового пространства сохраняется при каноническом преобразовании. Отсюда также следует теорема Лиувилля 20.1.12.

Третий критерий канонических преобразований формулируется в терминах скобок Пуассона.

Теорема 20.1.18. *Преобразование координат фазового пространства является каноническим тогда и только тогда, когда оно сохраняет скобки Пуассона*

$$[F, G]_{q,p} = [F, G]_{Q,P}. \tag{20.57}$$

где F и G – две произвольные дифференцируемые функции на фазовом пространстве, и скобки Пуассона вычислены, соответственно, в координатах q, p и Q, P с одинаковыми каноническими структурными функциями $\varpi^{-1\alpha\beta}$.

Доказательство. Канонические преобразования и только они сохраняют каноническую симплектическую форму. Сохранение этой формы эквивалентно сохранению матрицы $\varpi^{-1\alpha\beta}$, т.е. канонической пуассоновой структуры. □

Пример 20.1.15. Преобразование

$$Q^i = q^i, \qquad P_i = p_i + \frac{\partial F(q)}{\partial q^i}, \tag{20.58}$$

где $F(q) \in \mathcal{C}^2(\mathbb{R}^{\mathrm{N}})$ – произвольная функция от обобщенных координат q, является каноническим. Действительно, скобки Пуассона

$$[Q^i, Q^j] = 0, \qquad [Q^i, P_j] = \delta^i_j$$

остаются прежними. Нетрудно также проверить, что

$$[P_i, P_j] = \left[P_i, \frac{\partial F}{\partial q^j}\right] + \left[\frac{\partial F}{\partial q^i}, P_j\right] = 0.$$

Аналогично, преобразование

$$Q^i = q^i + \frac{\partial G(p)}{\partial p_i}, \qquad P_i = p_i,$$

где $G(p) \in \mathcal{C}^2(\mathbb{R}^{\mathrm{N}})$ – произвольная функция импульсов, также является каноническим □

Важность канонических преобразований сводится к следующему. Если удается найти такое каноническое преобразование, при котором гамильтониан упрощается настолько, что уравнения Гамильтона явно интегрируются, то тогда можно построить и решение исходной задачи. Соответствующие канонические преобразования зависят от конкретной задачи, и общего метода их нахождения не существует.

20.1.11 Производящие функции канонических преобразований

Широкий класс канонических преобразований можно описать на языке производящих функций, которые строятся следующим образом. Рассмотрим фазовое пространство $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \approx \mathbb{R}^{2\mathrm{N}}$ с координатами q, p и канонической пуассоновой структурой. Пусть 2N функций $Q^i(q,p)$ и $P_i(q,p)$ от 2N переменных задают каноническое преобразование. Тогда 1-форма

$$dqp - dQP = dS(q,p) \tag{20.59}$$

есть полный дифференциал.

Предположим, что в окрестности некоторой точки старые и новые координаты (q,Q) можно выбрать в качестве координат фазового пространства. Это значит, что

$$\det \frac{\partial(q,Q)}{\partial(q,p)} = \det \begin{pmatrix} \dfrac{\partial q}{\partial q} & \dfrac{\partial Q}{\partial q} \\ 0 & \dfrac{\partial Q}{\partial p} \end{pmatrix} = \det \frac{\partial Q}{\partial p} \neq 0. \tag{20.60}$$

Тогда функцию S можно локально выразить через эти координаты:

$$S(q,p) = S_1(q,Q).$$

Функция $S_1(q,Q)$ называется *производящей* функцией канонического преобразования. Эту функцию не следует рассматривать как функцию на фазовом пространстве, потому что координаты q и Q не являются канонически сопряженными. Производящая функция S_1 от 2N переменных задана в некоторой окрестности прямого произведения $\mathbb{R}^{\mathrm{N}}_q \times \mathbb{R}^{\mathrm{N}}_Q$.

Из определения производящей функции и уравнения (20.59) следует, что

$$p = \frac{\partial S_1}{\partial q}, \qquad P = -\frac{\partial S_1}{\partial Q}. \tag{20.61}$$

Конечно, эти формулы являются сокращенной записью равенств

$$p_i = \frac{\partial S_1}{\partial q^i}, \qquad P_i = -\frac{\partial S_1}{\partial Q^i}.$$

Полученные уравнения служат для нахождения канонического преобразования $Q(q,p)$, $P(q,p)$. Таким образом, каждая производящая функция $S_1(q,Q)$ такая, что

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

определяет некоторое каноническое преобразование по формулам (20.61). Последнее неравенство следует из первого уравнения (20.61) и условия (20.60).

Для того, чтобы получить гамильтониан $K(Q,P)$ для новых канонических переменных, необходимо просто подставить в старое выражение $H(q,p)$ функции $q = q(Q,P)$ и $p = p(Q,P)$:

$$K(Q,P) = H\Big(q(Q,P), p(Q,P)\Big). \tag{20.62}$$

Пример 20.1.16. Пусть $S_1 = q^i Q_i$. Тогда $Q_i = p_i$ и $P^i = -q^i$. Тем самым координаты и импульсы меняются местами. Этот пример показывает, что в каноническом формализме координаты и импульсы играют совершенно равноправную роль. □

Пусть гамильтониан не зависит от времени явно. Покажем, что если производящая функция $S_1(q, Q)$ сама удовлетворяет уравнению Гамильтона–Якоби по переменным q, то уравнения движения интегрируются в квадратурах. Для этого заметим, что если гамильтониан системы зависит только от новых координат $H = K(Q)$, то канонические уравнения

$$\dot{Q} = 0, \qquad \dot{P} = -\frac{\partial K}{\partial Q}$$

просто интегрируются

$$Q = Q_0, \qquad P = P_0 - t\,\frac{\partial K}{\partial Q}\bigg|_{Q_0}.$$

Теперь будем искать производящую функцию $S_1(q, Q)$ этого канонического преобразования. Из первого условия (20.61) следует, что такая производящая функция должна удовлетворять уравнению

$$H\left(q, \frac{\partial S_1}{\partial q}\right) = K(Q). \tag{20.63}$$

То есть при каждом фиксированном значении Q производящая функция $S_1(q, Q)$ должна удовлетворять укороченному уравнению Гамильтона–Якоби (20.37). Верно также и обратное утверждение.

Теорема 20.1.19 (Якоби). *Если найдено решение $S_1(q, Q)$ укороченного уравнения Гамильтона–Якоби (20.63), зависящее от* N *параметров Q^i и такое, что*

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

то канонические уравнения (20.24) решаются в квадратурах, причем функции $Q(q, p)$, определяемые уравнениями $\partial S_1/\partial q = p$, являются N *первыми интегралами уравнений Гамильтона.*

Доказательство. См., например, [**?**]. □

Теорема Якоби сводит решение канонических уравнений к нахождению полного интеграла укороченного уравнения Гамильтона–Якоби. Может показаться удивительным, что сведение более простого к более сложному доставляет эффективный метод решения конкретных задач. Между тем оказывается, что это – наиболее мощный из существующих в настоящее время методов интегрирования уравнений движения, и многие задачи, решенные методом Якоби, вообще не поддаются решению другими способами.

В качестве дополнительного аргумента производящей функции можно рассматривать время, $S_1 = S_1(q, Q, t)$. При этом все предыдущие формулы остаются в силе за исключением выражения (20.62) для нового Гамильтониана. Оно получает дополнительное слагаемое,

$$K = H + \frac{\partial S_1}{\partial t}, \tag{20.64}$$

что следует из выражения для действия (**??**) в расширенном конфигурационном пространстве.

Если после канонического преобразования новый гамильтониан тождественно равен нулю, то уравнения движения, очевидно, интегрируются. Для этого необходимо выполнение уравнения

$$\frac{\partial S_1}{\partial t} + H(q,p,t) = 0.$$

Учитывая выражение для импульсов через производящую функцию (20.61) получаем, что производящая функция в этом случае должна удовлетворять уравнению Гамильтона–Якоби для всех значений Q. Тем самым доказана

Теорема 20.1.20 (Якоби). *Если найдено решение $S_1(q,Q,t)$ уравнения Гамильтона–Якоби (20.34), зависящее от* N *параметров Q^i и такое, что*

$$\det \frac{\partial^2 S_1}{\partial q^i \partial Q^j} \neq 0,$$

то канонические уравнения (20.24) решаются в квадратурах, причем координаты $Q^i = \mathsf{const}$ и импульсы $P_i = -\partial S_1/\partial Q^i = \mathsf{const}$ дают 2N *первых интегралов уравнений Гамильтона.*

Частное решение уравнения Гамильтона–Якоби, зависящее от N + 1 параметра (по числу независимых переменных), называется *полным интегралом*. Функция от N + 1 параметра $S_1(q,Q,t) + \mathsf{const}$, где S_1 – решение уравнения Гамильтона–Якоби из предыдущей теоремы является полным интегралом уравнения Гамильтона–Якоби. В том случае, когда для уравнения Гамильтона–Якоби можно определить общее решение, оно будет зависеть от произвольных функций. Это значит, что полный интеграл уравнения Гамильтона–Якоби дает лишь незначительную часть всех решений.

Если известен полный интеграл уравнения Гамильтона–Якоби $S_1(q,Q,t) + \mathsf{const}$, то для нахождения траектории частицы, проходящей через заданную точку фазового пространства в начальный момент времени, необходимо решить дополнительные уравнения:

$$\frac{\partial S}{\partial Q^i} = c_i, \tag{20.65}$$

где c_i – некоторые постоянные, определяемые начальным положением частицы.

Продолжим изучение производящих функций. Выше был рассмотрен случай, когда производящая функция зависит от старых и новых координат. Аналогичным образом можно показать, что канонические преобразования генерируются производящими функциями $S_2(q,P,t)$, зависящими от старых координат и новых импульсов. Если

$$\det \frac{\partial^2 S_2}{\partial q^i \partial P_j} \neq 0,$$

то каноническое преобразование задается формулами

$$p = \frac{\partial S_2}{\partial q}, \qquad Q = \frac{\partial S_2}{\partial P}, \qquad K = H + \frac{\partial S_2}{\partial t}. \tag{20.66}$$

Переход от производящей функции $S_1(q,Q)$ к $S_2(q,P)$ является преобразованием Лежандра по переменной Q при этом

$$S_2(q,P,t) = S_1(q,Q,t) + PQ.$$

Пример 20.1.17. Каноническое преобразование, генерируемое производящей функцией $S_2 = qP := q^i P_i$ является тождественным. □

Пример 20.1.18. Пусть $S_2 = q^i S_i{}^j P_j$, где $S_i{}^j$ – произвольная невырожденная матрица. Тогда $P_i = S_i^{-1j} p_j$ и $Q^i = q^j S_j{}^i$. То есть это преобразование генерирует линейное однородное преобразование фазового пространства. □

Пример 20.1.19. Выберем производящую функцию бесконечно малых канонических преобразований в виде

$$S_2 = qP + \epsilon S(q, P, \epsilon), \qquad \epsilon \ll 1,$$

где $S(q, P, \epsilon)$ – произвольная функция, зависящая от параметра ϵ. Тогда

$$Q = q + \epsilon \frac{\partial S}{\partial P}, \qquad p = P + \epsilon \frac{\partial S}{\partial q}.$$

Отсюда следует, что бесконечно малые канонические преобразования удовлетворяют каноническим уравнениям

$$\left.\frac{dQ}{d\epsilon}\right|_{\epsilon=0} = \frac{\partial H}{\partial p}, \qquad \left.\frac{dP}{d\epsilon}\right|_{\epsilon=0} = -\frac{\partial H}{\partial q}$$

с функцией Гамильтона $H(q, p) = S(q, p, 0)$. Тем самым гамильтониан произвольной механической системы можно рассматривать, как генератор бесконечно малых канонических преобразований по времени $t = \epsilon$, которые происходят по мере эволюции механической системы. □

На самом деле движение частицы в фазовом пространстве $q(t)$, $p(t)$ можно рассматривать как последовательность канонических преобразований, параметризующихся одним параметром t. Допустим, что решение канонических уравнений движения определено для всех $t \in \mathbb{R}$. Тогда канонические уравнения движения (20.13) определяют однопараметрическую группу преобразований фазового пространства. Для каждого момента времени, по-определению, выполнены канонические коммутационные соотношения (20.14). Следовательно, для каждого момента времени t отображение фазового пространства

$$\mathbb{R}^{2N} \ni \Big(q(0), p(0)\Big) \mapsto \Big(q(t), p(t)\Big) \in \mathbb{R}^{2N}$$

представляет собой каноническое преобразование.

В заключение рассмотрим еще два вида производящих функций, часто встречающихся в приложениях. Если производящая функция $S_3(p, Q, t)$ зависит от старых импульсов и новых координат и при этом

$$\det \frac{\partial^2 S_3}{\partial p_i \partial Q^j} \neq 0,$$

то каноническое преобразование задается формулами:

$$q = -\frac{\partial S_3}{\partial p}, \qquad P = -\frac{\partial S_3}{\partial Q}, \qquad K = H + \frac{\partial S_3}{\partial t}. \tag{20.67}$$

При этом производящая функция S_3 является преобразованием Лежандра функции S_1 по переменной q:

$$S_3(p, Q, t) = S_1(q, Q, t) - pq.$$

Пример 20.1.20. Каноническое преобразование, генерируемое производящей функцией $S_3 = -p_i Q^i$, является тождественным. □

Производящая функция полезна в следующем случае. Допустим, что нам заранее известно, как удобно выбрать координаты в конфигурационном пространстве, т.е. заданы функции $q = q(Q)$. Например, если конфигурационное пространство инвариантно относительно действия некоторой группы преобразований, то новые координаты Q удобно выбрать в соответствии с симметрией задачи. Тогда производящая функция $S_3 = -p_i q^i(Q)$ дает выражения для новых обобщенных импульсов P_i, сопряженных к координатам Q^i.

Точечные преобразования, рассмотренные в разделе 20.1.3, относятся именно к этому типу. Точечные преобразования – это все преобразования координат, которые допустимы в лагранжевой механике. Рассмотренный выше пример показывает, что точечные преобразования составляют лишь небольшой класс канонических преобразований. Все множество канонических преобразований намного шире, и это является причиной большей гибкости канонического формализма по сравнению с лагранжевым.

Продолжим изучение производящих функций. Если производящая функция $S_4(p, P, t)$ зависит от старых и новых импульсов и

$$\det \frac{\partial^2 S_4}{\partial p_i \partial P_j} \neq 0,$$

то каноническое преобразование задается формулами:

$$q = -\frac{\partial S_4}{\partial p}, \qquad Q = \frac{\partial S_4}{\partial P}, \qquad K = H + \frac{\partial S_4}{\partial t}. \tag{20.68}$$

При этом производящая функция S_4 является двойным преобразованием Лежандра функции $S_1(q, Q)$ по переменным q и Q:

$$S_4(p, P, t) = S_1(q, Q, t) - pq + PQ.$$

Пример 20.1.21. Каноническое преобразование, генерируемое производящей функцией $S_4 = p_i P^i$, меняет местами координаты и импульсы: $Q_i = p_i$, $P^i = -q_i$. Это преобразование совпадает с каноническим преобразованием из примера 20.1.16. □

Пример 20.1.22 (Гармонический осциллятор). Продемонстрируем сложные построения последних разделов на примере гармонического осциллятора. Пусть задана функция Лагранжа

$$L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2,$$

где $\omega = \mathsf{const}$ – собственная частота гармонического осциллятора. Конфигурационное пространство представляет собой прямую $q \in \mathbb{R}$. Импульс, сопряженный координате q, равен скорости

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q}.$$

Преобразование Лежандра приводит к положительно определенному гамильтониану:

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2.$$

Фазовым пространством для гармонического осциллятора является евклидова плоскость $(q,p)\in\mathbb{R}^2$ с канонической пуассоновой структурой. Соответствующие уравнения движения имеют вид

$$\left.\begin{aligned}\dot q &= \;\; p,\\ \dot p &= -\omega^2 q,\end{aligned}\right\} \quad\Leftrightarrow\quad \ddot q+\omega^2 q=0. \tag{20.69}$$

Общее решение этих уравнений параметризуется двумя произвольными постоянными, амплитудой A_0 и фазой φ,

$$\begin{aligned} q &= \;\; A_0\cos(\omega t+\varphi),\\ p &= -A_0\omega\sin(\omega t+\varphi).\end{aligned} \tag{20.70}$$

Фаза φ соответствует выбору начала отсчета времени. Полученное решение соответствует решению задачи Коши с начальными данными

$$\begin{aligned} q_0 &= \;\; A_0\cos\varphi,\\ p_0 &= -A_0\omega\sin\varphi.\end{aligned}$$

В терминах начальных условий общее решение (20.70) перепишется в виде

$$\begin{aligned} q &= \;\; q_0\cos\omega t+\frac{p_0}{\omega}\sin\omega t,\\ p &= -\omega q_0\sin\omega t+p_0\cos\omega t.\end{aligned}$$

Траекториями гармонического осциллятора в фазовом пространстве являются эллипсы. Как видим, эволюция во времени представляет собой эллиптическое вращение фазового пространства $(q_0,p_0)\mapsto\big(q(t),p(t)\big)$. Если $\mathbb{V}_0\subset\mathbb{R}^2$ – произвольная ограниченная область фазового пространства, то в соответствии с теоремой Лиувилля (20.1.12) объем этой области сохраняется во времени:

$$\int_{\mathbb{V}_0} dq_0dp_0=\int_{\mathbb{V}(t)} dqdp=\int_{\mathbb{V}(t)} dq_0dp_0,$$

т.к. якобиан преобразования координат равен единице,

$$\det\frac{\partial(q,p)}{\partial(q_0,p_0)}=1.$$

Поскольку гамильтониан не зависит от времени явно, то энергия на каждой траектории постоянна $E=H=\mathsf{const}$.

Найдем функцию действия $S(q,t)$. Чтобы вычислить интеграл (20.32) необходимо исключить скорости $\dot q$ из лагранжиана. Поскольку интегрирование ведется вдоль траектории частицы, то в лагранжиан необходимо подставить

$$\dot q=p=\sqrt{2E-\omega^2q^2}.$$

В этом выражении E рассматривается, как некоторая постоянная. После несложных вычислений получаем функцию действия

$$\begin{aligned} S(q,t) &= -\int dtE+\int dq\sqrt{2E-\omega^2q^2}\\ &= -Et+\frac{E}{\omega}\arcsin\frac{\omega q}{\sqrt{2E}}+\frac{q}{2}\sqrt{2E-\omega^2q^2}\end{aligned}$$

с точностью до несущественной постоянной, определяемой начальными данными. Нетрудно проверить, что

$$\frac{\partial S}{\partial q} = \sqrt{2E - \omega^2 q^2} = p.$$

Поэтому уравнение Гамильтона–Якоби (20.34) выполняется. Укороченное действие имеет вид

$$W(q, E) = \frac{E}{\omega} \operatorname{arcsin} \frac{\omega q}{\sqrt{2E}} + \frac{q}{2}\sqrt{2E - \omega^2 q^2}.$$

Таким образом, нам известно укороченное действие, зависящее от параметра E, причем

$$\frac{\partial^2 W}{\partial q \partial E} \neq 0.$$

Поэтому его можно выбрать в качестве производящей функции канонического преобразования $q, p \to Q, P$

$$S_2(q, P) = \frac{P}{\omega} \operatorname{arcsin} \frac{\omega q}{\sqrt{2P}} + \frac{q}{2}\sqrt{2P - \omega^2 q^2}.$$

Воспользовавшись формулами (20.66), получаем явный вид канонических преобразований

$$q = \frac{\sqrt{2P}}{\omega} \sin(\omega Q),$$
$$p = \sqrt{2P} \cos(\omega Q).$$

Отсюда вытекает связь дифференциалов:

$$dq = \sqrt{2P} \cos(\omega Q) dQ + \frac{1}{\omega\sqrt{2P}} \sin(\omega Q) dP,$$
$$dp = -\omega\sqrt{2P} \sin(\omega Q) dQ + \frac{1}{\sqrt{2P}} \cos(\omega Q) dP.$$

Теперь нетрудно убедиться в том, что построенное преобразование действительно является каноническим

$$dp \wedge dq = dP \wedge dQ.$$

В новых переменных гамильтониан имеет вид $H = P$ и уравнения движения

$$\dot{Q} = P, \qquad \dot{P} = 0$$

тривиально интегрируются. Таким образом укороченная функция действия для гармонического осциллятора является производящей функцией к переменным "действие-угол". Фазовым пространством в переменных "действие-угол" является полуплоскость $Q \in \mathbb{R}$, $P \geq 0$, а траекториями – прямые линии $P = \mathsf{const}$. □

В настоящем разделе мы рассмотрели широкий класс канонических преобразований, которые генерируются четырьмя видами производящих функций: $S_1(q, Q, t)$, $S_2(q, P, t)$, $S_3(p, Q, t)$ и $S_4(p, P, t)$. Существуют также производящие функции более общего вида, которые мы сейчас опишем.

Предложение 20.1.3. *Пусть заданы две системы координат q, p и Q, P в фазовом пространстве пространстве $\mathbb{R}^{2N}$ с канонической пуассоновой структурой. Тогда среди $4N$ координатных функций q, p, Q, P всегда можно выбрать $2N$ независимых функций так, чтобы среди них не было ни одной пары канонически сопряженных функций q^i, p_i или Q^i, P_i.*

Доказательство. См., например, [**?**]. □

Из данного утверждения следует, что в качестве координат на фазовом пространстве $\mathbb{R}^{2\mathrm{N}}$ всегда можно выбрать набор функций

$$q^1,\dots,q^L,\ p_{L+1},\dots,p_{\mathrm{N}},\ Q^1,\dots,Q^M,\ P_{M+1},\dots,P^{\mathrm{N}},\qquad 0\le L,M\le \mathrm{N},$$

состоящий из N старых и N новых координат. Введем обозначения:

$$\{z^i\}:=\{q^1,\dots,q^L,p_{L+1},\dots,p_{\mathrm{N}}\},$$
$$\{Z^i\}:=\{Q^1,\dots,Q^M,P_{M+1},\dots,P^{\mathrm{N}}\}.$$

Допустим, что на фазовом пространстве задана функция $S(z,Z)$ такая, что

$$\det\frac{\partial^2 S}{\partial z^i\partial Z^j}\neq 0. \tag{20.71}$$

Тогда нетрудно проверить, что функция $S(z,Z)$ определяет каноническое преобразование:

$$\begin{aligned}
p_i &= \frac{\partial S}{\partial q^i}, && i=1,\dots,L,\\
q^i &= -\frac{\partial S}{\partial p_i}, && i=L+1,\dots,\mathrm{N},\\
P_i &= -\frac{\partial S}{\partial Q^i}, && i=1,\dots,M,\\
Q^i &= \frac{\partial S}{\partial P_i}, && i=M+1,\dots,\mathrm{N}.
\end{aligned}$$

Рассмотренное каноническое преобразование включает в себя четыре предыдущих при $L=0,\mathrm{N}$ и $M=0,\mathrm{N}$. В общем случае производящая функция $S(z,Z)$ зависит от старых координат и импульсов и новых координат и импульсов. При этом среди них нет ни одной пары канонически сопряженных переменных. Примеры 20.1.16 и 20.1.21 показывают, что различные производящие функции могут приводить к одинаковым каноническим преобразованиям. Во всех случаях от производящей функции требуется отличие от нуля определителя (20.71). Поэтому производящие функции описывают широкий класс канонических преобразований, но не все.

Пример 20.1.23. В примере 20.1.15 из предыдущего раздела было рассмотрено каноническое преобразование, которое не охватывается производящими функциями, рассмотренными в настоящем разделе. Действительно, для преобразования (20.58) формула (20.59) принимает вид

$$-dq\frac{\partial F(q)}{\partial q}=dS(q,p)=dq\frac{\partial S}{\partial q}+dp\frac{\partial S}{\partial p}.$$

Откуда следуют равенства:

$$\frac{\partial S}{\partial q}=-\frac{\partial F}{\partial q},\qquad \frac{\partial S}{\partial p}=0.$$

Таким образом, для канонических преобразований вида (20.58) функция $S(q,p)$ имеет вид

$$S(q,p)=-F(q)+\mathsf{const}$$

и не относится ни к одному типу производящих функций настоящего раздела. □

20.1.12 Разделение переменных в уравнении Гамильтона–Якоби

Теорема Якоби 20.1.20 утверждает, что если найден полный интеграл уравнения Гамильтона–Якоби, то канонические уравнения решаются в квадратурах. По-видимому, все известные в настоящее время полные интегралы могут быть найдены путем *разделения переменных* в уравнении Гамильтона–Якоби. Суть метода состоит в следующем.

Допустим, что одна из координат – обозначим ее q^{N} – и соответствующая ей производная $\partial S/\partial q^{\mathrm{N}}$ входят в уравнение Гамильтона–Якоби только в виде комбинации $f(q^{\mathrm{N}}, \partial S/\partial q^{\mathrm{N}})$, где f – некоторая функция двух переменных, не зависящая от других координат, соответствующих производных действия и времени. Тогда будем искать решение уравнения Гамильтона–Якоби (20.34) в виде

$$S(q,t) = \check{S}(\check{q},t) + S_{\mathrm{N}}(q^{\mathrm{N}}),$$

где $\check{S}(\check{q},t)$ – функция оставшихся переменных $\check{q} = q^1, \ldots, q^{\mathrm{N}-1}$ и времени, а $S_{\mathrm{N}}(q^{\mathrm{N}})$ зависит только от отделяемой координаты q^{N}. Тогда уравнение Гамильтона–Якоби принимает вид

$$\frac{\partial \check{S}}{\partial t} + H\left(\check{q}, \frac{\partial \check{S}}{\partial \check{q}}, f\left(q^{\mathrm{N}}, \frac{dS_{\mathrm{N}}}{dq^{\mathrm{N}}}\right), t\right) = 0. \tag{20.72}$$

Допустим, что решение этого уравнения найдено. Тогда, после подстановки найденного решения в уравнение (20.72) мы получаем тождество, справедливое для всех значений координат. При изменении координаты q^{N} меняется только функция f. Поэтому для выполнения уравнения (20.72) необходимо, чтобы функция f сама была равна постоянной, которую мы обозначим Q^{N}, для произвольного решения. Таким образом, уравнение Гамильтона–Якоби (20.72) распадается на два уравнения:

$$f\left(q^{\mathrm{N}}, \frac{dS_{\mathrm{N}}}{dq^{\mathrm{N}}}\right) = Q^{\mathrm{N}}, \tag{20.73}$$

$$\frac{\partial \check{S}}{\partial t} + H\left(\check{q}, \frac{\partial \check{S}}{\partial \check{q}}, Q^{\mathrm{N}}, t\right) = 0. \tag{20.74}$$

Первое уравнение представляет собой обыкновенное дифференциальное уравнение первого порядка, из которого можно определить функцию $S_{\mathrm{N}} = S_{\mathrm{N}}(q^{\mathrm{N}}; Q^{\mathrm{N}})$. Если функция f достаточно проста, то это уравнение иногда удается проинтегрировать явно. Постоянная Q^{N} является первым интегралом исходного уравнения Гамильтона–Якоби. После этого остается решить оставшееся уравнение (20.74), которое также имеет вид уравнения Гамильтона–Якоби, но для системы с меньшим числом степеней свободы. Эта система описывается гамильтонианом, который получается из исходного путем замены функции f для отделенной степени свободы на некоторую константу. Единственное ограничение на постоянную Q^{N} состоит в том, чтобы уравнение (20.73) имело решение.

Иногда этот процесс отделения переменных можно продолжить. Если таким образом можно разделить все N координат, то тогда полный интеграл уравнения Гамильтона–Якоби находится в квадратурах. В этом случае мы говорим, что уравнение Гамильтона–Якоби допускает *разделение переменных*.

Замечание. Разделение переменных в уравнении Гамильтона–Якоби для одной и той же механической системы возможно не во всякой системе координат конфигурационного пространства. Как правило системы координат, в которых переменные

делятся полностью или частично связаны с симметрией системы. Процедура нахождения такой системы координат, в которой переменные для интегрируемой системы разделяются, представляет собой отдельную задачу. □

Допустим, что механическая система консервативна и переменные делятся. Тогда полный интеграл уравнения Гамильтона–Якоби имеет вид

$$S(q,t) = -E(Q^1,\dots,Q^{\mathrm{N}})t + \sum_{k=1}^{\mathrm{N}} S_k(q^k;Q^k,\dots,Q^{\mathrm{N}}), \tag{20.75}$$

где каждая функция S_k зависит от одной координаты q^k, а энергия E как функция первых интегралов $Q^1,\dots,Q^{\mathrm{N}}$ получается подстановкой укороченного действия

$$W = \sum_{k=1}^{\mathrm{N}} S_k(q^k,Q^k,\dots,Q^{\mathrm{N}})$$

в укороченное уравнение Гамильтона–Якоби (20.37).

Пример 20.1.24 (Циклические координаты). Пусть q^{N} – циклическая координата, т.е. гамильтониан от нее вообще не зависит. В этом случае функция $f(q^{\mathrm{N}},\partial S/\partial q^{\mathrm{N}})$ сводится просто к $\partial S/\partial q^{\mathrm{N}}$. Тогда из уравнения (20.73) следует, что $S_{\mathrm{N}} = q^{\mathrm{N}}Q^{\mathrm{N}}$ (никакого суммирования). Тогда функция действия принимает вид

$$S(q,t) = \check{S}(\check{q},t) + q^{\mathrm{N}}Q^{\mathrm{N}}.$$

В этом случае первый интеграл $Q_{\mathrm{N}} = \partial S/\partial q^{\mathrm{N}} = p_{\mathrm{N}}$ имеет смысл обобщенного импульса, сопряженного циклической координате q^{N}.

Отметим, что если время t рассматривать, как обобщенную координату, то отделение времени в виде слагаемого $-Et$ в функции действия (20.36) для консервативной системы соответствует методу разделения переменных для циклической координаты t. При этом энергия E сохраняется и соответствует импульсу для обобщенной координаты t. □

Разделение переменных для консервативной системы возможно также в следующем более общем случае. Пусть гамильтониан механической системы имеет вид

$$H(q,p) = \frac{f_1(q^1,p_1) + \ldots + f_{\mathrm{N}}(q^{\mathrm{N}},p_{\mathrm{N}})}{g_1(q^1,p_1) + \ldots + g_{\mathrm{N}}(q^{\mathrm{N}},p_{\mathrm{N}})},$$

где $f_i(q^i,p_i)$ и $g_i(q^i,p_i)$ – некоторые функции только от указанных аргументов. Предыдущий случай возникает, например, при $g_i = \mathsf{const}$, $i = 1,\dots,\mathrm{N}$, $\sum_i g_i \neq 0$. Будем искать решение уравнения Гамильтона–Якоби в виде

$$S(q,t) = -Et + \sum_{i=1}^{\mathrm{N}} W_i(q^i).$$

Тогда укороченное уравнение Гамильтона–Якоби (20.37) примет вид

$$\sum_{i=1}^{\mathrm{N}} \left[f_i\left(q^i, \frac{dW_i}{dq_i}\right) - E g_i\left(q^i, \frac{dW_i}{dq_i}\right)\right] = 0.$$

Поскольку координаты q_i меняются независимо, то каждое слагаемое в данной сумме должно быть равно некоторой постоянной,

$$f_i\left(q^i, \frac{dW_i}{dq_i}\right) - E g_i\left(q^i, \frac{dW_i}{dq_i}\right) = c_i. \tag{20.76}$$

При этом не все постоянные c_i являются независимыми, т.к. должно быть выполнено равенство

$$c_1 + \ldots + c_{\text{N}} = 0.$$

Допустим, что все уравнения (20.76) можно разрешить относительно производных:

$$\frac{dW_i}{dq^i} = F_i(q^i, c_i).$$

Тогда полный интеграл уравнения Гамильтона–Якоби равен

$$S(q,t) = -Et + \sum_{i=1}^{\text{N}} \int dq^i F_i(q^i, c_i, E).$$

Пример 20.1.25 (**Задача Кеплера**). Рассмотрим частицу массы m, которая движется по плоскости с полярными координатами r, φ под действием центральной силы с потенциалом $U(r)$. Гамильтониан такой системы имеет вид

$$H(r,\varphi) = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) + U(r). \tag{20.77}$$

К этой задаче сводится задача движения двух тел, находящихся под действием гравитационного взаимодействия с потенциалом $U(r) = -m_1 m_2/r$ после отделения движения центра масс. Поскольку гамильтониан не зависит от угла, то координата φ является циклической, и переменные делятся. Поэтому функция действия имеет вид

$$S = -Et + W(r, E, M) + \varphi M, \tag{20.78}$$

где $p_\varphi = M = \mathsf{const}$ – сохраняющийся угловой импульс частицы (момент количества движения). В этом случае уравнение Гамильтона–Якоби сводится к одному обыкновенному дифференциальному уравнению первого порядка

$$\left(\frac{dW}{dr}\right)^2 = 2m(E-U) - \frac{M^2}{r^2}. \tag{20.79}$$

Произвол в решении данного уравнения фиксируется положением частицы в начальный момент времени. В соответствии с общими утверждениями, функция действия (20.78), где W – решение уравнения (20.79), зависит от двух постоянных: энергии E и момента количества движения M, и представляет собой полный интеграл уравнения Гамильтона–Якоби. Таким образом, задача двух тел, взаимодействующих посредством центральных сил с произвольным потенциалом $U(r)$, решается в квадратурах. Чтобы найти траекторию частицы, проходящую через заданную точку фазового пространства в начальный момент времени, в соответствии с теоремой Якоби (20.1.20) необходимо решить дополнительные уравнения:

$$\frac{\partial S}{\partial M} = \frac{\partial W}{\partial M} + \varphi = c_1, \qquad \frac{\partial S}{\partial E} = -t + \frac{\partial W}{\partial E} = c_2,$$

где $c_{1,2}$ – некоторые постоянные (20.66), определяемые начальным положением частицы.

Заметим, что переменные в задаче Кеплера на плоскости делятся в полярных координатах (в трехмерном пространстве – в сферических), которые связаны со сферической симметрией потенциала взаимодействия. В декартовой системе координат переменные в уравнении Гамильтона–Якоби не делятся. □

20.2 Гамильтонова динамика частиц со связями

В современной математической физике большую роль играют модели, инвариантные относительно действия локальных групп преобразований, которые принято называть калибровочными. С точки зрения гамильтонова формализма такие модели соответствуют системам со связями. В калибровочных моделях связи возникают не как дополнительные условия на канонические переменные, наложенные извне, а при переходе от вырожденного инвариантного лагранжиана к гамильтониану. Основы гамильтонова формализма для калибровочных моделей были заложены П. Дираком на примере точечных частиц [?, ?, ?]. Обобщение этого формализма на теорию поля представляет значительные трудности, которые мы кратко обсудим в разделе ??. Несмотря на важность, эта задача до сих пор не решена. В теории поля до настоящего времени, как правило, используют методы, развитые для точечных частиц. В простейших случаях это приводит к разумным результатам.

В последующих разделах мы обсудим вопрос о гамильтоновой динамике со связями для систем точечных частиц, оставив в стороне важную задачу перехода от вырожденных калибровочно инвариантных лагранжианов к гамильтонианам, которая подробно рассматривается в монографиях [?, ?, ?]. Квантование калибровочных моделей с помощью функционального интеграла в фазовом пространстве было развито в [?, ?].

20.2.1 Связи в гамильтоновом формализме

Рассмотрим систему N точечных частиц с действием (20.7). Предположим, что фазовое пространство представляет собой дифференцируемое многообразие $\mathbb{N}$ класса $\mathcal{C}^k$, $\dim \mathbb{N} = 2N$. Для простоты предположим, что фазовое пространство топологически тривиально $\mathbb{N} \approx \mathbb{R}^{2N}$ и покрывается одной картой. Координаты на фазовом пространстве $\mathbb{N}$, как и раньше, обозначим $(q,p) = (q^1, \dots, q^N, p_1, \dots, p_N)$. Мы предполагаем, что на фазовом пространстве задана каноническая пуассонова структура $[q^i, p_j] = \delta^i_j$.

Пусть на канонические переменные наложены *связи* в виде 2M алгебраических условий:

$$\Phi^\mu(q,p) = 0, \qquad \Phi^\mu \in \mathcal{C}^k(\mathbb{N}), \quad \mu = 1, \dots, 2M < 2N. \tag{20.80}$$

Мы пишем индекс связи вверху, поскольку в дальнейшем связи будут частью новых координатных функций. Условия (20.80) представляют собой систему уравнений на канонические переменные. Будем считать, что все Φ^μ функционально независимы, т.е. ранг матрицы $\partial\Phi^\mu/\partial x^\alpha$, где $x = (q,p)$, максимален и равен 2M в окрестности всех точек, координаты которых удовлетворяют системе уравнений (20.80). Если все связи принадлежат классу $\mathcal{C}^k$, то их совокупность, согласно теореме 3.10.4, определяет подмногообразие

$$\mathbb{M} = \{x \in \mathbb{N} : \quad \Phi = 0\}$$

того же класса гладкости и размерности $\dim \mathbb{M} = 2(N - M)$. Это подмногообразие называется *поверхностью связей*, которую будем обозначать равенством $\Phi = 0$ без индекса.

Замечание. Вообще говоря, $\Phi^\mu(q,p)$ не являются функциями на фазовом пространстве, т.к. могут иметь тензорные индексы. Например, в электродинамике первичная связь имеет вид $p_0 = 0$. □

Условие функциональной независимости связей накладывает жесткие ограничения на вид функций Φ^μ.

Пример 20.2.1. Пусть на фазовом пространстве $\mathbb{N}$ задана одна связь $\Phi(q,p)=0$. Условие функциональной независимости означает, что ранг матрицы $\partial_\alpha\Phi$, которая в рассматриваемом случае представляет собой строку, равен единице, т.е. по крайней мере одна из частных производных в каждой точке поверхности связей $\mathbb{M}$ отлична от нуля. Уравнение $\Phi^n=0$ при $0<n<1$ и $n>1$ определяет ту же поверхность связей $\mathbb{M}\subset\mathbb{N}$. Однако связь $\Phi^n=0$ не является функционально независимой. Действительно,

$$\partial_\alpha\Phi^n|_{\Phi=0}=n\Phi^{n-1}\partial_\alpha\Phi|_{\Phi=0}=\begin{cases}0, & n>1,\\ \infty, & 0<n<1.\end{cases}$$

Поэтому ранг матрицы $\partial_\alpha\Phi^n$ либо равен нулю, либо неопределен на поверхности связей, и связь $\Phi^n=0$ не является функционально независимой на $\mathbb{M}\subset\mathbb{N}$, хотя и определяет ту же поверхность связей. □

Рассмотренный пример показывает, что условие функциональной независимости связей означает не только определение поверхности связей $\mathbb{M}\subset\mathbb{N}$, но и возможность выбора функций Φ^μ в качестве трансверсальных координат к $\mathbb{M}$ в некоторой окрестности поверхности связей.

Уравнения, задающие поверхность связей, определены неоднозначно. Например, невырожденные линейные комбинации исходных связей

$$\Psi^\mu(q,p)=\Phi^\nu(q,p)S_\nu{}^\mu(q,p), \tag{20.81}$$

где матрица $S_\nu{}^\mu$ невырождена на всем фазовом пространстве $\mathbb{N}$, определяют ту же поверхность связей $\mathbb{M}\subset\mathbb{N}$ и функционально независимы.

Определение. Системы связей $\Phi=0$ и $\Psi=0$ называются *эквивалентными*, если они связаны между собой невырожденным линейным преобразованием (20.81) во всем фазовом пространстве. □

Замечание. Если потребовать выполнения равенства (20.81) не во всем фазовом пространстве, а только на поверхности связей, то этого недостаточно для эквивалентности. Действительно, пусть поверхность связей $\mathbb{M}\subset\mathbb{N}$ определена уравнениями $\Phi=0$. Определим новые связи $\Psi=0$ на поверхности связей с помощью некоторой невырожденной матрицы S, заданной на $\mathbb{M}$. После этого продолжим функции Ψ на все фазовое пространство таким образом, чтобы уравнения $\Psi=0$ определяли некоторое подмногообразие вида $\mathbb{M}\cup\mathbb{U}\subset\mathbb{N}$, где $\mathbb{U}$ – некоторое собственное подмногообразие $\mathbb{N}$ такое, что $\mathbb{U}\cap\mathbb{M}=\emptyset$. Никаких препятствий для этого нет. Таким образом, для новых связей условие (20.81) будет выполнено на $\mathbb{M}$, однако условия $\Psi=0$ выделяют в фазовом пространстве $\mathbb{N}$ другое подмногообразие. Это означает, что матрица S будет вырождена на $\mathbb{U}$. □

Предложение 20.2.1. *Любая функция $f(q,p)\in\mathcal{C}^\infty(\mathbb{N})$, обращающаяся в нуль на поверхности связей $\mathbb{M}\subset\mathbb{N}$, в некоторой окрестности $\mathbb{U}_0\subset\mathbb{N}$ произвольной точки $(q_0,p_0)\in\mathbb{M}$ представима в виде*

$$f(q,p)=\Phi^\mu(q,p)f_\mu(q,p),\qquad (q,p)\in\mathbb{U}_0, \tag{20.82}$$

с достаточно гладкими коэффициентами $f_\mu(q,p)\in\mathcal{C}^k(\mathbb{U}_0)$.

Доказательство. Поскольку связи Φ^μ функционально независимы, то в некоторой окрестности произвольной точки на поверхности связей $(q_0, p_0) \in \mathbb{M}$ их можно дополнить до системы координат $\{x^\alpha\} \to \{y^a, \Phi^\mu\}$, где y^a, $a = 1, \ldots, 2(\mathrm{N}-\mathrm{M})$, – координаты на поверхности связей. Тогда согласно предложению 3.6.2 произвольная достаточно гладкая функция $f \in \mathcal{C}^k(\mathbb{U}_0)$ представима в виде

$$f(q,p) = f(q_0,p_0) + y^a(q,p) f_a(q,p) + \Phi^\mu(q,p) f_\mu(q,p)$$

с достаточно гладкими коэффициентами f_μ, f_a. Так как функция f равна нулю на поверхности связей $\mathbb{M}$, то $f(q_0,p_0) = 0$ и $f_a(q,p) = 0$ для всех a и $(q,p) \in \mathbb{U}_0$. Поэтому справедливо представление (20.82). □

20.2.2 Гамильтонова динамика частиц со связями II рода

Начнем рассмотрение гамильтоновой динамики частиц со связями с обсуждения связей второго рода, т.к. она значительно проще. Связи первого рода будут рассмотрены в следующем разделе. Там будет показано, что динамика частиц со связями I рода после наложения канонической калибровки сводится к рассмотрению частиц со связями II рода.

Определение. Связи Φ^μ, $\mu = 1, \ldots, 2\mathrm{M}$, называются *связями II рода*, если определитель матрицы, составленной из скобок Пуассона связей между собой, отличен от нуля на поверхности связей:

$$\det [\Phi^\mu, \Phi^\nu]_{\Phi=0} \neq 0. \quad \square \tag{20.83}$$

Отметим, что элементы матрицы, составленной из скобок Пуассона $[\Phi^\mu, \Phi^\nu]$, в общем случае зависят от точки фазового пространства. Условие (20.83) означает отличие определителя от нуля в каждой точке подмногообразия $\mathbb{M} \subset \mathbb{N}$, определяемого связями (20.80). Из непрерывности функций Φ^μ следует, что определитель (20.83) отличен от нуля не только на $\mathbb{M}$, но и в некоторой окрестности подмногообразия $\mathbb{M}$.

Поскольку матрица, составленная из скобок Пуассона,

$$J^{\mu\nu} = -J^{\nu\mu} = [\Phi^\mu, \Phi^\nu], \tag{20.84}$$

антисимметрична, то ее определитель может быть отличен от нуля только при четном числе связей. Поэтому мы с самого начала предположили наличие четного числа связей $2\mathrm{M}$. Это значит, что поверхность связей также имеет четную размерность $2(\mathrm{N}-\mathrm{M})$.

С точки зрения вариационного принципа при наличии связей на фазовом пространстве мы имеем задачу на условный экстремум, которую можно решать методом неопределенных множителей Лагранжа. С этой целью рассмотрим *обобщенный (extended) гамильтониан*, который получается после добавления к исходному гамильтониану всех связей второго рода

$$H_{\mathrm{E}} = H + \lambda_\mu \Phi^\mu, \tag{20.85}$$

где $\lambda_\mu = \lambda_\mu(q,p,t)$ – неопределенные множители Лагранжа. Этому гамильтониану соответствует *обобщенное действие*

$$S_{\mathrm{E}} = \int_{t_1}^{t_2} dt (p\dot{q} - H - \lambda_\mu \Phi^\mu). \tag{20.86}$$

При варьировании этого действия мы считаем вариации координат на границе нулевыми, $\delta q(t_{1,2}) = 0$, а вариации импульсов и множителей Лагранжа могут быть произвольны, т.к. они входят в действие без производных. Канонические уравнения для обобщенного гамильтониана имеют вид

$$\begin{aligned} \dot{q}^i &= [q^i, H_{\text{E}}] = [q^i, H] + \lambda_\mu [q^i, \Phi^\mu] + [q^i, \lambda_\mu]\Phi^\mu, \\ \dot{p}_i &= [p_i, H_{\text{E}}] = [p_i, H] + \lambda_\mu [p_i, \Phi^\mu] + [p_i, \lambda_\mu]\Phi^\mu. \end{aligned} \tag{20.87}$$

Эти уравнения необходимо дополнить уравнениями связей (20.80), возникающими при варьировании обобщенного действия по множителям Лагранжа. В уравнениях движения (20.87) последние слагаемые можно отбросить, т.к. они равны нулю на поверхности связей (20.80). Таким образом, мы имеем $2(\text{N} + \text{M})$ уравнений (20.87), (20.80) на $2(\text{N} + \text{M})$ переменных q, p, λ, из которых 2N уравнений являются дифференциальными. Решение этой задачи можно провести следующим образом.

Предложение 20.2.2. *Для того, чтобы фазовая траектория гамильтоновой системы (20.86), проходящая через произвольную точку поверхности связей, целиком лежала на этой поверхности необходимо и достаточно, чтобы производная по времени от всех связей обращалась в нуль,*

$$\dot{\Phi}^\mu = [\Phi^\mu, H_{\text{E}}]_{\Phi=0} = [\Phi^\mu, H]_{\Phi=0} + \lambda_\nu [\Phi^\mu, \Phi^\nu]_{\Phi=0} = 0. \tag{20.88}$$

Доказательство. Достаточность. Пусть выполнено условие (20.88) для каждой связи. Тогда, если в начальный момент времени траектория находится на поверхности связей, то она на ней и останется во все последующие моменты времени. Необходимость. Допустим, что каждая траектория, имеющая хотя бы одну точку на поверхности связей, целиком принадлежит этой поверхности. Тогда производная вдоль траектории от каждой связи должна быть равна нулю на поверхности связей, поскольку мы можем произвольно менять исходную точку на поверхности. Это означает выполнение условий (20.88) для всех значений индекса μ. □

Уравнения (20.88) можно рассматривать как уравнения на множители Лагранжа на поверхности связей. При этом условие, определяющее связи второго рода (20.83), является необходимым и достаточным для однозначного определения множителей Лагранжа $\lambda_\mu(x)$ на поверхности связей $\mathbb{M}$. Вне поверхности связей множители Лагранжа можно продолжить любым достаточно гладким образом, поскольку это не влияет на динамику частицы на поверхности связей. Для определенности положим

$$\lambda_\mu = -J^{-1}_{\mu\nu} [\Phi^\nu, H] \tag{20.89}$$

в некоторой окрестности поверхности связей. Если продолжить множители Лагранжа вне поверхность связей каким либо иным образом, то это изменит только траектории, лежащие вне поверхности связей, которые нас не интересуют. Множители Лагранжа (20.89) подставляем в уравнения (20.87) и решаем задачу Коши для координат и импульсов. Тогда, если в начальный момент времени точка фазового пространства находилась на поверхности связей, то она там и останется при эволюции системы. Это означает, что метод неопределенных множителей Лагранжа позволяет из вариационного принципа для обобщенного действия (20.86) определить множители Лагранжа и эволюцию динамических переменных.

Решение (20.89) означает, что, если связи выполняются в начальный момент времени для некоторой траектории, то множители Лагранжа всегда можно подобрать таким образом, что связи будут выполнены и во все последующие моменты времени.

Исключим множители Лагранжа (20.89) из уравнений движения (20.87):

$$\dot{q}^i = [q^i, H] - [q^i, \Phi^\mu] J^{-1}_{\mu\nu} [\Phi^\nu, H],$$
$$\dot{p}_i = [p_i, H] - [p_i, \Phi^\mu] J^{-1}_{\mu\nu} [\Phi^\nu, H].$$

Эти уравнения можно записать в компактном виде с помощью нового важного понятия, которое было введено в [**?**].

Определение. Для любых функций $f, g \in \mathcal{C}^1(\mathbb{N})$ билинейная операция

$$[f, g]_{\text{D}} = [f, g] - [f, \Phi^\mu] J^{-1}_{\mu\nu} [\Phi^\nu, g], \tag{20.90}$$

называется *скобкой Дирака.* □

Скобка Дирака, очевидно, антисимметрична и билинейна. Кроме того, нетрудно проверить, что для нее справедливо правило Лейбница и тождества Якоби. Следовательно, скобка Дирака определяет на фазовом пространстве $\mathbb{N}$ новую пуассонову структуру (см. раздел 18.3). Напомним, что на фазовом пространстве уже существует каноническая пуассонова структура для координат и сопряженных импульсов $[f, g]$. Скобка Дирака определяет на $\mathbb{N}$ вторую *диракову пуассонову структуру*, которая обладает рядом замечательных свойств для гамильтоновых систем со связями II рода.

Сначала заметим, что уравнения движения (20.87) можно записать в компактном виде

$$\dot{q}^i = [q^i, H]_{\text{D}}, \qquad \dot{p}_i = [p_i, H]_{\text{D}}. \tag{20.91}$$

В эти уравнения не входят явно множители Лагранжа, потому что достаточная информация о связях содержится в определении скобки Дирака.

Из определения (20.90) следует, что скобка Дирака каждой связи (20.80) с произвольной функцией $F \in \mathcal{C}^1(\mathbb{N})$ равна нулю:

$$[F, \Phi^\mu]_{\text{D}} = 0.$$

Отсюда вытекает, что диракова пуассонова структура вырождена и все связи второго рода являются функциями Казимира для скобки Дирака. Покажем, что других функционально независимых функций Казимира не существует.

Предложение 20.2.3. *На поверхности связей* $\Phi = 0$ *существует такая система координат* y^a, $a = 1, \ldots, 2(\text{N} - \text{M})$, *что выполнены следующие условия:*

$$\mathsf{det}\, [y^a, y^b]_{\Phi=0} \neq 0, \qquad [y^a, \Phi^\mu]_{\Phi=0} = 0.$$

Доказательство. Поскольку связи функционально независимы, то их можно выбрать в качестве части координатных функций новой системы координат в некоторой окрестности поверхности связей. Выберем какие либо координаты $\tilde{y}^a$, $a = 1, \ldots, 2(\text{N} - \text{M})$, на поверхности связей. Тогда совокупность функций $\{\tilde{y}^a, \Phi^\mu\}$ образует систему координат в окрестности поверхности связей. В общем случае $[\tilde{y}^a, \Phi^\mu] \neq 0$. Тогда введем новые координаты на поверхности связей $y^a = \tilde{y}^a + \lambda^a_\mu \Phi^\mu$, где λ^a_μ – некоторые функции. На поверхности связей $\Phi = 0$ эти координаты совпадают со старыми. Выберем неизвестные функции λ^a_μ таким образом, чтобы на поверхности связей выполнялись уравнения:

$$[y^a, \Phi^\mu]_{\Phi=0} = [\tilde{y}^a, \Phi^\mu]_{\Phi=0} + \lambda^a_\nu [\Phi^\nu, \Phi^\mu]_{\Phi=0} = 0.$$

Это всегда можно сделать, поскольку связи Φ^μ второго рода. В новой системе координат структурные функции канонической пуассоновой структуры на поверхности связей примет вид

$$J = \begin{pmatrix} [y^a, y^b]_{\Phi=0} & 0 \\ 0 & [\Phi^\mu, \Phi^\nu]_{\Phi=0} \end{pmatrix}.$$

Поскольку она не может быть вырожденной, то $\det [y^a, y^b]_{\Phi=0} \neq 0$. □

В системе координат $\{y^a, \Phi^\mu\}$ из только что доказанного предложения структурные функции дираковой пуассоновой структуры примут вид

$$J_{\text{D}} = \begin{pmatrix} [y^a, y^b]_{\Phi=0} & 0 \\ 0 & 0 \end{pmatrix},$$

поскольку $[y^a, y^b]_{\text{D}\Phi=0} = [y^a, y^b]_{\Phi=0}$. Отсюда следует, что в окрестности связей ранг J_{D} равен $2(\text{N} - \text{M})$. Таким образом мы доказали следующее утверждение.

Теорема 20.2.1. *Связи второго рода Φ^μ и только они являются функциями Казимира для пуассоновой структуры Дирака (20.90).*

Используя скобку Дирака, для каждой связи Φ^μ можно построить векторное поле с компонентами

$$X^\alpha_{\Phi^\mu} = [\Phi^\mu, x^\alpha]_{\text{D}},$$

где $x = (q, p)$ – координаты на фазовом пространстве $\mathbb{N}$. Согласно предложению 18.3.1 коммутаторы этих векторных полей равны нулю:

$$[X_{\Phi^\mu}, X_{\Phi^\nu}] = X_{[\Phi^\mu, \Phi^\nu]_{\text{D}}} = 0.$$

Следовательно распределение векторных полей $\{X_{\Phi^\mu}\}$, $\mu = 1, \ldots, 2\text{M}$, находится в инволюции и определяет интегральные подмногообразия в $\mathbb{N}$. Таким образом, постоянные значения функций Казимира слоят фазовое пространство $\mathbb{N}$ на симплектические сечения, одним из которых является поверхность связей $\mathbb{M} \subset \mathbb{N}$.

Каноническая пуассонова структура на $\mathbb{N}$ не приспособлена для описания систем со связями, потому что ограничение скобки Пуассона двух функций на поверхность связей нельзя проводить до вычисления самой скобки:

$$\Big[f, g\Big]_{\Phi=0} \neq \Big[f_{\Phi=0}, g_{\Phi=0}\Big].$$

Например, мы сразу приходим к противоречию, если $f = \Phi^\mu$ и $g = \Phi^\nu$ для некоторых μ и ν. В то же время ограничение скобки Дирака на поверхность связей можно проводить до вычисления самой скобки:

$$\Big[f, g\Big]_{\text{D}}\Big|_{\Phi=0} = \Big[f_{\Phi=0}, g_{\Phi=0}\Big]_{\text{D}}. \tag{20.92}$$

Так как поверхность связей является подмногообразием $\mathbb{M} \subset \mathbb{N}$, то ее можно рассматривать, как вложение $\mathbb{M} \hookrightarrow \mathbb{N}$. Тогда равенство (20.92) означает, что вложение $\mathbb{M} \hookrightarrow \mathbb{N}$ является пуассоновым отображением для скобки Дирака, но не для канонической пуассоновой структуры на $\mathbb{N}$.

Поскольку скобки Дирака достаточно для описания эволюции всех переменных, то ее использование приносит существенные упрощения в описание динамики систем со связями II рода.

Скобка Дирака (20.90) определена с помощью канонической пуассоновой структуры на фазовом пространстве $\mathbb{N}$. В разделе 20.1.10 было показано, что канонические преобразования сохраняют вид канонической скобки Пуассона. Поэтому скобка Дирака двух функций также инвариантна относительно канонических преобразований, т.е. она имеет одинаковый вид в любых координатах на $\mathbb{N}$, связанных между собой каноническим преобразованием.

В физических приложениях важную роль играет специальная система координат на фазовом пространстве, которая строится с учетом вложения $\mathbb{M} \hookrightarrow \mathbb{N}$. Ее существование обеспечивается следующим важным утверждением.

Теорема 20.2.2. *Пусть задано фазовое пространство $\mathbb{N}$ и набор функционально независимых связей второго рода (20.80). Тогда существует такое каноническое преобразование*

$$\{q^i, p_i\} \;\to\; \{q^{*a}, p^*_a, Q^{\mathrm{A}}, P_{\mathrm{A}}\}, \quad a = 1, \ldots, \mathrm{N} - \mathrm{M}, \quad \mathrm{A} = 1, \ldots, \mathrm{M},$$

что набор связей $\Phi = 0$ эквивалентен связям

$$Q = 0, \qquad P = 0.$$

При этом координаты q^*, p^* являются канонически сопряженными координатами и импульсами на поверхности связей $\mathbb{M} \subset \mathbb{N}$.

Доказательство. Поскольку связи Φ^μ являются связями второго рода, то согласно теореме Дарбу существуют такие координаты $Q^{\mathrm{A}}, P_{\mathrm{A}}$, что $[Q^{\mathrm{A}}, P_{\mathrm{B}}] = \delta^{\mathrm{A}}_{\mathrm{B}}$ и уравнения $\Phi = 0$ эквивалентны уравнениям $Q = 0$, $P = 0$. После этого из теоремы Дарбу 18.5.3 следует существование системы координат $q^{*a}, p^*_a, Q^{\mathrm{A}}, P_{\mathrm{A}}$. Поскольку скобка Пуассона в новых координатах имеет канонический вид, то преобразование координат является каноническим. □

Эта теорема локальна. Если найдены координаты q^*, p^*, Q, P в явном виде, то ограничение функций на поверхность связей особенно просто: нужно просто положить $Q = 0$ и $P = 0$. В частности,

$$H_{\mathrm{E}}|_{\Phi=0} = H|_{Q=0,\,P=0}.$$

Поскольку

$$[q^*, Q] = 0, \quad [q^*, P] = 0 \quad \text{и} \quad [p^*, Q] = 0, \quad [p^*, P] = 0,$$

то уравнения движения (20.91) в новой системе координат примут вид

$$\dot{q}^* = [q^*, H_{\mathrm{ph}}], \quad \dot{p}^* = [p^*, H_{\mathrm{ph}}]$$
$$\dot{Q}^{\mathrm{A}} = 0, \quad \dot{P}_{\mathrm{A}} = 0,$$

где

$$H_{\mathrm{ph}}(q^*, p^*) = H_{\mathrm{E}}|_{\Phi=0} = H(q^*, p^*, Q, P)|_{Q=0,\,P=0}.$$

Мы видим, что динамика системы N частиц, на которую наложено 2M связей второго рода свелась к обычной динамике системы из $\mathrm{N} - \mathrm{M}$ частиц, на которую уже не наложено никаких связей. По этой причине мы говорим, что система имеет $\mathrm{N} - \mathrm{M}$ *физических степеней свободы* q^*, p^*. Координаты Q, P описывают *нефизические степени свободы*, поскольку определяют связи. Динамика физических степеней свободы

задается *эффективным (физическим) гамильтонианом* $H_{\rm ph}$, зависящим только от физических координат. Координаты на поверхности связей определены неоднозначно. Например, на $\mathbb{M}$ всегда можно совершить каноническое преобразование. Однако размерность поверхности связей (удвоенное число физических степеней свободы) фиксирована и всегда равна $2(\mathrm{N}-\mathrm{M})$. Мы говорим, что система N частиц с 2M связями II рода описывает $\mathrm{N}-\mathrm{M}$ физических степеней свободы.

Замечание. На практике найти в явном виде координаты q^*, p^* для физических степеней свободы удается только в простейших случаях. Кроме того, в теории поля связи часто представляют собой дифференциальные уравнения по пространственным координатам. В этих случаях переход к координатам q^*, p^* задается нелокальными выражениями, как, например, в свободной электродинамике (см. раздел 23.3.3). Поэтому, вычисления обычно проводят в исходных координатах q^i, p_i, используя теорему 20.2.2 для доказательства общих утверждений. □

Таким образом, если на фазовом пространстве $\mathbb{N}$ с канонической пуассоновой структурой задана система связей $\{\Phi^\mu\}$ второго рода, то она определяет некоторое подмногообразие $\mathbb{M} \subset \mathbb{N}$, на котором естественным образом определена каноническая пуассонова структура. Посмотрим на эту задачу с обратной точки зрения. Пусть задано вложение $\mathbb{M} \hookrightarrow \mathbb{N}$, и мы знаем, что подмногообразие $\mathbb{M}$ является фазовым пространством некоторой механической системы с канонической пуассоновой структурой. Возникает естественный вопрос может ли пространство-мишень также быть фазовым пространством ? Ниже мы покажем, что ответ на этот вопрос положительный.

Пусть задано вложение

$$\varphi: \quad \mathbb{M} \;\hookrightarrow\; \mathbb{N}, \qquad \dim \mathbb{M} = 2(\mathrm{N}-\mathrm{M}), \quad \dim \mathbb{N} = 2\mathrm{N}, \;\; \mathrm{M} < \mathrm{N},$$

фазового пространства $\mathbb{M}$ с канонической пуассоновой структурой. Тогда в окрестности каждой точки $\mathbb{M}$ существует такая система координат y^a, $a = 1, \ldots, 2(\mathrm{N}-\mathrm{M})$, в которой пуассонова структура имеет канонический вид

$$[y^a, y^b] = \varpi^{-1ab}.$$

Эта пуассонова структура определяет пуассонову структуру на образе $\varphi(\mathbb{M})$ с помощью дифференциала отображения φ_*. В дальнейшем образ отображения вложения мы отождествим с самим многообразием, т.е. положим $\mathbb{M} = \varphi(\mathbb{M}) \subset \mathbb{N}$. Если x^α, $\alpha = 1, \ldots, 2\mathrm{N}$, – система координат на $\mathbb{N}$, то индуцированная пуассонова структура на $\mathbb{M}$ в координатах $x^\alpha \in \mathbb{N}$ задается антисимметричной матрицей

$$J^{\alpha\beta} = \varpi^{-1ab} \partial_a x^\alpha \partial_b x^\beta, \tag{20.93}$$

где функции $x^\alpha(y)$ описывают вложение. Из свойств произведения матриц следует, что ранг этой пуассоновой структуры равен $2(\mathrm{N}-\mathrm{M})$, и поэтому она всегда вырождена. Следовательно, для индуцированной на $\mathbb{N}$ пуассоновой структуры существует 2M независимых функций Казимира c^μ, $\mu = 1, \ldots 2\mathrm{M}$. Выберем координаты в окрестности $\mathbb{M} \subset \mathbb{N}$ следующим образом

$$\{x^\alpha\} = \{y^a, c^\mu\}.$$

Тогда подмногообразие $\mathbb{M}$ задается постоянными значениями функций Казимира $c^\mu = \mathsf{const}$. Отсюда следует, что уравнения $c^\mu = \mathsf{const}$ эквивалентны исходной системе связей (20.80).

Проведенное построение показывает, что на фазовом пространстве $\mathbb{N}$ существует такая система координат, в которой индуцированная пуассонова структура имеет вид

$$J^{\alpha\beta} = \begin{pmatrix} \varpi^{-1ab} & 0 \\ 0 & 0 \end{pmatrix}.$$

Тем самым можно отождествить индуцированную пуассонову структуру с дираковой пуассоновой структурой, а функции Казимира c^μ со связями Φ^μ.

Ранее скобка Дирака была определена через каноническую пуассонову структуру на исходном многообразии $\mathbb{N}$ с помощью связей. В обратную сторону однозначного рецепта определения невырожденной пуассоновой структуры на $\mathbb{N}$ не существует, поскольку ранг индуцированной пуассоновой структуры $2(\mathrm{N}-\mathrm{M})$ меньше размерности многообразия $\dim \mathbb{N} = 2\mathrm{N}$. Здесь существует много возможностей. Например, можно положить $J^{\alpha\beta} = \varpi^{-1\alpha\beta}$ в координатах $\{y^a, c^\mu\}$. Тогда получим каноническую пуассонову структуру на $\mathbb{N}$.

20.2.3 Гамильтонова динамика частиц со связями I рода

Пусть действие для N точечных частиц в фазовом пространстве $\mathbb{N}$ имеет обычный вид (20.7). Будем искать решение канонических уравнений движения (20.24) при наличии $\mathrm{M} < \mathrm{N}$ связей на канонические переменные:

$$G_{\mathrm{A}}(q,p) = 0, \qquad \mathrm{A} = 1, \ldots, \mathrm{M} < \mathrm{N}. \tag{20.94}$$

Другими словами, будем считать, что частицы не могут покинуть $(2\mathrm{N}-\mathrm{M})$-мерного подмногообразия фазового пространства $\mathbb{U} \subset \mathbb{N}$ (*поверхность связей*), определенного уравнениями (20.94). Из дальнейшего рассмотрения будет ясно, почему мы выбрали число связей первого рода в два раза меньшим числа связей второго рода и почему индекс у связей первого рода пишется внизу.

Предположим, что связи являются достаточно гладкими функциями и функционально независимы на поверхности связей (20.94), т.е. ранг матрицы $\partial G_{\mathrm{A}}/\partial x^\alpha$, где $x = (q^1, \ldots, q^{\mathrm{N}}, p_1, \ldots, p_{\mathrm{N}})$ – координаты фазового пространства, максимален и равен M. Будем считать, что рассматриваемая гамильтонова система со связями находится в *инволюции*:

$$[G_{\mathrm{A}}, G_{\mathrm{B}}] = f_{\mathrm{AB}}{}^{\mathrm{C}} G_{\mathrm{C}} \approx 0, \tag{20.95}$$

$$[G_{\mathrm{A}}, H] = v_{\mathrm{A}}{}^{\mathrm{B}} G_{\mathrm{B}} \approx 0, \tag{20.96}$$

где $f_{\mathrm{AB}}{}^{\mathrm{C}}(q,p) = -f_{\mathrm{BA}}{}^{\mathrm{C}}(q,p)$ и $v_{\mathrm{A}}{}^{\mathrm{B}}(q,p)$ – некоторые функции от точки фазового пространства $(q,p) \in \mathbb{N}$. Волнистый знак равенства $\approx$ обозначает, что данная функция обращается в нуль на поверхности связей:

$$f \approx 0 \quad \Leftrightarrow \quad f|_{G=0} = 0.$$

При этом говорят, что скобки Пуассона связей между собой и с гамильтонианом *слабо равны нулю*.

Замечание. Для выполнения равенств (20.95) необходимо, чтобы количество связей не превосходило половины размерности фазового пространства $\mathrm{M} \leq \mathrm{N}$, что мы потребовали с самого начала (20.94). Действительно, при преобразованиях координат фазового пространства каноническая пуассонова структура не может вырождаться. В силу функциональной независимости функций G_{A} их можно выбрать в качестве части новых координат. Если $\mathrm{M} > \mathrm{N}$, то пуассонова структура была бы вырожденной, что невозможно для фазового пространства. □

Определение. Связи (20.94), наложенные на канонические переменные механической системы с гамильтонианом $H(q,p)$, которые удовлетворяют условиям (20.95), (20.96), называются *связями I рода.* □

Мы предполагаем, что все связи I рода учтены в системе уравнений (20.94), т.е. для заданной гамильтоновой системы не существует большего числа функционально независимых соотношений между каноническими переменными, для которых выполнены условия (20.95) и (20.96).

Скобки Пуассона (20.95) по виду напоминают коммутатор базисных векторных полей в алгебре Ли. Однако в рассматриваемом случае допускается нетривиальная зависимость *структурных функций* от точки фазового пространства: $f_{\mathrm{AB}}{}^{\mathrm{C}} = f_{\mathrm{AB}}{}^{\mathrm{C}}(q,p)$.

Напомним, что связи определены неоднозначно. А именно, невырожденные линейные комбинации связей определяют ту же поверхность связей. Это приведет к изменению структурных функций $f_{\mathrm{AB}}{}^{\mathrm{C}}$ и $v_{\mathrm{A}}{}^{\mathrm{B}}$, что важно при решении уравнений движения.

Функции $f_{\mathrm{AB}}{}^{\mathrm{C}}$ и $v_{\mathrm{A}}{}^{\mathrm{B}}$ не могут быть произвольными. По-определению, скобки Пуассона удовлетворяют тождеству Якоби. Рассмотрев скобки Пуассона $\Big[[G_{\mathrm{A}},G_{\mathrm{B}}],G_{\mathrm{C}}\Big]$ и $\Big[[G_{\mathrm{A}},G_{\mathrm{B}}],H\Big]$ и их циклические перестановки, получим ограничения на структурные функции:

$$\begin{aligned} f_{\mathrm{AB}}{}^{\mathrm{D}} f_{\mathrm{DC}}{}^{\mathrm{E}} + f_{\mathrm{BC}}{}^{\mathrm{D}} f_{\mathrm{DA}}{}^{\mathrm{E}} + f_{\mathrm{CA}}{}^{\mathrm{D}} f_{\mathrm{DB}}{}^{\mathrm{E}} + [f_{\mathrm{AB}}{}^{\mathrm{E}}, G_{\mathrm{C}}] + [f_{\mathrm{BC}}{}^{\mathrm{E}}, G_{\mathrm{A}}] + [f_{\mathrm{CA}}{}^{\mathrm{E}}, G_{\mathrm{B}}] = 0, \\ f_{\mathrm{AB}}{}^{\mathrm{D}} v_{\mathrm{D}}{}^{\mathrm{C}} + v_{\mathrm{B}}{}^{\mathrm{D}} f_{\mathrm{DA}}{}^{\mathrm{C}} - v_{\mathrm{A}}{}^{\mathrm{D}} f_{\mathrm{DB}}{}^{\mathrm{C}} + [f_{\mathrm{AB}}{}^{\mathrm{C}}, H] + [v_{\mathrm{B}}{}^{\mathrm{C}}, G_{\mathrm{A}}] - [v_{\mathrm{A}}{}^{\mathrm{C}}, G_{\mathrm{B}}] = 0. \end{aligned} \tag{20.97}$$

Тождество Якоби для двойной скобки Пуассона $\Big[[G_{\mathrm{A}},H],H\Big]$ удовлетворяются автоматически в силу уравнения (20.96).

Если структурные функции постоянны, $f_{\mathrm{AB}}{}^{\mathrm{C}} = \mathsf{const}$, то множество всех линейных комбинаций $a^{\mathrm{A}}G_{\mathrm{A}}$, $a^{\mathrm{A}} \in \mathbb{R}$, образует алгебру Ли с базисом G_{A}. Тогда уравнение (20.95) задает коммутатор базисных векторов, а соотношение (20.97) сводится к тождеству Якоби для структурных констант.

Предложение 20.2.4. *Для того, чтобы фазовые траектории для гамильтониана H, проходящие через произвольную точку на поверхности связей, целиком лежали на этой поверхности необходимо и достаточно, чтобы были выполнены условия (20.96).*

Доказательство. Совпадает с доказательством предложения 20.2.2. Достаточно только заметить, что

$$\dot{G}_{\mathrm{A}} = [G_{\mathrm{A}}, H] \approx 0, \tag{20.98}$$

и воспользоваться предложением 20.2.1. □

С точки зрения вариационного принципа решение канонических уравнений (20.24) при наличии связей (20.94) является задачей на условный экстремум. Применим к решению этой задачи метод неопределенных множителей Лагранжа. С этой целью построим *полный (total) гамильтониан*, добавив к исходному гамильтониану линейную комбинацию связей первого рода,

$$H_{\mathrm{T}} = H + \lambda^{\mathrm{A}} G_{\mathrm{A}}, \tag{20.99}$$

где $\lambda^{\text{A}} = \lambda^{\text{A}}(q,p,t)$ – неопределенные множители Лагранжа. Соответствующее действие, которое мы назовем *полным*, имеет вид

$$S_{\text{T}} = \int_{t_1}^{t_2} dt (p\dot{q} - H - \lambda^{\text{A}} G_{\text{A}}).$$

Из него вытекают уравнения Эйлера–Лагранжа:

$$\begin{aligned} \frac{\delta S_{\text{T}}}{\delta p_i} &= \quad \dot{q}^i - \frac{\partial H}{\partial p_i} - \lambda^{\text{A}} \frac{\partial G_{\text{A}}}{\partial p_i} = 0, \\ \frac{\delta S_{\text{T}}}{\delta q^i} &= -\dot{p}_i - \frac{\partial H}{\partial q^i} - \lambda^{\text{A}} \frac{\partial G_{\text{A}}}{\partial q^i} = 0, \end{aligned} \tag{20.100}$$

$$\frac{\delta S_{\text{T}}}{\delta \lambda^{\text{A}}} = -G_{\text{A}} = 0, \tag{20.101}$$

где в правых частях уравнений движения (20.100) мы отбросили слагаемые, пропорциональные связям. При варьировании действия мы считаем вариации координат на границе равными нулю, $\delta q(t_{1,2}) = 0$, а вариации импульсов и множителей Лагранжа могут быть произвольны.

Предложение 20.2.5. *Для того, чтобы фазовые траектории для полного действия S_{T}, проходящие через произвольную точку на поверхности связей, целиком лежали на этой поверхности достаточно, чтобы были выполнены условия (20.95) и (20.96).*

Доказательство. Повторяет доказательство предложения 20.2.2. Достаточно заметить, что

$$\dot{G}_{\text{A}} = [G_{\text{A}}, H] + \lambda^{\text{B}} [G_{\text{A}}, G_{\text{B}}] \approx 0. \qquad \square$$

Уравнения движения для систем со связями первого рода (20.100) существенно отличаются от уравнений для систем со связями второго рода. Дело в том, что они не позволяют определить ни одного множителя Лагранжа. Это связано с тем, что уравнения (20.88), из которых находились множители Лагранжа для связей второго рода в рассматриваемом случае (20.98) на поверхности связей вовсе не содержат множителей Лагранжа. Тем самым любое решение уравнений движения для систем со связями первого рода содержит произвольные функции $\lambda^{\text{A}}(t)$, число которых совпадает с числом связей. Причиной функционального произвола в решении уравнений движения является калибровочная инвариантность.

Теорема 20.2.3. *Полное действие S_{T} инвариантно относительно локальных инфинитезимальных преобразований, генерируемых каждой связью первого рода,*

$$\begin{aligned} \delta q^i = \epsilon^{\text{A}} [q^i, G_{\text{A}}] &= \quad \epsilon^{\text{A}} \frac{\partial G_{\text{A}}}{\partial p_i}, \\ \delta p_i = \epsilon^{\text{A}} [p_i, G_{\text{A}}] &= -\epsilon^{\text{A}} \frac{\partial G_{\text{A}}}{\partial q^i}, \\ \delta \lambda^{\text{A}} &= \dot{\epsilon}^{\text{A}} + \epsilon^{\text{B}} v_{\text{B}}{}^{\text{A}} + \epsilon^{\text{B}} \lambda^{\text{C}} f_{\text{BC}}{}^{\text{A}}, \end{aligned} \tag{20.102}$$

где $\epsilon^{\text{A}} = \epsilon^{\text{A}}(q,p,t)$ – малый параметр преобразований и $\dot{\epsilon} = \partial\epsilon/\partial t$. Параметр преобразований может быть произвольной функцией координат, импульсов и времени с нулевыми граничными условиями $\epsilon^{\text{A}}(t_{1,2}) = 0$ для всех q и p.

Доказательство. Вариация действия равна

$$\delta S_{\text{т}} = \int dt \left(-\epsilon^{\text{а}} \frac{\partial G_{\text{а}}}{\partial q^i} \dot{q}^i - \dot{p}_i \epsilon^{\text{а}} \frac{\partial G_{\text{а}}}{\partial p_i} + \epsilon^{\text{а}} [G_{\text{а}}, H] + \epsilon^{\text{а}} \lambda^{\text{в}} [G_{\text{а}}, G_{\text{в}}] - \delta\lambda^{\text{а}} G_{\text{а}} \right).$$

Подставляя вариацию множителей Лагранжа и интегрируя слагаемое $\dot{\epsilon}^{\text{а}} G_{\text{а}}$ по частям с учетом уравнений (20.95) и (20.96), получаем $\delta S_{\text{т}} = 0$. Условия на параметры калибровочных преобразований $\epsilon^{\text{а}}(t_{1,2}) = 0$ достаточны для того, чтобы интегрирование слагаемого $\dot{\epsilon}^{\text{а}} G_{\text{а}}$ по частям было возможно. □

Согласно второй теореме Нетер локальная инвариантность приводит к линейной зависимости уравнений движения:

$$\frac{\delta S_{\text{т}}}{\delta q} \frac{\partial G_{\text{а}}}{\partial p} - \frac{\delta S_{\text{т}}}{\delta p} \frac{\partial G_{\text{а}}}{\partial q} - \frac{\partial}{\partial t} \left(\frac{\delta S_{\text{т}}}{\delta \lambda^{\text{а}}} \right) + \frac{\delta S_{\text{т}}}{\delta \lambda^{\text{в}}} (v_{\text{а}}{}^{\text{в}} + \lambda^{\text{с}} f_{\text{ас}}{}^{\text{в}}) = 0, \tag{20.103}$$

в чем нетрудно убедиться и прямой проверкой. В этом случае для любого решения системы уравнений (20.100) будет автоматически выполнено уравнение

$$\dot{G}_{\text{а}} = (v_{\text{а}}{}^{\text{в}} + \lambda^{\text{с}} f_{\text{ас}}{}^{\text{в}}) G_{\text{в}}.$$

Для этой системы уравнений точка $G = 0$ является стационарной. Поэтому, если в начальный момент времени фазовая траектория находится на поверхности связей, то для любого решения системы уравнений (20.100) при любых множителях Лагранжа $\lambda^{\text{а}}(q, p, t)$ связи $G_{\text{а}} = 0$ будут автоматически удовлетворены. Это и является причиной возникновения функционального произвола в решениях уравнений движения.

Замечание. Преобразования координат фазового пространства q, p (20.102) с малым постоянным параметром $\epsilon = \mathsf{const}$ являются инфинитезимальной формой канонического преобразования, описанного в примере 20.1.19,

$$q^i(t) \;\mapsto\; q^i(t, \epsilon), \qquad p_i(t) \;\mapsto\; p_i(t, \epsilon),$$

определяемого уравнениями

$$\frac{\partial q^i}{\partial \epsilon^{\text{а}}} = [q^i, G_{\text{а}}], \qquad \frac{\partial p_i}{\partial \epsilon^{\text{а}}} = [p_i, G_{\text{а}}]. \qquad \square$$

Определение. Модели, функционал действия которых инвариантен относительно локальных преобразований, называют *калибровочными*, а сами преобразования – *калибровочными*. □

Замечание. Это название пришло из физики, где все калибровочные модели, в частности, электродинамика и поля Янга–Миллса обладают этим свойством. □

Тем самым мы доказали, что каждой связи первого рода соответствует локальная инвариантность полного действия, а сама связь является генератором калибровочных преобразований. По аналогии с генераторами групп Ли мы пишем индексы у связей первого рода внизу.

Проанализируем уравнения движения подробнее. Поскольку среди $2\text{N} + \text{M}$ уравнений Эйлера–Лагранжа (20.100) и (20.101) только 2N являются независимыми, то в отличии от задачи на условный экстремум, рассмотренной в разделе 19.1, этих уравнений недостаточно для определения всех неизвестных функций $q(t)$, $p(t)$ и $\lambda(t)$.

Допустим, что систему уравнений (20.100) можно решить относительно канонических переменных, для которых поставлена задача Коши: $q(t_1) = q_1$, $p(t_1) = p_1$. Тогда решение этой задачи будет зависеть от M произвольных функций времени, которыми являются множители Лагранжа, и 2N постоянных интегрирования. Все постоянные интегрирования определяются начальными данными. В этом случае через одну точку фазового пространства проходит множество фазовых траекторий, которые параметризуется множителями Лагранжа.

С физической точки зрения это означает следующее. Пусть некоторое физическое явление описывается калибровочной моделью. Тогда начальное состояние системы не определяет однозначно последующую эволюцию, что противоречит экспериментальным данным, если не принимать во внимание квантовомеханическую неопределенность. Тем не менее калибровочные модели в настоящее время широко используются в теоретической физике: классическим примером служит электродинамика.

Выход из этого противоречия прост. Для калибровочных моделей вводится постулат: все физические наблюдаемые калибровочно инвариантны, т.е. не зависят от произвольных функций, которые могут содержаться в решении уравнений движения. Отсюда следует, что физические наблюдаемые описываются функциями $f \in \mathcal{C}^k(\mathbb{N})$ на фазовом пространстве $\mathbb{N}$, которые являются калибровочно инвариантными. А именно, мы требуем, чтобы скобка Пуассона каждой наблюдаемой функции со связями обращалась в нуль на поверхности связей:

$$[f, G_{\mathrm{A}}] = d_{\mathrm{A}}{}^{\mathrm{B}} G_{\mathrm{B}}, \tag{20.104}$$

где $d_{\mathrm{A}}{}^{\mathrm{B}}(q,p)$ – некоторые достаточно гладкие функции канонических переменных. Тогда в уравнении движения для калибровочно инвариантной функции

$$\dot{f} = [f, H] + \lambda^{\mathrm{A}}[f, G_{\mathrm{A}}]$$

все слагаемые с множителями Лагранжа обратится в нуль на поверхности связей и, следовательно, на поверхности связей никакого произвола в эволюции калибровочно инвариантной функции нет.

Проведенное рассмотрение требует комментария, потому что каждой физической наблюдаемой соответствует не одна, а целый класс калибровочно инвариантных функций. Поскольку мы рассматриваем динамику частиц на поверхности связей (20.94), то физические наблюдаемые определяются значениями калибровочно инвариантных функций на поверхности связей. Пусть две калибровочно инвариантные функции f_1 и f_2 совпадают на поверхности связей $\mathbb{U}$. Тогда они могут отличаться только на линейную комбинацию связей:

$$f_2 = f_1 + \mu^{\mathrm{A}} G_{\mathrm{A}}, \tag{20.105}$$

где $\mu^{\mathrm{A}}(q,p)$ – некоторые достаточно гладкие функции. Следовательно, все множество калибровочно инвариантных функций разбивается на классы эквивалентности (20.105). При этом каждый класс эквивалентности соответствует одной физической наблюдаемой.

Другими словами, калибровочно инвариантная функция задается на поверхности связей $\mathbb{U} \subset \mathbb{N}$, а затем продолжается на все фазовое пространство в значительной степени произвольным образом. Степень произвола описывается произвольными функциями $\mu^{\mathrm{A}}(q,p)$. При этом физические наблюдаемые не зависят от способа продолжения.

Оценим произвол, с которым калибровочно инвариантная функция может быть задана на поверхности связей $\mathbb{U} \subset \mathbb{N}$. Условие калибровочной инвариантности (20.104) представляет собой систему M дифференциальных уравнений в частных производных первого порядка, которую можно переписать в виде

$$X_{G_\mathrm{A}} f = 0,$$

где X_{G_A} – векторное поле, соответствующее связи G_A (18.27). Для этой системы уравнений условиями совместности являются уравнения (20.95). Действительно,

$$[X_{G_\mathrm{A}}, X_{G_\mathrm{B}}] f = X_{[G_\mathrm{A}, G_\mathrm{B}]} f = f_{\mathrm{AB}}{}^{\mathrm{C}} X_{G_\mathrm{C}} f = 0,$$

где мы воспользовались равенством (18.28). Поэтому функция f однозначно определяется начальными данными на некотором собственном подмногообразии $\mathbb{M} \subset \mathbb{U} \subset \mathbb{N}$ размерности $2\mathrm{N} - \mathrm{M} - \mathrm{M} = 2(\mathrm{N} - \mathrm{M})$ и существенно зависит только от части координат на $\mathbb{U}$. Подмногообразие $\mathbb{M}$ можно задать с помощью M дополнительных функционально независимых соотношений между координатами и импульсами:

$$F^\mathrm{A}(q, p) = 0, \tag{20.106}$$

которые называются *калибровочными условиями*. Эти условия должны удовлетворять неравенству

$$\det [F^\mathrm{A}, G_\mathrm{B}] \Big|_{F=0,\, G=0} \neq 0, \tag{20.107}$$

т.к. только в этом случае на $\mathbb{M}$ можно задать начальные данные для системы уравнений (20.104). Условия (20.106) при выполнении (20.107) называются *канонической калибровкой*. Для точечных частиц отличие от нуля определителя (20.107) является необходимым и достаточным условием однозначного определения множителей Лагранжа в полном гамильтониане (20.99). В общем случае функции F^A, определяющие калибровочные условия, могут зависеть также от времени и множителей Лагранжа λ^A.

После фиксирования канонической калибровки на рассматриваемую гамильтонову систему будет наложено 2M связей (20.94) и (20.106). Введем для полной совокупности связей и калибровочных условий следующее обозначение

$$\{\Phi^\mu\} = (F^1, \ldots, F^\mathrm{M}, G_1, \ldots, G_\mathrm{M}), \qquad \mu = 1, \ldots, 2\mathrm{M}. \tag{20.108}$$

Очевидно, что

$$\det [\Phi^\mu, \Phi^\nu]_{\Phi=0} = \det \begin{pmatrix} [F^\mathrm{A}, F^\mathrm{B}] & [F^\mathrm{A}, G_\mathrm{B}] \\ [G_\mathrm{A}, F^\mathrm{B}] & [G_\mathrm{A}, G_\mathrm{B}] \end{pmatrix}_{\Phi=0} = \det{}^2 [F^\mathrm{A}, G_\mathrm{B}]_{\Phi=0}, \tag{20.109}$$

т.к. $[G_\mathrm{A}, G_\mathrm{B}] \approx 0$. Поскольку определитель скобок Пуассона для канонических калибровочных условий со связями первого рода, по-построению, отличен от нуля (20.107), то полная совокупность связей Φ^μ представляет собой систему связей второго рода, рассмотренную в предыдущем разделе. Таким образом, калибровочные модели в канонической калибровке сведены к гамильтоновым системам со связями второго рода, для которых метод множителей Лагранжа применим в полном объеме. Заметим, что значение скобок Пуассона для канонических калибровочных условий между собой $[F^\mathrm{A}, F^\mathrm{B}]$ несущественно.

После наложения калибровочных условий возникает обобщенное действие

$$S_\mathrm{E} = \int dt (p_i \dot{q}^i - H - \lambda^\mathrm{A} G_\mathrm{A} - \pi_\mathrm{A} F^\mathrm{A}) = \int dt (p_i \dot{q}^i - H - \lambda_\mu \Phi^\mu),$$

которое совпадает с (20.86).

Каноническая калибровка выделяет в фазовом пространстве $\mathbb{N}$ единственную траекторию, проходящую через данную точку физического подпространства $\mathbb{M}$. Это происходит благодаря тому, что каноническая калибровка однозначно определяет множители Лагранжа на поверхности связей $\Phi = 0$. Верно также обратное утверждение: произвольный выбор множителей Лагранжа эквивалентен некоторой канонической калибровке. Действительно, при фиксированных множителях Лагранжа граничная задача для уравнений (20.100) имеет единственное решение $x(t) = \{q(t), p(t)\}$. Перепишем для данного решения уравнения (20.100), но теперь уже с неопределенными множителями Лагранжа. В результате получим переопределенную, но совместную систему $2\mathrm{N}$ линейных алгебраических уравнений на M множителей Лагранжа. Тогда соответствующие соотношения $x^{\mathrm{A}} = x^{\mathrm{A}}(t)$ можно принять в качестве канонических калибровочных условий. Это следует из того, что, поскольку условия $F^{\mathrm{A}} = x^{\mathrm{A}} - x^{\mathrm{A}}(t)$ позволяют однозначно определить множители Лагранжа, то условие (20.107) выполнено. Напомним, что это условие является необходимым и достаточным для однозначного определения λ^{A} на поверхности связей, как следует из анализа, проведенного в предыдущем разделе, и (20.109).

Замечание. В приложениях часто рассматриваются неканонические калибровки. Например, лоренцева или временна́я калибровки в электродинамике не являются каноническими. Они не фиксируют калибровочную свободу полностью, и это создает определенные трудности, например, при квантовании. В электродинамике неканонические калибровочные условия можно использовать, т.к. уравнения движения линейны и их можно проанализировать. В общем случае существенно нелинейных моделей формализм для неканонических калибровок не развит и мы на них останавливаться не будем. □

Таким образом, динамика частиц со связями I рода в канонических калибровках сводится к гамильтоновым моделям со связями II рода. Это сведение не является однозначным, т.к. от калибровочных условий требуется выполнение только неравенств (20.107). Выбор той или иной системы калибровочных условий диктуется рассматриваемой задачей и соображениями простоты. Как правило, исследование калибровочных моделей проводится в различных калибровках, каждая из которых имеет свои преимущества и недостатки.

20.2.4 Калибровочная модель нерелятивистской частицы

В настоящем разделе мы рассмотрим динамику точечной частицы с точки зрения калибровочных моделей и покажем трудности в определении энергии, которые при этом возникают.

Рассмотрим движение точечной частицы в фазовом пространстве $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}})$ с координатами $q^* = \{q^{*a}\}$ и $p^* = \{p^*_a\}$, где $a = 1, \ldots, \mathrm{N}$. Пусть динамика частицы задана гамильтонианом $H^*(q^*, p^*)$, не зависящим явно от времени. Действие такой частицы имеет обычный вид

$$S = \int_{t_1}^{t_2} dt(\dot{q}^{*a} p^*_a - H^*), \tag{20.110}$$

где точка обозначает дифференцирование по времени t. Будем считать, что все координаты имеют фиксированные значения на границе $t = t_{1,2}$. Для каждой траектории

частицы $q^*(t), p^*(t)$, которая определяется каноническими уравнениями движения

$$\dot{q}^* = \frac{\partial H^*}{\partial p^*}, \qquad \dot{p}^* = -\frac{\partial H^*}{\partial q^*},$$

выполнено равенство $\dot{H}^* = 0$. Это значит, что на каждой траектории гамильтониан имеет постоянное численное значение, которое называется энергией точечной частицы.

Переформулируем модель точечной частицы, как калибровочную. С этой целью расширим фазовое пространство $\mathbb{T}^*(\mathbb{R}^{\mathrm{N}}) \to \mathbb{T}^*(\mathbb{R}^{\mathrm{N}+1})$ путем введения дополнительной пары сопряженных канонических переменных Q, P, и рассмотрим новое действие

$$S_{\mathrm{т}} = \int_{\tau_1}^{\tau_2} d\tau (\dot{Q}P + \dot{q}^* p_* - \lambda G), \qquad G = -Q + H^*, \tag{20.111}$$

где λ – множитель Лагранжа, а точка обозначает дифференцирование по некоторому параметру τ, играющему роль времени. Уравнения движения для этой модели имеют вид

$$\dot{Q} = 0, \tag{20.112}$$

$$\dot{P} = \lambda, \tag{20.113}$$

$$\dot{q}^* = \lambda \frac{\partial H^*}{\partial p_*}, \tag{20.114}$$

$$\dot{p}_* = -\lambda \frac{\partial H^*}{\partial q^*}, \tag{20.115}$$

$$G = 0. \tag{20.116}$$

Последнее уравнение представляет собой уравнение связи. Эта связь, очевидно, является связью первого рода, которая определяет гамильтониан системы.

Согласно общей теории действие (20.111) инвариантно относительно инфинитезимальных калибровочных преобразований с параметром $\epsilon(\tau)$, генерируемых связью G,

$$\begin{aligned} \delta Q &= \epsilon[Q, G] = 0, \\ \delta P &= \epsilon[P, G] = \epsilon, \\ \delta q^* &= \epsilon[q^*, G] = \epsilon \frac{\partial H^*}{\partial p_*}, \\ \delta p_* &= \epsilon[p_*, G] = -\epsilon \frac{\partial H^*}{\partial q^*}. \end{aligned} \tag{20.117}$$

При этом множитель Лагранжа преобразуется по правилу

$$\delta\lambda = \dot{\epsilon}. \tag{20.118}$$

Для того, чтобы вариации δq^* на границе были равны нулю, необходимо предположить, что $\epsilon(\tau_{1,2}) = 0$. Вариация множителя Лагранжа на границе несущественна, т.к. он входит в действие без производных. Отметим, что эволюцию во времени канонических переменных можно рассматривать, как калибровочное преобразование с параметром $\epsilon = \lambda d\tau$.

Приведенные выше калибровочные преобразования представляют собой бесконечно малые преобразования, соответствующие инвариантности действия (20.111)

относительно перепараметризации времени. При произвольном преобразовании временно́го параметра $\tau' = \tau'(\tau)$ мы постулируем, что координаты фазового пространства преобразуются, как скаляры

$$Q'(\tau') = Q(\tau), \qquad P'(\tau') = P(\tau)$$

(такие же формулы преобразования постулируются для q^* и p_*). При этом множитель Лагранжа преобразуется, как 1-форма:

$$\lambda'(\tau') = \frac{d\tau}{d\tau'}\lambda(\tau).$$

Нетрудно проверить, что действие (20.111) инвариантно относительно произвольной перепараметризации времени τ. Это и есть калибровочная инвариантность. При бесконечно малом преобразовании $\tau' = \tau + u(\tau)$ для вариаций формы функций имеем

$$\begin{aligned}
\delta Q &= -u\dot{Q} = 0,\\
\delta P &= -u\dot{P} = -u\lambda,\\
\delta q^* &= -u\dot{q}^* = -u\lambda\frac{\partial H^*}{\partial p_*},\\
\delta p_* &= -u\dot{p}_* = u\lambda\frac{\partial H^*}{\partial q^*},\\
\delta\lambda &= -u\dot{\lambda} - \dot{u}\lambda = -\frac{d}{d\tau}(u\lambda),
\end{aligned} \tag{20.119}$$

где в первых четырех уравнениях были использованы уравнения движения. Полученные преобразования совпадают с инфинитезимальными калибровочными преобразованиями (20.117), (20.118) при $\epsilon = -u\lambda$.

Таким образом мы показали, что действие (20.111) инвариантно относительно произвольной перепараметризации времени. В модели имеется одна связь первого рода, и для фиксирования соответствующего произвола в решениях уравнений движения необходимо наложить одно калибровочное условие. Поскольку множитель Лагранжа λ является произвольной функцией времени, то из уравнения движения (20.113) следует, что импульс P также произволен. Чтобы устранить этот произвол, зафиксируем калибровку

$$F_1 = P - \mathsf{const} = 0.$$

Тогда из уравнений движения определяется множитель Лагранжа $\lambda = 0$. При этом для действия получаем следующее выражение

$$S_{\text{т}}|_{F_1=0,\,G=0} = \int d\tau \dot{q}^* p^*.$$

Это значит, что в выбранной калибровке мы получили "замороженную" теорию, в которой не происходит никакой эволюции. При этом вся эволюция заменяется калибровочным преобразованием.

Можно рассмотреть класс калибровок, явно зависящих от времени. Пусть калибровочное условие имеет вид

$$F_2 = P - \tau = 0.$$

В этом случае из уравнений движения следует $\lambda = 1$, и эффективный гамильтониан для физических степеней свободы становится нетривиальным. Поскольку

$$\dot{Q}P = \frac{d}{d\tau}(QP) - Q\dot{P},$$

и

$$Q\dot{P}|_{F_2=0,G=0} = H^*(q^*, p^*).$$

то эффективное действие для физических степеней свободы равно

$$S_{\text{т}}|_{F_2=0,G=0} = \int d\tau(\dot{q}^* p^* - H^*),$$

что совпадает с исходным действием (20.110) для точечной частицы. Мы видим, что в калибровке $F_2 = 0$ нетривиальный эффективный гамильтониан возникает из кинетического слагаемого $\dot{Q}P$ для нефизической степени свободы.

Можно рассмотреть более общий класс калибровок

$$F_3 = P - f(\tau) = 0,$$

где $f(\tau)$ – произвольная функция времени с положительной производной, $\dot{f} > 0$. Для этой калибровки $\lambda = \dot{f}$, и полное действие принимает вид

$$S_{\text{т}}|_{F_3=0,G=0} = \int d\tau(\dot{q}^* p_* - \dot{f} H^*).$$

После перепараметризации траектории $dt = d\tau\dot{f}$, мы возвращаемся к исходному действию для точечной частицы (20.110).

Таким образом, действие (20.111) калибровочно инвариантно и после наложения калибровочного условия (из достаточно широкого класса калибровок) эквивалентно обычному действию для точечной частицы. В рассмотренном примере нефизическую степень свободы удалось в явном виде исключить из теории после решения связи и калибровочного условия. В подавляющем большинстве моделей математической физики это сделать не удается. Даже в электродинамике связи и калибровочные условия нельзя решить в явном виде. Ситуация в моделях Янга–Миллса и гравитации еще более сложная. Поэтому для проведения вычислений в калибровочных моделях используют методы, учитывающие как физические, так и нефизические степени свободы.

А теперь обратимся к вопросу об определении энергии в калибровочно инвариантных теориях. Во многих моделях математической физики исходное действие в гамильтоновой форме имеет вид (20.111). В таком виде гамильтониан системы при выполнении уравнений движения тождественно равен нулю, и принимать его за энергию системы не имеет никакого смысла. В рассмотренной модели за энергию частицы естественно принять численное значение гамильтониана H^* для физических степеней свободы. Для того, чтобы его построить, исходя из действия (20.111), необходимо сначала зафиксировать калибровку, зависящую явно от времени, а затем решить уравнения движения для нефизических степеней свободы и связь. При этом, выбирая различные функции времени в калибровочном условии $F_3 = 0$, можно получить, что множитель Лагранжа и, следовательно, эффективный гамильтониан будут явно зависеть от времени. Для простоты картины, следует выбрать такую функцию времени, чтобы эта зависимость исчезла. Если это возможно, то построенный таким образом гамильтониан следует принять за определение энергии, а соответствующий ему временно́й параметр назвать временем $\tau = t$.

20.2.5 Частица в псевдоримановом пространстве

Рассмотрим точечную частицу постоянной массы m, которая движется в произвольном псевдоримановом пространстве $(\mathbb{M}, g)$ размерности n. Пусть q^α, $\alpha = 0, 1, \ldots, n-1$,

– локальная система координат в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$. Будем считать, что координаты выбраны таким образом, что q^0 является временно́й координатой, т.е. $g_{00} > 0$, и все сечения $q^0 = \mathsf{const}$ пространственноподобны, т.е. пространственная часть метрики $g_{\mu\nu}$, $\mu, \nu = 1, \ldots, n-1$ отрицательно определена или $N^2 > 0$, где $N > 0$ – функция хода в АДМ параметризации метрики (см. раздел 22.2).

Рассмотрим кривую $\gamma = q(t)$, $t \in [t_0, t_1] \subset \mathbb{R}$, соединяющую две точки этой окрестности $q_0(t_0), q_1(t_1) \in \mathbb{U}$ и целиком лежащую в $\mathbb{U}$. Действие для точечной частицы пропорционально длине траектории (17.16) и имеет вид

$$S = \int_\gamma dt L(q, \dot{q}) = -m \int_\gamma dt \sqrt{g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta}. \tag{20.120}$$

Поскольку траектория частицы предполагается времениподобной, т.е.

$$\dot{q}^2 := g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta > 0,$$

то подынтегральное выражение определено. В рассматриваемом действии метрика $g_{\alpha\beta}(q)$ является внешним заданным полем и по ней варьирование не производится пока не включено взаимодействие с гравитационным полем, т.е. не добавлено, например, действие Гильберта–Эйнштейна. Обобщенные импульсы для координат q^α имеют вид

$$p_\alpha := \frac{\partial L}{\partial \dot{q}^\alpha} = -m \frac{g_{\alpha\beta} \dot{q}^\beta}{\sqrt{\dot{q}^2}}. \tag{20.121}$$

Отсюда следует, что импульсы удовлетворяют соотношению

$$p^2 := g^{\alpha\beta}(q) p_\alpha p_\beta = m^2, \tag{20.122}$$

которое должно быть выполнено для всех траекторий, вдоль которых частица может двигаться. Поэтому соотношение

$$\widetilde{G} := p^2 - m^2 = 0 \tag{20.123}$$

является первичной связью для точечной частицы.

Пример 20.2.2. В пространстве Минковского связь (20.123) имеет вид

$$p_0^2 + \eta^{\mu\nu} p_\mu p_\nu - m^2 = 0$$

и выделяет в фазовом пространстве $(q, p) \in \mathbb{T}^*(\mathbb{R}^{1,n-1})$ двуполостный гиперболоид в кокасательном пространстве, умноженный на пространство Минковского $\mathbb{R}^{1,n-1}$, которое соответствует координатам q. Поверхность связей является несвязным подмногообразием в фазовом пространстве $\mathbb{T}^*(\mathbb{R}^{1,n-1})$ и состоит из двух компонент, определяемых неравенствами $p_0 > 0$ и $p_0 < 0$. Топологически поверхность связей в произвольном лоренцевом многообразии устроена также. □

Используя АДМ параметризацию метрики, связь (20.123) перепишем в виде

$$p^2 - m^2 = \left[\frac{1}{N}(p_0 - N^\mu p_\mu) + \sqrt{\hat{p}^2 + m^2} \right] \left[\frac{1}{N}(p_0 - N^\mu p_\mu) - \sqrt{\hat{p}^2 + m^2} \right] = 0, \tag{20.124}$$

где введено обозначение $\hat{p}^2 := -\hat{g}^{\mu\nu} p_\mu p_\nu > 0$ для положительно определенного квадрата пространственных компонент импульса частицы. Равенство нулю одного из сомножителей выделяет в фазовом пространстве одну полу "гиперболоида". Из определения импульсов (20.121) следует равенство

$$p_0 - N^\mu p_\mu = -\frac{mN^2}{\sqrt{\dot{q}^2}} \dot{q}^0. \tag{20.125}$$

Поэтому при $\dot{q}^0 > 0$ второй сомножитель в уравнении (20.124) отрицательно определен и не обращается в нуль на поверхности связи. Если $\dot{q}^0 < 0$, то на поверхности связи отличен от нуля первый сомножитель. Следовательно, связь (20.123) эквивалентна связи

$$G := \sqrt{\hat{p}^2 + m^2} - \frac{1}{N}|p_0 - N^\mu p_\mu| = 0, \tag{20.126}$$

где мы использовали знак модуля, чтобы объединить оба случая $\dot{q}^0 > 0$ и $\dot{q}^0 < 0$. В таком виде первичная связь будет удобна для дальнейших вычислений.

Перейдем к вычислению гамильтониана $H = p_\alpha \dot{q}^\alpha - L$. Поскольку в теории есть первичная связь, то гессиан модели вырожден и из определения обобщенных импульсов (20.121) нельзя определить все скорости, как функции импульсов и координат. Ранг гессиана в рассматриваемом случае равен $n-1$, что позволяет определить пространственные компоненты скорости. Чтобы их найти, заметим, что

$$\dot{q}^2 = N^2(\dot{q}^0)^2 + g_{\mu\nu}(\dot{q}^\mu + N^\mu \dot{q}^0)(\dot{q}^\nu + N^\nu \dot{q}^0). \tag{20.127}$$

Далее, из определения импульсов (20.121) следует равенство

$$p_\mu \sqrt{\dot{q}^2} = -m(N_\mu \dot{q}^0 + g_{\mu\nu}\dot{q}^\nu).$$

Возведение этого равенства в квадрат с помощью метрики $\hat{g}^{\mu\nu}$ позволяет найти квадрат пространственных компонент скорости,

$$g_{\mu\nu}(\dot{q}^\mu + N^\mu \dot{q}^0)(\dot{q}^\nu + N^\nu \dot{q}^0) = -\frac{N^2\hat{p}^2}{\hat{p}^2 + m^2}(\dot{q}^0)^2.$$

Подстановка полученного выражения в (20.127) приводит к равенству

$$\sqrt{\dot{q}^2} = \frac{Nm}{\sqrt{\hat{p}^2 + m^2}}|\dot{q}^0|,$$

которое позволяет решить выражение для импульсов (20.121) относительно скоростей:

$$\frac{\dot{q}^0}{|\dot{q}^0|} = -\frac{p_0 - N^\nu p_\nu}{N\sqrt{\hat{p}^2 + m^2}},$$
$$\frac{\dot{q}^\mu}{|\dot{q}^0|} = -\frac{N^2\hat{g}^{\mu\nu}p_\nu - N^\mu(p_0 - N^\nu p_\nu)}{N\sqrt{\hat{p}^2 + m^2}}.$$

Мы видим, что уравнения (20.121) определяют только знак производной $\dot{q}^0$ и пространственные компоненты скорости $\dot{q}^\mu$. При этом модуль временно́й компоненты скорости $|\dot{q}^0|$ является произвольной функцией.

Теперь нетрудно вычислить гамильтониан, который удобно записать в виде

$$H := p\dot{q} - L = N|\dot{q}^0|G, \tag{20.128}$$

где мы использовали равенство

$$(p_0 - N^\mu p_\mu)\dot{q}^0 = -|p_0 - N^\mu p_\mu||\dot{q}^0|,$$

которое следует из (20.125). Таким образом, гамильтониан однозначно определен и пропорционален первичной связи. Действие (20.120) в гамильтоновой форме принимает вид

$$S = \int dt(p_\alpha \dot{q}^\alpha - H) = \int dt\left(p_\mu \dot{q}^\mu - N|\dot{q}^0|\sqrt{\hat{p}^2 + m^2} + N^\mu p_\mu\right). \tag{20.129}$$

Отметим сокращение кинетических слагаемых $p_0\dot{q}^0$. Можно проверить, что гамильтоновы уравнения движения для данного действия (вариации по p_μ, q^μ и q^0) эквивалентны исходным уравнениям Эйлера–Лагранжа для действия (20.120).

Заметим, что из определения обобщенных импульсов (20.121) следует не только выражение для пространственных компонент скоростей $\dot{q}^\mu$, но и знак производной, $\mathsf{sign}\,\dot{q}^0 = -\,\mathsf{sign}\,(p_0 - N^\mu p_\mu)$. Тем самым вырожденность гессиана приводит к тому, что только модуль временно́й компоненты скорости $|\dot{q}^0|$ остается неопределенным, и его нельзя исключить из гамильтониана (20.128).

В соответствии с общим методом, полный гамильтониан на первом этапе получается путем добавления первичной связи. Так как исходный гамильтониан уже пропорционален связи, то

$$H_{\text{т}} = \lambda G, \tag{20.130}$$

где $\lambda = \lambda(t)$ – неопределенный множитель Лагранжа.

Поскольку скобка Пуассона связи с собой равна нулю,

$$[G, G] = 0$$

(она должна быть антисимметрична, что невозможно для одной связи), то вторичных связей не возникает и связь $G = 0$ является связью первого рода. Поэтому система находится в инволюции, т.е. выполнены условия (20.95), (20.96). Отсюда вытекает, что полный гамильтониан системы (20.99) определяется единственной связью первого рода (20.130).

Гамильтониан (20.130) приводит к уравнениям движения, которые эквивалентны уравнениям движения для гамильтониана

$$\widetilde{H}_{\text{т}} = \mu\widetilde{G}, \tag{20.131}$$

где $\mu = \mu(t)$ – множитель Лагранжа, на поверхности связей, т.к. связи $\widetilde{G}$ и G эквивалентны. Действительно, связи G и $\widetilde{G}$ отличаются на отличный от нуля множитель (для каждой полы гиперболоида):

$$\widetilde{G} = fG, \qquad f(q,p) \neq 0.$$

Поэтому

$$\dot{q}^\alpha = \mu[q^\alpha, \widetilde{G}] = \mu[q^\alpha, f]G + \mu f[q^\alpha, G].$$

На поверхности связи $G = 0$, и поэтому первое слагаемое в правой части исчезает. Такой же вид имеет уравнение для импульсов. Следовательно, замена связи в полном гамильтониане $H_{\text{т}}$ приводит к переопределению множителя Лагранжа: $\lambda = \mu f$.

В дальнейшем из соображений удобства мы будем выбирать тот или иной вид полного гамильтониана.

Рассмотрим гамильтоновы уравнения движения (20.100) для гамильтониана (20.131)

$$\begin{aligned} \dot{q}^\alpha &= 2\mu g^{\alpha\beta} p_\beta, \\ \dot{p}_\alpha &= \mu p^\beta p^\gamma \partial_\alpha g_{\beta\gamma}. \end{aligned} \tag{20.132}$$

Для сравнения полученных уравнений движения с уравнениями для экстремалей (17.17), продифференцируем первое уравнение по времени и воспользуемся вторым уравнением для исключения $\dot{p}$. В результате получим уравнение для координат частицы

$$\ddot{q}^\alpha = \frac{\dot{\mu}}{\mu}\dot{q}^\alpha - \Gamma_{\beta\gamma}{}^\alpha \dot{q}^\beta \dot{q}^\gamma. \tag{20.133}$$

С точностью до первого слагаемого в правой части оно совпадает с уравнением для экстремалей (17.17) (для краткости знак тильды у Γ опущен). Это слагаемое связано с произволом в выборе параметризации мировой линии частицы. Действительно, после преобразования $t \mapsto t'(t)$, где функция $t'(t)$ удовлетворяет дифференциальному уравнению

$$\frac{dt'}{dt} = \mu(t), \qquad \mu > 0,$$

гамильтоновы уравнения движения будут иметь вид (20.133), но с $\mu = 1$. В этом случае уравнения движения совпадут с уравнениями для экстремалей, и, значит, параметр t' является каноническим параметром вдоль экстремали.

Поскольку модель содержит связь первого рода, то ей соответствует калибровочная инвариантность полного действия

$$S_{\text{т}} = \int_\gamma dt(p_\alpha \dot{q}^\alpha - \mu \widetilde{G}). \tag{20.134}$$

Чтобы найти соответствующие преобразования симметрии, рассмотрим инфинитезимальные преобразования, которые генерируются связью (20.102):

$$\begin{aligned} \delta q^\alpha &= \epsilon[q^\alpha, \widetilde{G}] = 2\epsilon g^{\alpha\beta} p_\beta, \\ \delta p_\alpha &= \epsilon[p_\alpha, \widetilde{G}] = \epsilon p^\beta p^\gamma \partial_\alpha g_{\beta\gamma}, \\ \delta\mu &= \dot{\epsilon}. \end{aligned}$$

Сравнение этих преобразований с уравнениями движения (20.132) показывает, что эволюция во времени канонических переменных представляет собой последовательность калибровочных преобразований, где $\epsilon = \mu dt$. Эта калибровочная симметрия описывает произвол в выборе параметра t вдоль траектории частицы. Уравнения движения (20.132) вместе со связью (20.126) не определяют множитель Лагранжа μ. Для его определения необходимо зафиксировать калибровку.

Если учесть выражение для импульсов (20.121), то кинетический член в действии (20.134) в следующем виде

$$p_\alpha \dot{q}^\alpha = -m\sqrt{\dot{q}^2}.$$

После интегрирования кинетического члена получается исходное действие (20.120).

Временна́я калибровка

Продолжим исследование модели в соответствии с общей схемой. Зафиксируем *временну́ю калибровку*

$$F_{\text{т}} := q^0 - bt = 0, \qquad b = \mathsf{const} \neq 0. \tag{20.135}$$

Модуль свободного параметра b связан с выбором единицы измерения времени, что не существенно. Поэтому, без ограничения общности, положим $|b| = 1$, т.е. $b = 1$, если время q^0 для внешнего наблюдателя увеличивается при увеличении параметра t вдоль мировой линии частицы, и $b = -1$, если при увеличении t время q^0 уменьшается.

Скобка Пуассона этого калибровочного условия со связью равна

$$[F_{\text{т}}, G] = -\frac{1}{N}\left[q^0, |p_0 - N^\mu p_\mu|\right] = \begin{cases} 1/N, & b > 0, \\ -1/N, & b < 0, \end{cases} \tag{20.136}$$

где мы использовали равенство (20.125) для определения знака выражения, стоящего под модулем. Поскольку функция хода отлична от нуля, $N \neq 0$, то скобка Пуассона (20.136) отлична от нуля во всем фазовом пространстве и, в частности, на поверхности связи $G = 0$. Следовательно, условие $q^0 = bt$ определяет каноническую калибровку.

Расширенный гамильтониан с учетом связи и калибровочного условия имеет вид

$$H_{\text{E}} = \lambda G + \pi F_{\text{T}},$$

где λ и π – множители Лагранжа. Из условий сохранения связи и калибровочного условия во времени,

$$\dot{G} = \lambda[G, G] + \pi[G, F_{\text{T}}] = \pi[G, F_{\text{T}}] \approx 0,$$
$$\dot{F}_{\text{T}} = \frac{\partial F_1}{\partial t} + \lambda[F_1, G] + \pi[F_1, F_1] = -b + \lambda[F_{\text{T}}, G] \approx 0,$$

находим множители Лагранжа:

$$\lambda = N > 0, \qquad \pi = 0,$$

где мы объединили оба случая, $b > 0$ и $b < 0$, и учли равенство $|b| = 1$. Таким образом в данной канонической калибровке полный гамильтониан (20.130) совпадает с исходным гамильтонианом (20.128), т.к. $|\dot{q}^0| = |b| = 1$.

Связь (20.126) и калибровочное условие (20.135) можно решить относительно нефизических переменных q^0 и p_0:

$$q^0 = bt,$$
$$p_0 = \pm N\sqrt{\hat{p}^2 + m^2} + N^\mu p_\mu.$$

Поскольку $\dot{q}^0 = b$, то знак $\pm$ в выражении для p_0 можно включить в постоянную b:

$$p_0 = -bN\sqrt{\hat{p}^2 + m^2} + N^\mu p_\mu.$$

Так как решения уравнений Эйлера–Лагранжа можно подставлять в действие (см. раздел 19.4), то эффективное действие для физических переменных q^μ, p_μ принимает вид

$$S_{\text{eff}} = \int dt \left(p_0\dot{q}^0 + p_\mu\dot{q}^\mu\right) = \int dt \left(p_\mu\dot{q}^\mu - H_{\text{eff}}\right),$$

где эффективный гамильтониан для физических степеней свободы имеет вид

$$H_{\text{eff}} = N\sqrt{\hat{p}^2 + m^2} - bN^\mu p_\mu. \tag{20.137}$$

Как и в случае нерелятивистской точечной частицы эффективный гамильтониан для физических степеней свободы полностью определяется кинетическим слагаемым $p_0\dot{q}^0$ для нефизических степеней свободы. Эффективный гамильтониан зависит только от физических степеней свободы, которыми являются пространственные координаты точечной частицы и соответствующие импульсы, $\{q^\mu, p_\mu\}$, $\mu = 1, \ldots, n-1$. Компоненты метрики N, N^μ и $g_{\mu\nu}$ входят в гамильтониан в качестве внешних полей. Уравнения движения для физических степеней свободы во временно́й калибровке (20.135) имеют вид

$$\dot{q}^\mu = \frac{Np^\mu}{\sqrt{\hat{p}^2 + m^2}} - bN^\mu, \tag{20.138}$$

$$\dot{p}_\mu = -\sqrt{\hat{p}^2 + m^2}\partial_\mu N - \frac{N\partial_\mu\hat{g}^{\nu\rho}p_\nu p_\rho}{2\sqrt{\hat{p}^2 + m^2}} + b\partial_\mu N^\nu p_\nu, \tag{20.139}$$

где $p^\mu = \hat{g}^{\mu\nu} p_\nu$.

Чтобы дать физическую интерпретацию двух возможных ориентаций мировой линии $b = \pm 1$ рассмотрим следующий

Пример 20.2.3. В пространстве Минковского $\mathbb{R}^{1,n-1}$ компоненты метрики имеют вид

$$N = 1, \qquad N^\mu = 0, \qquad g_{\mu\nu} = \eta_{\mu\nu} = \mathsf{diag}\,(-\ldots-),$$

и уравнения движения (20.138), (20.139) существенно упрощаются:

$$\dot{q}^\mu = \frac{p^\mu}{\sqrt{\hat{p}^2 + m^2}},$$
$$\dot{p}_\mu = 0.$$

Эти уравнения имеют хорошо известный в специальной теории относительности вид. Возводя первое уравнение в квадрат с помощью пространственной метрики $g_{\mu\nu}$, получим выражение для пространственного импульса частицы через ее скорость:

$$p^\mu = \frac{m\dot{q}^\mu}{\sqrt{1 - \boldsymbol{u}^2}}, \tag{20.140}$$

где мы ввели квадрат пространственной скорости частицы $\boldsymbol{u}^2 = -\dot{q}^\mu \dot{q}^\nu g_{\mu\nu} > 0$. Тогда уравнения движения для свободной точечной частицы в пространстве Минковского во временно́й калибровке сводятся просто к условию сохранения импульса: $p_\mu = \mathsf{const}$.

Для того, чтобы дать физическую интерпретацию двум возможным ориентациям $\dot{q}^0 = \pm 1$ мировой линии частицы относительно временно́й координаты q^0 внешнего наблюдателя, рассмотрим взаимодействие частицы с внешним магнитным полем в четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$. Такое взаимодействие описывается дополнительным слагаемым в лагранжиане:

$$L \mapsto L + eA_\alpha \dot{q}^\alpha,$$

где $A_\alpha(x)$ – четырехмерный потенциал электромагнитного поля и $e = \mathsf{const}$ – заряд частицы. С точки зрения внешнего наблюдателя, для которого временем является координата q^0, добавочное слагаемое в лагранжиане имеет вид

$$gA_\alpha \frac{dq^\alpha}{dq^0},$$

где $g = be = \pm e$ – наблюдаемый заряд частицы. Таким образом, для внешнего наблюдателя частица имеет либо положительный, либо отрицательный заряд. В остальном частицы совпадают. Такие частицы в физике принято называть частицей и античастицей. Например, электрон и позитрон (они, правда, имеют спин $1/2$, в то время как в рассматриваемом случае спин частиц равен нулю).

Таким образом, свобода в выборе ориентации мировой линии частицы относительно временно́й координаты внешнего наблюдателя соответствует двум возможным зарядам частицы $\pm e$. Это означает, что исходное действие (20.120) описывает частицу и античастицу, у которых массы совпадают, а заряды при включении внешнего электромагнитного поля имеют противоположный знак. Если заряд равен нулю, $e = 0$, то действие (20.120) описывает одну нейтральную частицу.

Гамильтоновы уравнения движения для точечной заряженной частицы принимают особенно простой вид при движении в постоянном магнитном поле. В этом случае

временна́я компонента потенциала равна нулю, а пространственные компоненты зависят только от пространственных координат:

$$\{A_\alpha\} = \{A_0 = 0, A_\mu(\boldsymbol{x})\}.$$

Тогда все изменения в гамильтоновом формализме сводятся к переопределению пространственных компонент импульсов:

$$p_\mu \mapsto \tilde{p}_\mu := \frac{\partial(L + eA_\nu\dot{q}^\nu)}{\partial\dot{q}^\mu} = p_\mu + eA_\mu.$$

В частности, гамильтониан для физических степеней свободы и уравнения движения принимают вид

$$\begin{aligned} H_{\text{eff}} &= \sqrt{\tilde{p}^2 + m^2}, \\ \dot{q}^\mu &= \frac{p^\mu + eA^\mu}{\sqrt{\tilde{p}^2 + m^2}}, \\ \dot{p}_\mu &= -e\frac{\partial_\mu A_\nu(p^\nu + eA^\nu)}{\sqrt{\tilde{p}^2 + m^2}}, \end{aligned}$$

где введено обозначение

$$\tilde{p}^2 := -g^{\mu\nu}(p_\mu + eA_\mu)(p_\nu + eA_\nu).$$

Как видим, при взаимодействии с внешним полем импульс больше не сохраняется, и траектория частицы в общем случае отличается от экстремали. □

Точечные частицы двигаются в пространстве-времени вдоль экстремалей. В разделе 17.2 было показано, что через данную точку в данном направлении проходит одна и только одна экстремаль. Это значит, что при постановке задачи Коши для экстремали достаточно задать точку $x \in \mathbb{M}$ и вектор $X \in \mathbb{T}_x(\mathbb{M})$. Анализ настоящего раздела показывает, что эта информация является избыточной. Действительно, длина касательного вектора к экстремали постоянна вдоль экстремали. Это значит, что для однозначного восстановления экстремали, проходящей через данную точку, достаточно задать не сам вектор, а его направление, которое определяется $n-1$ параметром. Кроме того, остается еще произвол в выборе параметризации. Во временно́й калибровке в качестве параметра выбирается наблюдаемое время $q^0 = \pm t$. Следовательно, для задания траектории частицы достаточно задать пространственные координаты q^μ и пространственные компоненты импульсов p_μ в начальный момент времени q^0. Поэтому точечная частица на псевдоримановом многообразии описывает $n-1$ степень свободы.

Калибровка светового конуса

Для многих приложений, например, в суперсимметричных моделях, удобно использовать калибровку светового конуса. Эта калибровка упрощает многие формулы, если частица движется в пространстве Минковского $\mathbb{R}^{1,n-1}$, что мы и предположим.

Чтобы определить калибровку светового конуса, вместо двух первых координат частицы q^0 и q^1 введем новые конусные переменные

$$q^\pm = \frac{1}{\sqrt{2}}(q^0 \pm q^1),$$

оставив остальные координаты без изменения. Поскольку известно выражение новых координат через старые, то совершим каноническое преобразование с производящей функцией

$$S_3 = -\frac{1}{\sqrt{2}}(p_0 + p_1)q^+ - \frac{1}{\sqrt{2}}(p_0 - p_1)q^- - p_2q^2 - \ldots - p_{n-1}q^{n-1},$$

зависящей от новых координат и старых импульсов (см. раздел 20.1.11). Отсюда следуют выражения для новых импульсов, сопряженных $q^\pm$:

$$p_\pm = -\frac{\partial S_3}{\partial q^\pm} = \frac{1}{\sqrt{2}}(p_0 \pm p_1).$$

Остальные импульсы при этом не меняются.

В новых канонических переменных первичная связь (20.123) принимает вид

$$\widetilde{G} = 2p_+p_- + p^{\text{A}}p_{\text{A}} - m^2 = 0, \tag{20.141}$$

где суммирование ведется только по $n-2$ значениям индексов: $\text{A} = 2, \ldots, n-1$. Поскольку квадратичная форма $p^{\text{A}}p_{\text{A}} - m^2$ отрицательно определена, то $p_+p_- > 0$. Полы гиперболоида, соответствующего поверхности связей, определяются неравенствами ($p_+ > 0$, $p_- > 0$) и ($p_+ < 0$, $p_- < 0$). И в любом случае $p_+p_- \neq 0$. Связь (20.141) просто решается

$$p_+ = \frac{1}{2p_-}(-p^{\text{A}}p_{\text{A}} + m^2), \tag{20.142}$$

или

$$p_- = \frac{1}{2p_+}(-p^{\text{A}}p_{\text{A}} + m^2).$$

Зафиксируем калибровку *светового конуса*

$$F_{\text{LC}} := q^+ - at = 0, \qquad a = \pm 1. \tag{20.143}$$

Легко видеть, что

$$[\widetilde{G}, F_{\text{LC}}] = -2p_-.$$

Поскольку на поверхности связей $p_- \neq 0$, то условие (20.143) определяет каноническую калибровку.

Так как связи $\widetilde{G}$ и G пропорциональны, то обобщенный гамильтониан можно записать в виде

$$H_{\text{E}} = \lambda\widetilde{G} + \pi F_{\text{LC}},$$

где λ и π – множители Лагранжа. Из условий сохранения связи и калибровочного условия во времени,

$$\dot{\widetilde{G}} = \lambda[\widetilde{G}, \widetilde{G}] + \pi[\widetilde{G}, F_{\text{LC}}] = \pi[\widetilde{G}, F_{\text{LC}}] \approx 0,$$
$$\dot{F}_{\text{LC}} = \frac{\partial F_{\text{LC}}}{\partial t} + \lambda[F_{\text{LC}}, \widetilde{G}] + \pi[F_{\text{LC}}, F_{\text{LC}}] = -a + \lambda[F_{\text{LC}}, \widetilde{G}] \approx 0,$$

находим множители Лагранжа:

$$\lambda = \frac{a}{2p_-}, \qquad \pi = 0.$$

В калибровке светового конуса физическими переменными являются $\{q^-, q^{\text{A}}, p_-, p_{\text{A}}\}$, $\text{A} = 2, \ldots, n-1$. После подстановки решения уравнения связи (20.142) и калибровочного условия (20.143) в действие,

$$S_{\text{eff}} = \int dt(p_+\dot{q}^+ + p_-\dot{q}^- + p_{\text{A}}\dot{q}^{\text{A}}) = \int dt(p_-\dot{q}^- + p_{\text{A}}\dot{q}^{\text{A}} - H_{\text{eff}}),$$

получаем выражение для эффективного гамильтониана

$$H_{\text{eff}} = a\frac{p^{\text{A}}p_{\text{A}} - m^2}{2p_-}. \tag{20.144}$$

Как и раньше, нетривиальный эффективный гамильтониан для физических степеней свободы возникает из кинетического слагаемого $p_+\dot{q}^+$ для нефизической степени свободы.

Уравнения движения для физических степеней свободы имеют вид

$$\begin{aligned}
\dot{q}^- &= -a\frac{p^{\text{A}}p_{\text{A}} - m^2}{2p_-^2},\\
\dot{q}^{\text{A}} &= a\frac{p^{\text{A}}}{p_-},\\
\dot{p}_- &= 0,\\
\dot{p}_{\text{A}} &= 0.
\end{aligned}$$

Мы видим, что как и во временно́й калибровке уравнения движения в калибровке светового конуса сводятся к сохранению обобщенных импульсов свободной точечной частицы.

Замечание. В калибровке светового конуса параметр эволюции $t \in \mathbb{R}$ целиком лежит на световом конусе в пространстве Минковского, который является характеристикой волнового уравнения (см. раздел 17.6). □

Эффективный гамильтониан (20.144) в калибровке светового конуса мало чем напоминает эффективный гамильтониан (20.137) во временно́й калибровке. Тем на менее оба гамильтониана описывают одну и ту же массивную частицу. Ее траектории в конфигурационном пространстве – это экстремали. Выбор той или иной калибровки является существенным для анализа уравнений движения и квантования. Часто калибровка светового конуса упрощает вычисления, особенно в квантовой теории поля.

20.2.6 Граничные слагаемые в калибровочных моделях

В настоящем разделе мы рассмотрим простой пример вариационной задачи, который проанализируем с различных точек зрения. Этот пример позволяет продемонстрировать тонкости вариационной задачи в теории поля, важную роль граничных слагаемых в действии и связь вариационной задачи на условную стационарную точку с фиксированием калибровки в калибровочных моделях.

Лагранжева формулировка

Обозначим декартовы координаты на евклидовой плоскости $\mathbb{R}^2$ через $x^0 = \tau$, $x^1 = \sigma$. Будем называть координату τ временем, а σ – пространством, хотя мы не предполагаем наличие на $\mathbb{R}^2$ какой либо метрики. Пусть в конечном прямоугольнике задано

действие

$$S = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma (p\dot{q} + p^*\dot{q}^*), \tag{20.145}$$

зависящее от четырех полей $q, p, q^*, p^* \in \mathcal{C}^1(\mathbb{R}^2)$. В действии точка обозначает дифференцирование по τ. В дальнейшем пределы интегрирования, для краткости, будем опускать. Рассмотрим задачу на условный экстремум для действия (20.145). Пусть на поля наложены две связи:

$$G := -\partial_1 q + H^*(q^*, p^*) = 0, \tag{20.146}$$

$$F := p = 0. \tag{20.147}$$

Будем считать, что функция $H^*(q^*, p^*) \geq 0$ зависит только от полей и не зависит от их производных. Предположим также, что поля q и q^* имеют определенные граничные условия при $\tau = \tau_{1,2}$. Этого достаточно для того, чтобы избежать граничных вкладов в вариацию действия (20.145), возникающих при интегрировании по частям. Тогда для действия (20.145) определена задача на условный экстремум (см. раздел 19.1.4).

Решим эту задачу прямым методом и методом множителей Лагранжа. В первом случае исключим из действия переменные q и p с помощью уравнений связей. Поскольку на поверхности связей $p = 0$, то, независимо от вида функции $q(\tau, \sigma)$, первое слагаемое в (20.145) обращается в нуль, и мы получаем эффективное действие

$$S\Big|_{G=0,F=0} = \int d\tau d\sigma p^* \dot{q}^*, \tag{20.148}$$

в котором переменные q^* и p^* уже рассматриваются, как независимые переменные. Поскольку вариация δq^* равна нулю на границе $\tau = \tau_{1,2}$, то из вариационного принципа следуют только уравнения Эйлера–Лагранжа, которые просто интегрируются:

$$\frac{\delta S}{\delta p^*} = \quad \dot{q}^* = 0, \qquad \Rightarrow \qquad q^* = q^*(\sigma), \tag{20.149}$$

$$\frac{\delta S}{\delta q^*} = -\dot{p}^* = 0, \qquad \Rightarrow \qquad p^* = p^*(\sigma). \tag{20.150}$$

Мы видим, что решением уравнений Эйлера–Лагранжа являются произвольные функции от σ. Вид произвольной функции q^* находится из граничных условий, которые должны быть заданы одинаковыми при $\tau = \tau_{1,2}$. Затем можно определить q из уравнения связи (20.146). Таким образом, задача на условный экстремум имеет решение, хотя и не для очень широкого класса граничных условий.

Поскольку с помощью уравнений связей мы исключили переменные q и p, то будем называть их нефизическими, а переменные q^* и p^* – физическими.

Теперь воспользуемся методом неопределенных множителей Лагранжа. Построим расширенное действие

$$S_{\text{E}} = \int d\tau d\sigma (p\dot{q} + p^*\dot{q}^* - \lambda G - \mu F), \tag{20.151}$$

где $\lambda, \mu \in \mathcal{C}^1(\mathbb{R}^2)$ – множители Лагранжа. Для этого действия нефизическими полями

являются q, p, λ и μ. Полная система уравнений Эйлера–Лагранжа имеет вид

$$\frac{\delta S_{\text{E}}}{\delta p} = \dot{q} - \mu = 0, \tag{20.152}$$

$$\frac{\delta S_{\text{E}}}{\delta q} = -\dot{p} - \partial_1 \lambda = 0, \tag{20.153}$$

$$\frac{\delta S_{\text{E}}}{\delta p^*} = \dot{q}^* - \lambda \frac{\partial H^*}{\partial p^*} = 0, \tag{20.154}$$

$$\frac{\delta S_{\text{E}}}{\delta q^*} = -\dot{p}^* - \lambda \frac{\partial H^*}{\partial q^*} = 0, \tag{20.155}$$

$$\frac{\delta S_{\text{E}}}{\delta \lambda} = -G = 0, \tag{20.156}$$

$$\frac{\delta S_{\text{E}}}{\delta \mu} = -F = 0. \tag{20.157}$$

При вариации слагаемого $-\lambda \partial_1 q$ по q возникает также граничное условие на множитель Лагранжа

$$\lambda|_{\sigma_{1,2}} = 0, \tag{20.158}$$

поскольку значения переменной q фиксированы только на пространственноподобной границе $\tau = \tau_{1,2}$ и, следовательно, вариации δq на времениподобной границе $\sigma = \sigma_{1,2}$ произвольны.

Перейдем к анализу уравнений Эйлера–Лагранжа. Решение последней связи (20.157) тривиально. Решение связи (20.156) имеет вид

$$q = \int_{\sigma_1}^{\sigma} d\sigma' H^* + q_0(\tau), \tag{20.159}$$

где q_0 – произвольная функция τ. Отсюда следует, что задание граничных условий $q|_{\sigma_2}$, будет противоречить уравнению связи (20.156), т.к. значение поля q при $\sigma = \sigma_2$ определяется значениями физических полей q^* и p^* во внутренних точках области. Дифференцируя решение (20.159) по τ и используя уравнения (20.154) и (20.155), получим $\dot{q} = \dot{q}_0$. Затем решаем уравнения (20.152) и (20.153) относительно множителей Лагранжа:

$$\mu = \dot{q}_0, \qquad \lambda = \lambda_0(\tau), \tag{20.160}$$

где $\lambda_0(\tau)$ – произвольная функция. Таким образом, мы решили уравнения движения для нефизических переменных q, p и множителей Лагранжа λ, μ, и это решение зависит от двух произвольных функций $q_0(\tau)$ и $\lambda_0(\tau)$. Для физических переменных q^* и p^* остаются уравнения (20.154), (20.155), которые имеют вид обычной гамильтоновой системы. Эта система уравнений действительно воспроизводит уравнения Эйлера–Лагранжа на условный экстремум (20.149), (20.150) при $\lambda_0 = 0$. Заметим, что только это значение согласуется с граничным условием (20.158).

Таким образом, мы решили вариационную задачу для действия (20.145) с заданными граничными условиями для полей q и q^* на границе $\tau = \tau_{1,2}$ прямым способом и методом множителей Лагранжа. Как и следовало ожидать, результат одинаков, а класс решений очень беден.

Однако для расширенного действия (20.151) можно поставить более содержательную вариационную задачу. Предположим, что нефизическая переменная задана на всей границе $q|_{\sigma_{1,2}}$ и $q|_{\tau_{1,2}}$, а физическое поле – только на пространственноподобной границе $q^*|_{\tau_{1,2}}$. В этом случае вариации δq равны нулю на границе и граничного

условия на множитель Лагранжа (20.158) не возникнет. Тогда, при $\lambda_0 \neq 0$, вместо τ можно ввести новую переменную t, определяемую дифференциальным уравнением

$$\frac{dt}{d\tau} = \lambda_0(\tau).$$

Это уравнение определяет координату t с точностью до сдвига на постоянную величину, что несущественно. Тогда уравнения для физических полей (20.154), (20.155) примут вид

$$\begin{aligned} \frac{dq^*}{dt} &= \frac{\partial H^*}{\partial p^*}, \\ \frac{dp^*}{dt} &= -\frac{\partial H^*}{\partial q^*}. \end{aligned} \tag{20.161}$$

Таким образом, мы получили систему гамильтоновых уравнений движения для физических полей, динамика которых определяется гамильтонианом $H^*(q^*, p^*)$. Это показывает, что модель, основанная на действии S_{E} с множителями Лагранжа допускает постановку более широкого класса вариационных задач, чем исходная задача на условную стационарную точку, и является более содержательной.

Здесь выявляется специфика полевых моделей, поскольку важно, что поля зависят не только от времени τ, но и от пространственной координаты σ. Действительно, если бы связь имела вид $-q + H^* = 0$, то уравнение (20.153) приняло бы вид $-\dot{p} + \lambda_0 = 0$. Откуда следовало бы единственное решение $\lambda_0 = 0$ при $p = 0$. (Мы употребили термины время и пространство, исходя из аналогии с теорией относительности, не смотря на то, что на плоскости $\mathbb{R}^2$ никакой метрики не задано.)

Заметим, что подстановка решения связей в расширенное действие S_{E} снова приводит к тривиальному действию (20.148), которое не воспроизводит уравнения Эйлера–Лагранжа (20.161). Это показывает, что подстановки решения части уравнений Эйлера–Лагранжа в действие и в оставшиеся уравнения в общем случае не эквивалентны. Это связано с тем, что в общем случае мы не можем накладывать граничные условия на нефизические поля произвольным образом и рассматривать исчезающие на границе вариации. Действительно, вариация δq определяется физическими полями и их вариациями во внутренних точках области:

$$\delta q = \int_{\sigma_1}^{\sigma} d\sigma' \left(\frac{\partial H^*}{\partial q^*} \delta q^* + \frac{\partial H^*}{\partial p^*} \delta p^* \right), \tag{20.162}$$

и в общем случае не равна нулю на границе $\sigma = \sigma_2$. То есть задание граничного условия $q|_{\sigma_2}$ противоречит уравнению связи (20.156).

Изменим постановку вариационной задачи таким образом, чтобы уравнения Эйлера–Лагранжа остались прежними, а исключение нефизических полей в действии приводило бы к новому действию, воспроизводящему уравнения (20.154), (20.155). Введем новое действие

$$S_{\text{ph}} := S_{\text{E}} - \int_{\tau_1}^{\tau_2} d\tau (\lambda q)\Big|_{\sigma=\sigma_1}^{\sigma_2}, \tag{20.163}$$

которое отличается от расширенного действия с множителями Лагранжа граничным слагаемым. Как и раньше, мы считаем, что на границе $\tau_{1,2}$ заданы значения переменных q и q^*. Добавление граничного члена не меняет уравнений Эйлера–Лагранжа, но меняет граничные условия. Этот граничный член подобран таким образом, чтобы компенсировать граничный вклад в вариацию действия, обусловленный вариацией

δq. Его необходимо добавить, если мы не хотим получить граничное условие $\lambda|_{\sigma_2} = 0$ при произвольной вариации δq на границе. Теперь нетрудно проверить, что на решениях уравнений (20.153) и (20.156) действие принимает вид

$$S_{\rm ph}\Big|_{G=0,F=0} = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma p^* \dot{q}^* - \int_{\tau_1}^{\tau_2} d\tau \lambda(\tau,\sigma_2) q(\sigma_2) = \int_{\tau_1}^{\tau_2} d\tau \int_{\sigma_1}^{\sigma_2} d\sigma (p^* \dot{q}^* - \lambda_0 H^*), \quad (20.164)$$

где мы положили $\lambda_0(\tau) = \lambda(\tau,\sigma_2)$. Это действие воспроизводит гамильтоновы уравнения движения для физических полей. При этом мы отбросили слагаемое с $q_0(\tau)$, которое не влияет на уравнения Эйлера–Лагранжа, т.к. в этом действии функция $\lambda_0(\tau)$ рассматривается как заданная и не варьируется. Таким образом, добавление граничного слагаемого не меняет уравнений Эйлера–Лагранжа, позволяет избежать граничных условий на множители Лагранжа и, что самое важное, разрешает подстановку решений уравнений Эйлера–Лагранжа непосредственно в действие. При этом возникает нетривиальный гамильтониан для физических полей.

В предыдущем построении была выделена роль точки σ_2 в определении функции $\lambda_0(\tau)$. Это не существенно и связано с выбором начальной точки в решении уравнения связи (20.159). Замена $\sigma_1 \mapsto \sigma_2$ в нижнем пределе этого интеграла приведет к переопределению $\lambda_0(\tau) = \lambda(\tau,\sigma_1)$.

В принципе, можно было бы ограничиться случаем $\lambda = 0$ и не добавлять граничный член. Однако исходное действие в калибровочных моделях и в моделях, инвариантных относительно общих преобразований координат, в канонической формулировке имеют вид расширенного действия $S_{\rm E}$, содержащего множители Лагранжа. При этом многие решения, важные с физической точки зрения, соответствуют $\lambda \neq 0$. Например, решение Шварцшильда соответствует нетривиальному множителю Лагранжа, роль которого играет функция хода N.

Гамильтонова формулировка

Обозначения в предыдущем разделе были выбраны не случайно. По сути дела модель (20.151) уже записана в гамильтоновой форме, при этом переменные p и p^* являются импульсами, сопряженными координатам q и q^*. Нашей исходной точкой будет полное действие

$$S_{\rm T} = \int d\tau d\sigma (p\dot{q} + p^* \dot{q}^* - \lambda G), \quad (20.165)$$

которое получится из действия (20.151), если положить $\mu = 0$. Координаты τ и σ будем считать временно́й и пространственной соответственно. В рассматриваемом случае гамильтониан системы задан единственной связью

$$H = \int d\sigma \lambda G, \qquad G := -\partial_1 q + H^*(q^*, p^*). \quad (20.166)$$

Будем считать, что интегрирование в (20.165) проводится по всей плоскости τ, σ. При этом все возникающие интегралы предполагаются сходящимися.

Таким образом, модель описывается двумя парами канонически сопряженных переменных q, p и q^*, p^*, на которые наложена одна связь $G = 0$. Уравнения движения имеют прежний вид (20.152)–(20.156) (при $\mu = 0$), где точка обозначает дифференцирование по времени. Нетрудно проверить, что связь G является связью первого рода:

$$[G, G'] = 0,$$

где штрих обозначает, что соответствующие полевые переменные рассматриваются в точке σ'. Поэтому на поверхности связей

$$\dot{G} = [G, H] \approx 0,$$

и никаких дополнительных связей в модели не возникает. Наличие связи первого рода означает, что действие (20.165) калибровочно инвариантно. Генератором калибровочных преобразований для канонических переменных является функционал

$$T = \int d\sigma \epsilon G,$$

где $\epsilon(\tau, \sigma)$ – малый параметр локальных преобразований. Бесконечно малые преобразования имеют вид

$$\begin{aligned} \delta q &= [q, T] = 0, \\ \delta p &= [p, T] = -\partial_1 \epsilon, \\ \delta q^* &= [q^*, T] = \quad \epsilon \frac{\partial H^*}{\partial p^*}, \\ \delta p^* &= [p^*, T] = -\epsilon \frac{\partial H^*}{\partial q^*}. \end{aligned} \tag{20.167}$$

Если дополнить эти преобразования преобразованием множителя Лагранжа

$$\delta\lambda = \dot{\epsilon},$$

то, как нетрудно убедиться с помощью прямой подстановки, действие (20.165) является калибровочно инвариантным (см. раздел 20.2.3). Отметим, что эволюция во времени канонических переменных (20.152)–(20.155) является в рассматриваемом случае калибровочным преобразованием с параметром $\epsilon = \lambda d\tau$.

Согласно второй теореме Нетер калибровочная инвариантность приводит к зависимости уравнений движения:

$$\partial_1 \left(\frac{\delta S_{\text{т}}}{\delta p} \right) + \frac{\delta S_{\text{т}}}{\delta q^*} \frac{\partial H^*}{\partial p^*} - \frac{\delta S_{\text{т}}}{\delta p^*} \frac{\partial H^*}{\partial q^*} - \partial_\tau \frac{\delta S_{\text{т}}}{\delta \lambda} = 0.$$

С математической точки зрения наличие калибровочной инвариантности отражается в том, что решение уравнений движения зависит от произвольной функции λ, которая ничем не фиксирована. Основным предположением в моделях с калибровочной симметрией является утверждение о том, что все физические наблюдаемые калибровочно инвариантны. В данном случае это означает, что наблюдаемые функции от канонических переменных не зависят от λ. Чтобы исключить произвол в решениях уравнений движения и исключить нефизические переменные необходимо наложить калибровочное условие. Согласно канонической процедуре фиксирования калибровки мы должны наложить одно калибровочное условие по числу связей первого рода. Выберем его в виде

$$F = p - p_0(\sigma) = 0, \tag{20.168}$$

где $p_0(\sigma)$ – произвольная функция только от σ. Скобка Пуассона связи с калибровочным условием имеет вид

$$[G, F'] = \delta'(\sigma' - \sigma),$$

где δ' обозначает производную от δ-функции. Поскольку скобка Пуассона связи с калибровочным условием не обращается в нуль при выполнении связи, то вместе они не представляют собой систему связей первого рода. В то же время пара функций G, F не представляет собой также и систему связей второго рода, поскольку $\det[G, F'] = 0$, т.к. у δ' нетривиально ядро, состоящее из констант.

Канонически сопряженные переменные q, p являются нефизическими переменными, и могут быть исключены из рассмотрения. С этой целью решим их уравнения движения и связь, как это было сделано в предыдущем разделе. В результате получим, что модель описывает одну физическую пару канонически сопряженных полей q^*, p^* с эффективным гамильтонианом $\lambda_0 H^*(q^*, p^*)$. При этом подстановка связи и калибровочного условия в действие (20.165) приводит к неверному результату, который не воспроизводит уравнения движения для физических полей q^* и p^*. Причина этого и решение проблемы то же, что и в предыдущем разделе – к действию необходимо добавить граничный член (20.163). Заметим, что при выполнении уравнения связи граничный член в исходном действии превращается в интеграл по пространству от некоторой гамильтоновой плотности для физических переменных.

Исключение нефизических полей из уравнений движения не зависело от глобальной структуры пространства-времени, т.к. при этом решаются только уравнения движения, связи и калибровочные условия, которые локальны. Нетривиальный эффективный гамильтониан из "нулевого" исходного гамильтониана (20.166) для замкнутых многообразий можно получить следующим образом. Предположим, что пространство-время является прямым произведением $\mathbb{R} \times \mathbb{S}^1$, где первый сомножитель соответствует времени, а второй – пространству. Поскольку окружность $\mathbb{S}^1$ – компактное многообразие без края, то, казалось бы, никаких граничных вкладов не возникает, и можно свободно интегрировать по частям. Однако при внимательном рассмотрении оказывается, что при таком подходе можно потерять много решений уравнений движения, представляющих физический интерес. опишем это подробнее.

Пусть пространственная координата $\sigma \in [0, 2\pi]$ параметризует окружность $\mathbb{S}^1$. Поскольку физические поля q^* и p^* ничем не ограничены, то их можно считать достаточно гладкими функциями на окружности. В то же время нефизические поля должны удовлетворять уравнению связи (20.146), которое в общем случае не имеет непрерывных решений (20.159) на окружности:

$$q(0) = 0, \qquad q(2\pi) = \int_0^{2\pi} d\sigma H^* \neq 0.$$

Кроме того, вариация нефизического поля (20.162) в общем случае не может быть определена как непрерывная функция на окружности. Это значит, что при постановке вариационной задачи необходимо сделать разрез и добавить к действию (20.165) граничный член, который и приведет к нетривиальному эффективному действию для физических переменных. Поэтому, если пространство представляет собой окружность, то в физическом действии и эффективном гамильтониане достаточно просто изменить пределы интегрирования по σ.

Покажем, что нетривиальный эффективный гамильтониан возникает также в более общих калибровках, зависящих от времени явно. Пусть калибровочное условие имеет вид

$$F = p - p_0(\tau, \sigma) = 0,$$

где функция $p_0(\tau, \sigma)$ задана. В этом случае действие на поверхности связей получит дополнительный вклад за счет слагаемого $p\dot{q}$. Нетрудно проверить, что дополнитель-

ный вклад не меняет окончательного ответа. Действительно,

$$\triangle S\Big|_{F=0,G=0} = \int d\tau d\sigma\, p\dot{q} = \int d\tau d\sigma\, (-\dot{p}q),$$

поскольку интегрирование по частям по времени τ допустимо. Используя уравнение (20.153), определяющее λ, и интегрируя по частям, получим

$$\triangle S\Big|_{F=0,G=0} = \int d\tau d\sigma(-\lambda\partial_1 q) + \int d\tau\, (\lambda q)\Big|_{\sigma=\sigma_1}^{\sigma_2}.$$

Теперь первое слагаемое воспроизводит эффективный гамильтониан λH^*, где $\lambda = \lambda(\tau,\sigma)$, а второе слагаемое сокращается с граничным членом, добавленным в физическое действие (20.163).

Таким образом, при подстановке связей и калибровочных условий в действие, необходимо проявлять осторожность. Рассмотренный пример показывает, что эффективный гамильтониан и действие для физических полей может полностью определяться граничным членом в исходном действии. Причина этого кроется в том, что связи могут не иметь решений, убывающих в бесконечности, и предположение о финитности вариации неправомерно. К сожалению, для определения явного вида граничных членов необходим глубокий анализ уравнений связей, что не всегда возможно, из-за их сложности.

Глава 21

Основы общей теории относительности

В настоящей главе мы приступим к изложению основ общей теории относительности, которая в настоящее время рассматривается в качестве основной модели гравитационных взаимодействий. После вступительного раздела, будут написаны основные уравнения и поставлена задача, которая решается в теории гравитации.

21.1 Пространство-время, метрика и гравитация

В общей теории относительности мы предполагаем, что все законы природы формулируются на четырехмерном многообразии (пространстве-времени) $\mathbb{M}$, на котором задана метрика $g = dx^\alpha \otimes dx^\beta g_{\alpha\beta}$ лоренцевой сигнатуры, $\mathsf{sign}\, g_{\alpha\beta} = \mathsf{diag}\,(+---)$. В общей теории относительности не предполагается, что пространство-время снабжено какой либо линейной структурой, как это было в механике Ньютона и специальной теории относительности.

В общем случае топология пространства-времени $\mathbb{M}$ может быть нетривиальной и отличаться от топологии пространства Минковского. Поскольку топология $\mathbb{M}$ не фиксирована, то вводится новое требование. Пространство-время, по-определению, должно быть максимально продолжено вдоль геодезических (экстремалей). Это значит, что любая геодезическая в пространстве-времени может быть продолжена либо до бесконечного значения канонического параметра в обе стороны, либо при конечном значении канонического параметра она попадет в сингулярную точку, где какой либо из геометрических инвариантов обращается в бесконечность. Поскольку канонический параметр вдоль экстремалей определен с точностью до линейных преобразований (см., главу 17), то данное требование инвариантно, т.е. не зависит от выбора системы координат.

Замечание. Требование максимального продолжения пространства-времени вдоль геодезических нельзя заменить на более жесткое требование геодезической полноты, т.к. многие важные точные решения уравнений Эйнштейна не являются геодезически полными. Например, для решений, описывающих черные дыры, времениподобные геодезические линии достигают сингулярной точки, в которой квадрат тензора кривизны обращается в бесконечность, при конечном значении собственного времени. □

Еще одно очень важное отличие общей теории относительности от специальной заключается в следующем. В специальной теории относительности метрика Лоренца

в пространстве Минковского была постулирована. В общей теории относительности постулируется, что метрика на $\mathbb{M}$ должна удовлетворять уравнениям Эйнштейна (21.1). Тем самым компоненты метрики пространства-времени удовлетворяют некоторой системе уравнений движения так же, как и все другие поля.

Все поля, кроме метрики, мы делим на поле излучения, описывающее безмассовое электромагнитное поле, и поля материи, описывающие массивные частицы. Многие современные модели математической физики содержат дополнительные безмассовые поля, например, глюоны, существование которых в природе пока экспериментально не подтверждено. Для определенности, все безмассовые поля мы также относим к полям излучения. Таким образом, все поля, за исключением метрики, мы делим на два класса: *излучение* (безмассовые поля или частицы) и *поля материи* (массивные поля или частицы).

Теперь скажем несколько слов о гравитационном взаимодействии. Для описания движения планет в солнечной системе с хорошей точностью используется механика Ньютона и закон всемирного тяготения. Мы говорим, что между планетами действуют гравитационные силы, которые определяют их движение. При этом движение происходит в плоском трехмерном евклидовом пространстве $\mathbb{R}^3$, а время играет роль параметра. Выделим из всех частиц, участвующих в гравитационном взаимодействии одну и рассмотрим ее движение в гравитационном поле, создаваемом другими телами. При этом будем считать, что гравитационное поле, которое создается выбранной частицей, пренебрежимо мало. Такую частицу назовем *пробной*. Основное свойство гравитационного взаимодействия заключается в том, что движение пробной частицы, при заданных начальных условиях, не зависит от ее массы.

Пример 21.1.1. Ускорение свободного падения на Земле не зависит от массы падающего тела. Это утверждение в настоящее время экспериментально проверено с высокой степенью точности. □

Независимость ускорения от массы частицы означает, что при одних и тех же начальных условиях траектории и мировые линии пробных частиц разной массы совпадают.

Другое важное свойство гравитационного взаимодействия заключается в том, что оно всегда носит характер притяжения. До сих пор частицы отрицательной массы в природе не наблюдались.

Рассмотрим движение пробной частицы в специальной теории относительности в инерциальной системе отсчета. По-определению, если гравитационное поле отсутствует и на частицу не действуют никакие другие силы, то она движется равномерно и прямолинейно. Теперь рассмотрим движение той же частицы, но в неинерциальной системе отсчета, которая движется, например, с постоянным ускорением относительно инерциальной системы отсчета. В этой системе координат свободная частица движется с ускорением и наблюдатель в этой системе отсчета может сказать (если не наблюдает за другими телами), что его система инерциальна, а частица движется в постоянном и однородном гравитационном поле, которое и вызывает ускорение. При этом ускорение не зависит от массы частицы и определяется только неинерциальной системой координат.

Принцип эквивалентности. Движение свободной пробной частицы в неинерциальной системе отсчета такое же, как движение частицы в инерциальной системе отсчета, но находящейся под действием некоторого гравитационного поля. □

Замечание. В дальнейшем мы увидим, что никакой эквивалентности нет. А именно, обратное утверждение неверно. Если пробная частица движется в гравитационном

поле, то в общем случае не существует такой системы координат, в которой она бы двигалась равномерно и прямолинейно. □

В инерциальной системе координат в пространстве Минковского метрика диагональна и имеет постоянные компоненты, $\eta_{\alpha\beta} = \mathsf{diag}\,(+ - - -)$. В такой системе координат все экстремали и только они являются прямыми линиями. Поэтому можно сказать, что свободная пробная частица движется вдоль одной из экстремалей пространства Минковского. Если перейти в неинерциальную (криволинейную) систему координат, то в общем случае метрика перестанет быть диагональной, и ее компоненты станут зависеть от координат точки пространства-времени. В этой системе координат траектория свободной частицы уже не будет выглядеть прямолинейной, а движение – равномерным. Тем не менее траектория, конечно, будет оставаться экстремалью пространства Минковского, т.к. понятие экстремали инвариантно и не зависит от выбора системы координат.

Замечание. В дифференциальной геометрии метрика является инвариантным объектом и определяется независимо от выбора системы координат. Метрика Лоренца не зависит от точки пространства Минковского. Однако ее компоненты могут быть непостоянными функциями от координат точки: это зависит от системы координат. □

Таким образом, утверждение о том, что свободная пробная частица движется в пространстве Минковского вдоль одной из экстремалей инвариантно относительно выбора системы координат и лежит в основе перехода от механики Ньютона к общей теории относительности. Как уже было сказано, в общей теории относительности мы предполагаем, что пространство-время представляет собой четырехмерное многообразие $\mathbb{M}$, на котором задана метрика лоренцевой сигнатуры. Мы постулируем, что любая пробная частица движется вдоль одной из экстремалей пространства-времени. Этот постулат согласуется с упомянутыми выше свойствами гравитационного взаимодействия: мировая линия пробной частицы не зависит от ее массы и частиц с отрицательной массой в природе не наблюдается. При этом принцип эквивалентности является наводящим соображением о том, что метрика с нетривиальными компонентами описывает гравитационное взаимодействие.

Тем самым метрика пространства-времени в общей теории относительности играет выделенную роль. Мы считаем, что метрика описывает гравитационные взаимодействия материальных тел и излучения. А именно, если частица движется в плоском пространстве-времени, которое изометрично пространству Минковского с лоренцевой метрикой, то на нее не действуют гравитационные силы. В этом случае частица в инерциальной системе координат движется равномерно и прямолинейно. Если гравитационное поле нетривиально, то частицы (массовые и безмассовые) движутся по экстремалям в искривленном пространстве-времени, т.е. по многообразию $\mathbb{M}$ с метрикой g и связностью Леви–Чивиты $\widetilde{\Gamma}$, для которой тензор кривизны отличен от нуля. В этом случае отсутствует понятие инерциальной системы отсчета, а экстремали отличаются от прямых линий.

Поскольку в пространстве-времени $\mathbb{M}$ задана метрика, то она однозначно определяет связность Леви–Чивиты или символы Кристоффеля. Это позволяет использовать аппарат ковариантного дифференцирования для построения инвариантов и записи ковариантных уравнений движения. Введение связности Леви–Чивиты на многообразии $\mathbb{M}$ является постулатом общей теории относительности. То есть в теории тяготения Эйнштейна мы постулируем, что кручение и неметричность аффинной связности тождественно равны нулю.

Замечание. В настоящее время теория тяготения Эйнштейна имеет много обобщений. Большой класс таких обобщений представляют собой модели, в которых на многообразии $\mathbb{M}$ помимо метрики задается также независимая аффинная связность с нетривиальным кручением и неметричностью. Эти обобщения естественны с геометрической точки зрения, т.к. метрика и аффинная связность являются совершенно независимыми геометрическими объектами. В общем случае, даже если ограничиться инвариантными лагранжианами, приводящими к уравнениям движения второго порядка, существует очень много возможностей для построения моделей, которые в настоящее время не исследованы в полной мере.

Считается, что общая теория относительности согласуется со всеми экспериментальными данными. Однако, поскольку мы не знаем экспериментальных следствий упомянутых выше геометрических обобщений теории тяготения, говорить о том, что они противоречат экспериментальным данным нельзя. Различные геометрические обобщения теории тяготения Эйнштейна представляют самостоятельный математический интерес и могут быть полезны при построении квантовой теории гравитации и единых моделей. В современной математической физике такие модели привлекают исследователей постоянно со времен создания общей теории относительности. □

21.2 Действие Гильберта–Эйнштейна

В общей теории относительности постулируется, что пространство-время является псевдоримановым многообразием $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой лоренцевой сигнатуры $g_{\alpha\beta}$. При этом считается, что метрика описывает гравитационные взаимодействия. Мы намеренно не ограничиваемся наиболее интересным случаем $n = 4$, потому что модели гравитации в большем и меньшем числе измерений также важны для приложений.

Мы рассматриваем метрику пространства-времени в качестве одной из полевых переменных и постулируем для нее *уравнения Эйнштейна:*

$$\kappa\left(\widetilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\widetilde{R}\right) + g_{\alpha\beta}\frac{n-2}{2}\Lambda = -\frac{1}{2}T_{\text{м}\alpha\beta}. \tag{21.1}$$

В левой части этой системы уравнений для метрики стоит *тензор Эйнштейна*

$$G_{\alpha\beta} = \widetilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\widetilde{R}, \tag{21.2}$$

умноженный на *гравитационную постоянную* κ, и *космологическая постоянная* Λ. В правой части стоит *тензор энергии-импульса материи* $T_{\text{м}\alpha\beta}$. Эти уравнения при $\Lambda = 0$ и $n = 4$ были впервые предложены в [**?**].

Тензор энергии-импульса материи зависит от рассматриваемой модели, и в общем случае уравнения Эйнштейна необходимо дополнить уравнениями для полей материи. То есть сама по себе система уравнений Эйнштейна неполна.

Замечание. В уравнении (21.1) мы оставили знак тильды, чтобы подчеркнуть, что тензор кривизны строится только по метрике при нулевом кручении и неметричности. То есть метрика g на $\mathbb{M}$ определяет связность Леви–Чивиты (символы Кристоффеля), которые в свою очередь задают тензор кривизны. □

Замечание. Вклад космологической постоянной в уравнения Эйнштейна (21.1) можно перенести в правую часть

$$\kappa\left(\widetilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\widetilde{R}\right) = -\frac{1}{2}\big(T_{\text{м}\alpha\beta} + T_{\Lambda\alpha\beta}\big),$$

где

$$T_{\Lambda\alpha\beta} = (n-2)\Lambda g_{\alpha\beta}$$

и рассматривать его как дополнение к тензору энергии-импульса материи $T_{\text{м}\alpha\beta}$. Сравнивая это выражение с тензором энергии-импульса непрерывной среды (21.117), его можно интерпретировать, как вклад среды с давлением $\mathcal{P} = -(n-2)\Lambda$ и плотностью энергии противоположного знака $\mathcal{E} = -\mathcal{P} = (n-2)\Lambda$. Разность знаков давления и плотности энергии не позволяет интерпретировать космологическую постоянную, как постоянное распределение некоторой обычной материи. □

Обсудим некоторые общие свойства уравнений Эйнштейна и введем терминологию.

Тензор Эйнштейна (21.2) инвариантен относительно вейлевского преобразования метрики с постоянным параметром

$$g_{\alpha\beta} \to k g_{\alpha\beta}, \qquad k = \mathsf{const} \neq 0.$$

Уравнения Эйнштейна при заданном тензоре энергии-импульса представляют собой систему из $n(n+1)/2$, где n – размерность пространства-времени, нелинейных уравнений в частных производных второго порядка для метрики. В частности, в четырехмерном пространстве-времени мы имеем десять уравнений. Уравнения Эйнштейна чрезвычайно сложны, и в настоящее время известны лишь отдельные классы решений, часть из которых будет обсуждаться в дальнейшем.

Уравнения Эйнштейна можно переписать в другом виде. След (21.1) эквивалентен уравнению

$$\kappa\widetilde{R} = n\Lambda + \frac{1}{n-2}T_{\text{м}},$$

где $T_{\text{м}} = T_{\text{м}\alpha}{}^{\alpha}$ – след тензора энергии-импульса материи. Исключив скалярную кривизну из (21.1) с помощью этого равенства, получим эквивалентную систему уравнений

$$\kappa\widetilde{R}_{\alpha\beta} - \Lambda g_{\alpha\beta} = -\frac{1}{2}\rho_{\alpha\beta}. \tag{21.3}$$

где

$$\rho_{\alpha\beta} := T_{\text{м}\alpha\beta} - \frac{1}{n-2}g_{\alpha\beta}T_{\text{м}}.$$

Пространство-время называется *пустым*, если тензор энергии-импульса материи всюду равен нулю. В этом случае уравнения Эйнштейна (21.3) принимают вид

$$\kappa\widetilde{R}_{\alpha\beta} = \Lambda g_{\alpha\beta}. \tag{21.4}$$

Это – *вакуумные уравнения Эйнштейна* с космологической постоянной. Отсюда следует, что скалярная кривизна пустого пространства постоянна:

$$\widetilde{R} = \frac{n\Lambda}{\kappa}.$$

Замечание. Коэффициент перед космологической постоянной в уравнениях Эйнштейна (21.1) подобран таким образом, чтобы вакуумные уравнения Эйнштейна имели вид (21.4) и не зависели от размерности пространства-времени. □

При ненулевой космологической постоянной уравнения (21.4) означают, что тензор Риччи пропорционален метрике. Частным случаем таких пространств являются пространства постоянной кривизны, рассмотренные в разделе **??**. При нулевой космологической постоянной, $\Lambda = 0$, пустое пространство является *Риччи плоским*:

$$\widetilde{R}_{\alpha\beta} = 0. \tag{21.5}$$

Следовательно, в этом случае скалярная кривизна также равна нулю, $\widetilde{R} = 0$.

Пример 21.2.1. Для метрики Лоренца тензор кривизны равен нулю. Следовательно, пространство Минковского является Риччи плоским. Ясно, что метрика Лоренца удовлетворяет вакуумным уравнениям Эйнштейна с нулевой космологической постоянной. □

В двумерном пространстве-времени полный тензор кривизны однозначно восстанавливается по скалярной кривизне, а в трехмерном – по тензору Риччи и скалярной кривизне. Следовательно, полный тензор кривизны равен нулю, если выполнено условие (21.5). Это значит, что двумерное и трехмерное Риччи плоское пространство локально является пространством Минковского. То есть может быть либо пространством Минковского, либо цилиндром или тором. В четырех измерениях и выше равенства нулю тензора Риччи недостаточно для обращения в нуль полного тензора кривизны.

Пример 21.2.2. Решение Шварцшильда, описывающее черную и белую дыру, является Риччи плоским, но полный тензор кривизны отличен от нуля. А именно, отличен от нуля тензор Вейля. □

Физическая интерпретация уравнений Эйнштейна при нулевой космологической постоянной следующая. В общей теории относительности постулируется, что метрика пространства-времени не является метрикой Лоренца, а находится, как решение уравнений Эйнштейна. Таким образом, пространство-время представляет собой псевдоримановo многообразие с метрикой специального вида, удовлетворяющей уравнениям (21.1). Эти пространства называются *пространствами Эйнштейна*. Следующий постулат состоит в том, что пробные частицы под действием гравитационных сил двигаются по экстремалям в пространстве Эйнштейна. При этом в правой части уравнений Эйнштейна подразумевается тензор энергии-импульса всей остальной материи. При этом мы говорим следующее. Пустое пространство при нулевой космологической постоянной и отсутствии гравитационных волн является пространством Минковского, и точечные частицы двигаются по прямым линиям. Это соответствует отсутствию сил тяготения. При наличии полей материи в уравнениях Эйнштейна появляется нетривиальная правая часть, что приводит к тому, что пространство-время становится нетривиальным псевдоримановым многообразием. В этом пространстве-времени экстремали уже не являются прямыми линиями, что интерпретируется, как наличие сил тяготения. Мы говорим, что пробная частица движется в поле тяготения, созданном остальной материей. При этом закон всемирного тяготения является следствием уравнений Эйнштейна в определенном приближении, которое рассмотрено в разделе **??**.

Существующие экспериментальные данные свидетельствуют о том, что в отсутствие сил тяготения пространство-время в масштабах солнечной системы близко к пространству Минковского. Это значит, что если космологическая постоянная существует, то является в определенном смысле малой величиной. Отметим, что равенство или неравенство космологической постоянной нулю имеет принципиальное

значение. Действительно, наличие даже малой космологической постоянной приводит к тому, что метрика пространства Минковского уже не будет удовлетворять вакуумным уравнениям Эйнштейна.

Теперь обсудим принцип наименьшего действия для уравнений Эйнштейна. Левую часть уравнений (21.1) можно получить из *действия Гильберта–Эйнштейна* [**?**, **?**]

$$S_{\text{не}} = \int_{\mathbb{M}} d^n x \sqrt{|g|} \left(\kappa \widetilde{R} - (n-2)\Lambda \right), \tag{21.6}$$

где интегрирование ведется по всему пространству-времени $\mathbb{M}$. Конечно, мы предполагаем, что интеграл сходится. Это действие было впервые предложено Д. Гильбертом в 1915 году в четырехмерном пространстве-времени. Он предложил действие в более общем виде, включающем также электромагнитное поле [**?**]. Несколько позже А. Эйнштейн тоже рассмотрел это действие для вывода уравнений общей теории относительности в такой системе координат, где $\det g_{\alpha\beta} = 1$ [**?**].

В следующем разделе мы покажем, что вариационная производная действия Гильберта–Эйнштейна по метрике имеет вид

$$\frac{\delta S_{\text{не}}}{\delta g_{\alpha\beta}} = -\sqrt{|g|}\kappa \left(\widetilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \widetilde{R} \right) - \sqrt{|g|} \frac{n-2}{2} \Lambda g^{\alpha\beta}. \tag{21.7}$$

При доказательстве этого равенства были отброшены все граничные вклады, возникающие при интегрировании по частям. Это подразумевает либо отсутствие границы пространства-времени, либо достаточно быстрое убывание всех полей вблизи границы. К последнему случаю мы относим как границу пространства-времени, лежащую на конечном расстоянии относительно метрики $g_{\alpha\beta}$, так и бесконечно удаленную границу. В последнем случае говорят, что все поля достаточно быстро убывают на бесконечности. Однако это предположение может противоречить уравнениям Эйнштейна, которые допускают решения с недостаточно быстрым убыванием полей. Соответствующие граничные вклады в действие будут обсуждаться в разделе **??**.

При наличии полей материи чаще удобнее варьировать по обратной метрике, что приводит к изменению знака вариационной производной

$$\frac{\delta S_{\text{не}}}{\delta g^{\alpha\beta}} = \sqrt{|g|}\kappa \left(\widetilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \widetilde{R} \right) + \sqrt{|g|} \frac{n-2}{2} \Lambda g_{\alpha\beta}. \tag{21.8}$$

Полное действие для гравитационного поля и полей материи имеет вид суммы

$$S = S_{\text{не}} + S_{\text{м}}, \tag{21.9}$$

где $S_{\text{м}}$ – действие для полей материи. Обычно действие для полей материи в теории гравитации получают путем *минимальной подстановки*: выбирают лоренц-инвариантное действие в пространстве Минковского, заменяют лоренцеву метрику на псевдоримано-ву $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}(x)$, обычные производные – на ковариантные $\partial_\alpha \mapsto \nabla_\alpha$, и умножают лагранжиан на определитель репера $\sqrt{|g|}$, чтобы получить инвариантную меру интегрирования. В результате получим действие для полей материи, инвариантное относительно общих преобразований координат. Сравнивая правую часть уравнений Эйнштейна (21.1) с вариационной производной (21.8), получаем выражение для тензора энергии-импульса материи

$$T_{\text{м}\alpha\beta} = \frac{2}{\sqrt{|g|}} \frac{\delta S_{\text{м}}}{\delta g^{\alpha\beta}}. \tag{21.10}$$

Эту вариационную производную часто принимается за определение *тензора энергии-импульса* полей материи в общей теории относительности. При таком определении тензор энергии-импульса всегда симметричен. В ряде случаев, например, для скалярного и калибровочного полей, это определение совпадает с ковариантным обобщением канонического тензора энергии-импульса, т.е. получается из выражения (19.43) путем минимальной подстановки. Однако в общем случае это не так, потому что действие для полей материи (например, спинорных полей) не всегда может быть выражено через метрику.

Размерности постоянных и полей

В теории поля важную роль играет анализ размерностей. Он помогает контролировать проведение вычислений путем сравнения размерностей различных слагаемых и в некоторых случаях делать общие выводы. В квантовой теории поля, например, размерность констант связи позволяет судить о перенормируемости моделей. Здесь принят упрощенный вариант подсчета размерностей, где все измеряется в единицах длины. Однако, для сравнения общей теории относительности с теорией гравитации Ньютона нам этого будет недостаточно. Поэтому мы опишем оба подхода к определению размерностей.

В квантовой теории поля, по-определению, действие и метрика являются безразмерными величинами, а координаты имеют размерность длины (скорость света мы принимаем за единицу):

$$[S_{\text{не}}] := [g_{\alpha\beta}] := 1, \qquad [x^\alpha] := l.$$

Тогда скалярная кривизна имеет размерность

$$[\widetilde{R}] = l^{-2},$$

т.к. содержит две производные. Отсюда следует, что гравитационная и космологическая постоянные являются размерными величинами:

$$[\kappa] = l^{2-n}, \qquad [\Lambda] = l^{-n}.$$

Для контроля вычислений такого описания размерностей, как правило, достаточно.

Теперь опишем размерности величин, как это принято в системе СГС, которые нам понадобятся в дальнейшем. При этом мы считаем, что размерность пространства-времени равна четырем, $n = 4$. Исходными являются размерности массы (грамм), расстояния (сантиметр) и времени (секунда):

$$[m] := \text{г}, \qquad [x^\mu] := \text{см}, \qquad [t] := \text{сек},$$

где индекс $\mu = 1, 2, 3$ пробегает только пространственные значения. Координата $x^0 := ct$, где c – скорость света, также измеряется в сантиметрах. По-определению, компоненты метрики безразмерны:

$$[g_{\alpha\beta}] = [g^{\alpha\beta}] := 1.$$

Поскольку кривизна содержит две производные по координатам, то

$$[\widetilde{R}_{\alpha\beta\gamma\delta}] = [\widetilde{R}_{\alpha\beta}] = [\widetilde{R}] := \frac{1}{\text{см}^2}.$$

Действие, как это принято в механике, имеет ту же размерность, что и произведение импульса на скорость $p\,dq$ или энергии на время Edt:

$$[S] := \frac{\text{г} \cdot \text{см}^2}{\text{сек}}.$$

Учитывая, что в действии $[dt\,d^3x] = \text{сек} \cdot \text{см}^3$, определяем размерность гравитационной постоянной:

$$[\kappa] = \frac{\text{г} \cdot \text{см}}{\text{сек}^2}. \tag{21.11}$$

В разделе 21.9 мы сравним эту гравитационную постоянную с той, которая входит во всемирный закон тяготения. Наконец, размерность космологической постоянной равна

$$[\Lambda] = \frac{\text{г}}{\text{см} \cdot \text{сек}^2}.$$

В дальнейшем для упрощения формул мы часто будем полагать $c = 1$ и $\kappa = 1$. Там, где это необходимо, степени скорости света легко восстановить, исходя из соображений размерности.

21.3 Вариация действия Гильберта–Эйнштейна

Докажем равенство (21.7) в более общем виде, который полезен при рассмотрении моделей, основанных на геометрии Римана–Картана или аффинной геометрии. А именно, рассмотрим инвариантное действие

$$S = \int_{\mathbb{M}} d^n x\, L = \int_{\mathbb{M}} d^n x\, \sqrt{|g|}\varphi R, \tag{21.12}$$

зависящее от скалярного поля $\varphi(x) \in \mathcal{C}^2(\mathbb{M})$ и скалярной кривизны $R(g, \Gamma)$, построенной по метрике $g_{\alpha\beta}$ и аффинной связности общего вида $\Gamma_{\alpha\beta}{}^{\gamma}$. Мы предполагаем, что компоненты метрики и связности являются достаточно гладкими функциями, и интеграл (21.12) сходится. Кроме этого мы предположим, что всеми граничными слагаемыми, возникающими при интегрировании по частям, можно пренебречь.

Замечание. Подстановка в действие (21.12) римановой кривизны $\widetilde{R}$, зависящей только от метрики, приводит к чрезвычайно трудоемкой вариационной задаче. Это связано с тем, что при дифференцировании по частям необходимо дифференцировать также и скалярное поле. Поскольку скалярная кривизна $\widetilde{R}$ содержит вторые производные от метрики, то дифференцировать по частям необходимо два раза, и это приводит к большому числу слагаемых. Значительное упрощение вносят последовательные действия. Сначала варьируем по метрике $g_{\alpha\beta}$ и связности $\Gamma_{\alpha\beta}{}^{\gamma}$, рассматривая их, как независимые переменные, а затем подставляем вариацию связности, выраженную через вариацию метрики, тензора кручения и неметричности. В общей теории относительности такой подход называется *формализмом первого порядка.* □

Начнем с нескольких вспомогательных формул, необходимых в дальнейшем. Варьируя определение обратной метрики

$$g^{\alpha\beta} g_{\beta\gamma} = \delta^{\alpha}_{\gamma},$$

получаем тождество

$$\delta g^{\alpha\beta} g_{\beta\gamma} + g^{\alpha\beta} \delta g_{\beta\gamma} = 0.$$

Отсюда следует связь между вариацией самой метрики и ее обратной:

$$\delta g^{\alpha\beta} = -g^{\alpha\gamma} g^{\beta\delta} \delta g_{\gamma\delta}. \tag{21.13}$$

Из теории матриц известно, что для произвольной квадратной матрицы $A = \{A_{\alpha\beta}\}$ справедливо тождество

$$\delta \det A = \det A A^{-1\alpha\beta} \delta A_{\alpha\beta}.$$

Отсюда следует, что вариация определителя метрики $g := \det g_{\alpha\beta}$ равна

$$\delta g = g g^{\alpha\beta} \delta g_{\alpha\beta} = -g g_{\alpha\beta} \delta g^{\alpha\beta}. \tag{21.14}$$

Эту вариацию мы записали в двух видах, так как в приложениях часто бывает удобнее варьировать действие не по самой метрики, а по ее обратной. Наличие квадратного корня в мере объема $\sqrt{|g|} = \sqrt{|\det g_{\alpha\beta}|}$ приводит к появлению множителя $1/2$. Поэтому для ее вариации справедливы равенства

$$\delta\sqrt{|g|} = \frac{1}{2}\sqrt{|g|} g^{\alpha\beta} \delta g_{\alpha\beta} = -\frac{1}{2}\sqrt{|g|} \delta g^{\alpha\beta} g_{\alpha\beta}. \tag{21.15}$$

Приступим к вариации действия (21.12). Вариационная производная по скалярному полю φ очевидна

$$S_{,\varphi} := \frac{\delta S}{\delta \varphi} = \sqrt{|g|} R. \tag{21.16}$$

Метрика входит в действие (21.12) дважды: в форму объема и в определение скалярной кривизны (7.78), причем без производных. Поэтому нетрудно проверить, что

$$S_{,}{}^{\alpha\beta} := \frac{\delta S}{\delta g_{\alpha\beta}} = -\sqrt{|g|}\varphi \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right). \tag{21.17}$$

Вариация действия по аффинной связности $\Gamma_{\alpha\beta}{}^{\gamma}$, все компоненты которой рассматриваются, как независимые переменные, более трудоемка. Это связано с тем, что приходится интегрировать по частям, так как тензор кривизны (7.65) зависит от производных аффинной связности. Прямые вычисления приводят к следующему выражению для вариации

$$\begin{aligned} \delta L = &-\sqrt{|g|}\partial_\alpha \varphi g^{\alpha\gamma} \delta\Gamma_{\beta\gamma}{}^{\beta} + \sqrt{|g|}\partial_\gamma \varphi g^{\alpha\beta} \delta\Gamma_{\alpha\beta}{}^{\gamma} - \sqrt{|g|}\varphi \left(Q_\alpha{}^{\alpha\gamma} - \frac{1}{2} Q^\gamma \right) \delta\Gamma_{\beta\gamma}{}^{\beta} \\ &+ \sqrt{|g|}\varphi \left[-\left(T_{\delta\gamma}{}^{\delta} + \frac{1}{2} Q_\gamma \right) g^{\alpha\beta} - T_\gamma{}^{\beta\alpha} - T_\gamma{}^{\alpha\beta} + Q_\gamma{}^{\alpha\beta} \right] \delta\Gamma_{\alpha\beta}{}^{\gamma}, \end{aligned} \tag{21.18}$$

где $T_{\alpha\beta}{}^{\gamma}$ и $Q_{\alpha\beta\gamma}$ – тензоры кручения и неметричности. Для облегчения вычислений следует помнить, что выражение, стоящее перед вариацией связности, должно быть тензорным полем. Это поможет правильно сгруппировать слагаемые. Заметим также, что выражение, стоящее в квадратных скобках, симметрично по индексам α и β.

Теперь вычислим очень важную для приложений вариационную производную действия (21.12) по метрике $g_{\alpha\beta}$ в римановой геометрии. Обозначим соответствующее действие, зависящее только от скалярного поля и метрики, через

$$\widetilde{S} = \int_{\mathbb{M}} d^n x\, \widetilde{L} = \int_{\mathbb{M}} d^n x\, \sqrt{|g|}\varphi \widetilde{R}. \tag{21.19}$$

Поскольку кручение и неметричность в римановой геометрии равны нулю, то из (21.17) и (21.18) следует, что вариация подынтегрального выражения в действии по метрике равна

$$\delta\widetilde{L} = -\sqrt{|g|}\varphi\left(\widetilde{R}^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}\widetilde{R}\right) - \sqrt{|g|}\partial_\gamma\varphi g^{\gamma\beta}\delta\widetilde{\Gamma}_{\alpha\beta}{}^{\alpha} + \sqrt{|g|}\partial_\gamma\varphi g^{\alpha\beta}\delta\widetilde{\Gamma}_{\alpha\beta}{}^{\gamma}.$$

Выразив вариацию связности $\delta\widetilde{\Gamma}_{\alpha\beta}{}^{\gamma}$ через вариацию метрики, проинтегрировав по частям и приведя подобные члены, получим окончательное выражение для вариационной производной

$$\widetilde{S},^{\alpha\beta} := \frac{\delta\widetilde{S}}{\delta g_{\alpha\beta}} = -\sqrt{|g|}\varphi\left(\widetilde{R}^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}\widetilde{R}\right) + \sqrt{|g|}\big(\widetilde{\Box}\varphi g^{\alpha\beta} - \widetilde{\nabla}^{\alpha}\widetilde{\nabla}^{\beta}\varphi\big). \tag{21.20}$$

Напомним, что в римановой геометрии ввиду симметрии символов Кристоффеля по нижним индексам вторая ковариантная производная от скалярного поля симметрична: $\widetilde{\nabla}^{\alpha}\widetilde{\nabla}^{\beta}\varphi = \widetilde{\nabla}^{\beta}\widetilde{\nabla}^{\alpha}\varphi$ (сравните с (7.83)).

Если скалярное поле равно единице, $\varphi = 1$, то действие (21.19) совпадает с действием Гильберта–Эйнштейна (21.6) без космологической постоянной, и мы получаем выражение для вариационной производной (21.7).

Формула для вариационной производной (21.20) со скалярным полем важна, например, при вычислении скобок Пуассона в канонической формулировке общей теории относительности.

Вариация действия (21.12) была проведена в пространстве произвольной размерности. Отметим, что в двумерном случае тензор Эйнштейна (21.2) тождественно равен нулю, и первое слагаемое в (21.20) отсутствует (см. раздел 24.1).

21.4 Зависимость уравнений Эйнштейна

В настоящем разделе мы считаем, что кручение и неметричность равны нулю, а связностью является связность Леви–Чивиты, построенная по заданной метрике.

Важным обстоятельством в общей теории относительности является линейная зависимость уравнений Эйнштейна (21.1). Предположим, что эти уравнения получены вариацией по метрике действия (21.9), которое инвариантно относительно общих преобразований координат. Если действие инвариантно относительно локальных преобразований, то согласно второй теореме Нетер между уравнениями движения существует линейная зависимость (19.63). Рассмотрим эту зависимость в случае общих преобразований координат. Для простоты предположим, что действие полей материи зависит только от некоторого набора скалярных полей $\varphi^a(x)$, $a = 1, \ldots, N$. При бесконечно малых преобразованиях координат с параметром $\epsilon^\alpha(x)$ метрика и скалярные поля преобразуются по правилам (3.102) и (3.96)

$$\begin{aligned}\delta g_{\alpha\beta} &= -\widetilde{\nabla}_\alpha\epsilon_\beta - \widetilde{\nabla}_\beta\epsilon_\alpha,\\ \delta\varphi^a &= -\epsilon^\alpha\partial_\alpha\varphi^a,\end{aligned}$$

где $\epsilon_\alpha = g_{\alpha\beta}\epsilon^\beta$. Следовательно инвариантность действия записывается в виде

$$\delta S = \int dx\left(\frac{\delta S}{\delta g_{\alpha\beta}}(-2\widetilde{\nabla}_\beta\epsilon_\alpha) + \frac{\delta S}{\delta\varphi^a}(-\epsilon^\alpha\partial_\alpha\varphi^a)\right) = 0.$$

После интегрирования по частям первого слагаемого получаем искомую зависимость уравнений движения

$$2\widetilde{\nabla}_\beta\left(\frac{\delta(S_{\text{не}}+S_{\text{м}})}{\delta g_{\alpha\beta}}\right)-\partial_\alpha\varphi^a\frac{\delta S_{\text{м}}}{\delta\varphi^a}=0,$$

т.к. действие Гильберта–Эйнштейна не зависит от полей материи. Это – тождества, которые выполняются независимо от того удовлетворяют поля уравнениям движения или нет. Поскольку каждое слагаемое в действии инвариантно само по себе, то выполняются два тождества:

$$\widetilde{\nabla}_\beta G^{\beta\alpha}=0, \tag{21.21}$$

$$\sqrt{|g|}\widetilde{\nabla}_\beta T^\beta_{\text{м}\,\alpha}-\frac{\delta S_{\text{м}}}{\delta\varphi^a}\partial_\alpha\varphi^a=0. \tag{21.22}$$

Первое из этих уравнений представляет собой свернутые тождества Бианки (7.108), а второе – ковариантный “закон сохранения” тензора энергии-импульса материи. Действительно, если выполнены уравнения для полей материи,

$$\frac{\delta S_{\text{м}}}{\delta\varphi^a}=0,$$

то дивергенция тензора энергии-импульса материи обращается в нуль

$$\widetilde{\nabla}_\beta T^\beta_{\text{м}\,\alpha}=0. \tag{21.23}$$

Нетрудно видеть, что аналогичные выкладки можно проделать для любого набора полей материи. При этом второе слагаемое в (21.22) может усложниться, но оно всегда будет пропорционально уравнениям движения для полей материи. Единственное условие – это инвариантность действия. Таким образом получаем следующее утверждение.

Теорема 21.4.1. *Если действие полей материи инвариантно относительно общих преобразований координат и поля материи удовлетворяют своим уравнениям Эйлера–Лагранжа, то дивергенция тензора энергии-импульса (21.23) равна нулю.*

На формулу (21.23) можно взглянуть с другой точки зрения. Допустим, что нам заданы уравнения Эйнштейна (21.1), а про инвариантное действие, приводящее к этим уравнениям, ничего не известно. Уравнения Эйнштейна – это система дифференциальных уравнений на метрику, и у них есть условия интегрируемости. Чтобы их получить возьмем ковариантную производную от обеих частей уравнений Эйнштейна. Дивергенция тензора Эйнштейна равна нулю (21.21) как следствие тождеств Бианки (7.108). Дивергенция метрики тоже равна нулю, т.к. связность Леви–Чивиты является метрической. Следовательно, ковариантный “закон сохранения” тензора энергии-импульса материи (21.23) является условием интегрируемости системы дифференциальных уравнений Эйнштейна для метрики (21.1). Это важно учитывать в тех случаях, когда тензор энергии-импульса материи получен не из принципа наименьшего действия, а из каких либо других соображений.

Пример 21.4.1. Если в качестве материи рассматривать жидкость или газ, для которой уравнения движения не следуют из принципа наименьшего действия, то условие (21.23) является независимым уравнением. □

21.5 Действие для полей материи в обобщенных моделях гравитации

В настоящем разделе мы покажем, что при минимальной подстановке ковариантное обобщение канонического тензора энергии-импульса материи является источником в уравнении для репера $e_\alpha{}^a$, а ковариантное обобщение спинового момента полей материи – источником в уравнении для лоренцевой связности $\omega_{\alpha a}{}^b$.

В различных обобщенных моделях гравитации мы обычно предполагаем, что инвариантное действие состоит из двух слагаемых

$$S = S_{\text{G}} + S_{\text{M}}, \tag{21.24}$$

где S_{G} – гравитационная часть действия и S_{M} – действие для полей материи, в которое для простоты мы включили также калибровочные поля (электромагнитное поле и поле Янга–Миллса). В общей теории относительности $S_{\text{G}} = S_{\text{HE}}$ – это действие Гильберта–Эйнштейна (21.6), равное интегралу от скалярной кривизны с возможным добавлением космологической постоянной. Это действие зависит только от метрики или репера. В более общих моделях гравитационная часть действия может включать также инварианты более высокого порядка по кривизне, кручению, тензору неметричности и их ковариантных производных. В таких случаях мы рассматриваем в качестве независимых переменных, по которым производится варьирование, переменные Картана: репер $e_\alpha{}^a$ и линейная $\mathbb{GL}(n,\mathbb{R})$-связность $\omega_{\alpha a}{}^b$ (см. раздел 6.4).

Если действие для полей материи может быть выражено через метрику и аффинную связность, то в качестве независимых переменных можно рассматривать также метрику, кручение и неметричность. Однако это не всегда имеет место. Например, для спинорного поля в геометрии Римана–Картана лоренцева связность не может быть выражена через метрику и аффинную связность. В этом случае введение репера необходимо (**??**). Таким образом, в общем случае использование переменных Картана предпочтительнее.

В настоящем разделе мы обсудим общие свойства уравнений движения для полей материи, не используя конкретный вид гравитационной части действия S_{G}.

Остановимся более подробно на действии S_{M} для полей материи. К полям материи в настоящем разделе мы относим скалярные, спинорные поля, электромагнитное поле, поля Янга–Миллса и все другие поля, кроме репера и линейной связности. Обозначим всю совокупность полей материи через $\varphi = \{\varphi^{\text{A}}\}$, $\text{A} = 1, 2, \ldots$. Пусть действие для полей материи в плоском пространстве-времени Минковского $\mathbb{R}^{1,n-1}$ с координатами x^a, $a = 0, 1, \ldots, n-1$, имеет вид

$$S_{\text{M}} = \int_{\mathbb{R}^{1,n-1}} d^n x L_{\text{M}}(\eta, \hat{\eta}, \varphi, \partial\varphi),$$

где лагранжиан полей материи L_{M} зависит от метрики Минковского $\eta = \{\eta_{ab}\}$, инвариантной метрики в пространстве-мишени для полей материи $\hat{\eta} = \{\eta_{\text{AB}}\}$ (во многих случаях это просто символ Кронекера $\eta_{\text{AB}} = \delta_{\text{AB}}$), полей материи $\varphi = \{\varphi^{\text{A}}\}$ и их частных производных первого порядка $\partial\varphi = \{\partial_a \varphi^{\text{A}}\}$.

В моделях гравитации мы предполагаем, что пространство-время $\mathbb{M}$ является многообразием той же размерности, что и исходное пространство Минковского $\mathbb{R}^{1,n-1}$. Обозначим локальные координаты на $\mathbb{M}$ через x^α, $\alpha = 0, 1, \ldots, n-1$. В простейшем случае инвариантное действие для полей материи получается из действия в про-

странстве Минковского с помощью *минимальной подстановки*:

$$\begin{aligned}
\mathbb{R}^{1,n-1} \to \mathbb{M}, \qquad & \dim \mathbb{M} = n, \\
L_{\text{м}} &\mapsto \sqrt{|g|} L_{\text{м}}, \\
\eta_{ab} &\mapsto g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b g_{ab}, \\
\eta_{\text{AB}} &\mapsto g_{\text{AB}}(x), \\
\varphi^{\text{A}} &\mapsto \varphi^{\text{A}}, \\
\partial_a \varphi^{\text{A}} \mapsto \nabla_\alpha \varphi^{\text{A}}, \qquad & \nabla_\alpha \varphi^{\text{A}} := \partial_\alpha \varphi^{\text{A}} + \omega_{\alpha\text{B}}{}^{\text{A}} \varphi^{\text{B}}.
\end{aligned}$$

В случае общей линейной группы преобразований $\mathbb{GL}(n,\mathbb{R})$ инвариантной метрики в пространстве-мишени не существует (если поля материи преобразуются по точному представлению $\mathbb{GL}(n,\mathbb{R})$), и мы вынуждены рассматривать нетривиальную метрику в пространстве-мишени $g_{\text{AB}}(x)$ как дополнительное поле, по которому, однако, варьирование действия не проводится. Метрика g_{ab} – это произвольная симметричная невырожденная матрица лоренцевой сигнатуры, с помощью которой определяется репер.

Пример 21.5.1. Рассмотрим набор массивных скалярных полей одинаковой массы, лагранжиан имеет вид (см. раздел 23.1)

$$L_{\text{м}} = \frac{1}{2}\eta^{ab}\partial_a\varphi^{\text{A}}\partial_b\varphi^{\text{B}}\eta_{\text{AB}} - \frac{1}{2}m^2\varphi^{\text{A}}\varphi^{\text{B}}\eta_{\text{AB}}. \tag{21.25}$$

Мы считаем, что скалярные поля преобразуются по некоторому представлению группы Лоренца (индекс A), и η_{AB} – инвариантная метрика. Этот лагранжиан инвариантен относительно глобальных преобразований Лоренца, действующих на координаты x^a и поля φ^{A}. После минимальной подстановки он примет вид

$$L_{\text{м}} = \sqrt{|g|}\left(\frac{1}{2}g^{\alpha\beta}\nabla_\alpha\varphi^{\text{A}}\nabla_\beta\varphi^{\text{B}}g_{\text{AB}} - \frac{1}{2}m^2\varphi^{\text{A}}\varphi^{\text{B}}g_{\text{AB}}\right),$$

где ковариантная производная определяется линейной связностью (см. раздел 6.2)

$$\nabla_\alpha\varphi^{\text{A}} = \partial_\alpha\varphi^{\text{A}} + \omega_{\alpha\text{B}}{}^{\text{A}}\varphi^{\text{B}}, \qquad \omega_{\alpha\text{B}}{}^{\text{A}} = \omega_\alpha{}^{ab}L_{ab\text{B}}{}^{\text{A}}, \tag{21.26}$$

g_{AB} – метрика в пространстве-мишени и $g^{\alpha\beta} = e^\alpha{}_a e^\beta{}_b g^{ab}$. В приведенной формуле $L_{ab\text{B}}{}^{\text{A}}$ – представление генераторов линейной группы для выбранного набора скалярных полей. Соответствующее действие инвариантно относительно общих преобразований координат и локальных $\mathbb{GL}(n,\mathbb{R})$ вращений в пространстве-мишени:

$$\begin{aligned}
e_\alpha{}^b &\mapsto e_\alpha{}^b S^{-1}{}_b^a, \\
g_{ab} &\mapsto S_a{}^c S_b{}^d g_{cd}, \\
\omega_{\alpha a}{}^b &\mapsto S_a{}^c \omega_{\alpha c}{}^d S^{-1}{}_d^b + \partial_\alpha S_a{}^c S^{-1}{}_c^b, \\
\varphi^{\text{A}} &\mapsto \varphi^{\text{B}} S^{-1}{}_{\text{B}}^{\text{A}}, \\
g_{\text{AB}} &\mapsto S_{\text{A}}{}^{\text{C}} S_{\text{B}}{}^{\text{D}} g_{\text{CD}},
\end{aligned}$$

где $S_a{}^b(x) \in \mathbb{GL}(n,\mathbb{R})$ – матрица локальных вращений и $S_{\text{A}}{}^{\text{B}}$ – ее представление, которое соответствует выбранному набору скалярных полей. Если ограничить группу $\mathbb{GL}(n,\mathbb{R})$ до группы Лоренца $\mathbb{O}(1,n-1)$, то в качестве метрики $g_{ab}(x)$ можно выбрать инвариантную метрику Лоренца η_{ab}, которая уже не будет зависеть от точки многообразия. Ей соответствует некоторая инвариантная метрика η_{AB} в пространстве-мишени. □

Из общего выражения для действия (21.24) следуют уравнения Эйлера–Лагранжа для геометрических переменных $e_\alpha{}^a$, $\omega_{\alpha a}{}^b$ и полей материи φ^{A}:

$$\frac{\delta S}{\delta e_\alpha{}^a} = 0, \tag{21.27}$$

$$\frac{\delta S}{\delta \omega_{\alpha a}{}^b} = 0, \tag{21.28}$$

$$\frac{\delta S}{\delta \varphi^{\text{A}}} = \frac{\delta S_{\text{M}}}{\delta \varphi^{\text{A}}} = 0. \tag{21.29}$$

Поскольку при минимальной подстановке репер входит в действие полей материи только в качестве общего множителя $\sqrt{|g|} = \det e_\alpha{}^a$ и в ковариантную производную с латинским индексом $\nabla_a \varphi^{\text{A}} = e^\alpha{}_a \nabla_\alpha \varphi^{\text{A}}$, то первое уравнение можно переписать в виде

$$\frac{\delta S_{\text{G}}}{\delta e_\alpha{}^a} = \sqrt{|g|}\frac{\partial L_{\text{M}}}{\partial(\nabla_\alpha \varphi^{\text{A}})} e^\beta{}_a \nabla_\beta \varphi^{\text{A}} - \sqrt{|g|} e^\alpha{}_a L_{\text{M}} = \sqrt{|g|} e^\beta{}_a T_{\text{M}\beta}{}^\alpha, \tag{21.30}$$

где

$$T_{\text{M}\beta}{}^\alpha = -\frac{1}{\sqrt{|g|}}\frac{\delta S_{\text{M}}}{\delta e_\alpha{}^a} e_\beta{}^a = \nabla_\beta \varphi^{\text{A}} \frac{\partial L_{\text{M}}}{\partial(\nabla_\alpha \varphi^{\text{A}})} - \delta^\alpha_\beta L_{\text{M}}.$$

Сравнивая последнее выражение с каноническим тензором энергии-импульса (19.43), полученным из теоремы Нетер, мы видим, что в правой части гравитационных уравнений (21.30), полученных после варьирования по реперу, стоит его ковариантное обобщение. Поэтому говорят, что тензор энергии-импульса материи является источником для репера.

Линейная связность $\omega_\alpha{}^{ab}$ входит в действие полей материи только через ковариантную производную. Поэтому уравнение (21.28) можно переписать в виде

$$\frac{\delta S_{\text{G}}}{\delta \omega_\alpha{}^{ab}} = -\sqrt{|g|}\frac{\partial L_{\text{M}}}{\partial(\nabla_\alpha \varphi^{\text{A}})} L_{ab\text{B}}{}^{\text{A}} \varphi^{\text{B}} = \sqrt{|g|} S_{\text{M}ab}{}^\alpha,$$

где

$$S_{\text{M}ab}{}^\alpha = -\frac{\partial L_{\text{M}}}{\partial(\nabla_\alpha \varphi^{\text{A}})} L_{ab\text{B}}{}^{\text{A}} \varphi^{\text{B}}. \tag{21.31}$$

В геометрии Римана–Картана общая линейная группа $\mathbb{GL}(n, \mathbb{R})$ сужается до группы Лоренца $\mathbb{O}(1, n-1)$, линейная связность $\omega_\alpha{}^{ab}$ становится лоренцевой связностью, $L_{ab\text{B}}{}^{\text{A}}$ – генераторы группы Лоренца, $g_{ab} = \eta_{ab}$, g_{AB} – инвариантная метрика в пространстве-мишени, которая всегда существует, т.к. группа Лоренца проста. В этом случае выражение (21.31) является ковариантным обобщением спинового момента полей материи (19.53). Поэтому мы говорим, что спиновый момент полей материи является источником для лоренцевой связности.

Выбор переменных Картана для моделей гравитации позволил дать физическую интерпретацию источников гравитационного поля: тензор энергии-импульса материи является источником для репера и спиновый момент – источником для лоренцевой связности.

Замечание. При выборе в качестве независимых переменных в геометрии Римана–Картана метрики и кручения, в правой части соответствующих уравнений будут стоять выражения, которые не имеют столь простой и привлекательной интерпретации. □

21.6 Скалярно-тензорные модели

В качестве одного из возможных обобщений эйнштейновской теории гравитации, основанного на римановой геометрии, рассматривают скалярно-тензорные модели, в которых гравитационное взаимодействие описывается метрикой $g_{\alpha\beta}$ и скалярным полем ϕ. Лагранжиан скалярно-тензорных моделей обычно записывают в виде

$$L = \kappa\sqrt{|g|}\phi R + \sqrt{|g|}\frac{\omega(\phi)}{\phi}\partial\phi^2 - \sqrt{|g|}V(\phi) + L_{\text{м}}, \tag{21.32}$$

где $R(g)$ – псевдориманова кривизна (в настоящем разделе знаки тильды опущены),

$$\partial\phi^2 := g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi$$

– кинетический член для скалярного поля, κ – гравитационная постоянная, $\omega(\phi)$ и $V(\phi)$ – некоторые дифференцируемые функции от скалярного поля, характеризующие неминимальность взаимодействия и потенциал самодействия. Все остальные поля включены в лагранжиан полей материи $L_{\text{м}}$.

С точки зрения общей теории относительности можно сказать следующее. Лагранжиан (21.32) описывает гравитационное взаимодействие полей материи, причем гравитационная "постоянная" $\kappa\phi(x)$ зависит от точки пространства-времени, то есть является скалярным полем, которое удовлетворяет своим уравнениям движения. При этом связь этого скалярного поля с метрикой неминимальна и включается некоторое самодействие, описываемое потенциалом V.

При $n = 2$ лагранжиан вида (21.32) описывает двумерную дилатонную гравитацию общего вида, а скалярное поле ϕ называется полем дилатона. Этот класс моделей характеризуется двумя произвольными функциями $U(\phi) = 2\omega(\phi)/\phi$ и $V(\phi)$ и приводит к интегрируемым уравнениям движения (см. раздел **??**).

Впервые действие вида (21.32) было рассмотрено М. Фирцем в 1956 году [**?**]. Скалярно-тензорные модели гравитации привлекли значительное внимание после работ П. Йордана [**?**] и С. Бранса и Р. Дике [**?**]. Основная идея этих исследований восходит к работе П. Дирака, который предположил, что гравитационная постоянная может меняться со временем. ??? М. Фирц пошел дальше, выдвинув гипотезу о том, что гравитационная постоянная описывается независимым скалярным полем, удовлетворяющим некоторому нелинейному уравнению движения.

Получим уравнения движения для скалярно-тензорных моделей (21.32). Используя вид вариационной производной (21.20), нетрудно получить вариационные производные действия (21.32) по обратной метрике и скалярному полю

$$\frac{1}{\sqrt{|g|}}\frac{\delta S}{\delta g^{\alpha\beta}}: \qquad \kappa\phi\left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\right) - \kappa(\Box\phi g_{\alpha\beta} - \nabla_\alpha\nabla_\beta\phi)$$
$$+ \frac{\omega}{\phi}\left(\partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}g_{\alpha\beta}\partial\phi^2\right) + \frac{1}{2}g_{\alpha\beta}V + \frac{1}{2}T_{\text{м}\alpha\beta} = 0. \tag{21.33}$$

$$\frac{1}{\sqrt{|g|}}\frac{\delta S}{\delta\phi}: \qquad \kappa R - 2\frac{\omega}{\phi}\Box\phi - \frac{\omega'}{\phi}\partial\phi^2 + \frac{\omega}{\phi^2}\partial\phi^2 - V' = 0, \tag{21.34}$$

где $T_{\text{м}\alpha\beta}$ – тензор энергии-импульса материи (21.10), а штрих обозначает дифференцирование функции по аргументу. След уравнения (21.33) имеет вид

$$\phi(\tfrac{n}{2} - 1)\left(\kappa R + \frac{\omega}{\phi^2}\partial\phi^2\right) + \kappa(n-1)\Box\phi - \frac{n}{2}V - T_{\text{м}} = 0,$$

где $T_{\text{м}} := T_{\text{м}\alpha}{}^{\alpha}$ – след тензора энергии-импульса материи. Это уравнение при $n \neq 2$ позволяет исключить сумму

$$\kappa R + \frac{\omega}{\phi^2}\partial\phi^2$$

из уравнения для скалярного поля (21.34). В результате получим эквивалентное уравнение

$$2\left(\frac{n-1}{n-2}\kappa + \omega\right)\Box\phi + \frac{1}{n-2}(nV + 2T_{\text{м}}) - \omega'\partial\phi^2 - \phi V' = 0. \tag{21.35}$$

Таким образом систему уравнений движения (21.33), (21.34) можно переписать в эквивалентном виде, заменив уравнение для скалярного поля (21.34) на уравнение (21.35). В четырехмерном пространстве-времени уравнения движения обычно записывают в виде

$$\begin{aligned}\kappa\phi\left(R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R\right) =& -\frac{1}{2}T_{\text{м}\alpha\beta} + \kappa(\Box\phi g_{\alpha\beta} - \nabla_\alpha\nabla_\beta\phi) \\ &- \frac{\omega}{\phi}\left(\partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}g_{\alpha\beta}\partial\phi^2\right) - \frac{1}{2}g_{\alpha\beta}V, \qquad (21.36)\\ \Box\phi =& \frac{1}{2\omega + 3\kappa}(2V + T_{\text{м}} - \omega'\partial\phi^2 - \phi V'). \qquad (21.37)\end{aligned}$$

С математической точки зрения уравнения движения скалярно-тензорных моделей значительно сложнее уравнений общей теории относительности и изучены недостаточно полно. Поэтому практически ничего нельзя сказать о виде функций $\omega(\phi)$ и $V(\phi)$, которые приводят к удовлетворительным результатам с теоретической точки зрения и не противоречат существующим экспериментальным данным.

21.7 Полиномиальная форма действия Гильберта–Эйнштейна

Рассмотрим пространство-время $\mathbb{M}$ произвольной размерности $\dim\mathbb{M} = n \geq 3$. Обозначим локальные координаты пространства-времени через x^α, $\alpha = 0, 1, \ldots, n-1$. Уравнения движения для метрики $g_{\alpha\beta}(x)$ в общей теории относительности без полей материи следуют из вариационного принципа для действия Гильберта–Эйнштейна (21.6).

Действие Гильберта–Эйнштейна неполиномиально по компонентам метрики по двум причинам. Во-первых, оно содержит неполиномиальный элемент объема $\sqrt{|g|}$. Во-вторых, выражение для скалярной кривизны содержит обратную метрику $g^{\alpha\beta}$, компоненты которой неполиномиальны по $g_{\alpha\beta}$. Поэтому действие Гильберта–Эйнштейна в теории возмущений представляет собой очень сложный бесконечный ряд, что является основной технической трудностью при анализе уравнений движения и квантовании. По этим же причинам действие Гильберта–Эйнштейна неполиномиально по компонентам обратной метрики.

Покажем, что конфигурационное пространство общей теории относительности можно расширить, включив определитель метрики в качестве дополнительной независимой переменной таким образом, что действие примет полиномиальный вид. Эквивалентность полиномиального действия исходному действию достигается за счет наложения в расширенном конфигурационном пространстве связи, которая также полиномиальна по полям. Изложение настоящего раздела следует [?, ?].

Координатами конфигурационного пространства $\mathcal{M}$ общей теории относительности являются компоненты метрики $g_{\alpha\beta}(x)$. Размерность этого пространства равна

$$\mathsf{dim}\,\mathcal{M} = \frac{n(n+1)}{2} \times \infty^{(n-1)},$$

где символический множитель $\infty^{(n-1)}$ соответствует точкам пространства в пространстве-времени $\mathbb{M}$.

Рассмотрим другое конфигурационное пространство $\mathcal{N}$ с координатами $\varrho(x), k_{\alpha\beta}(x)$. При этом будем считать, что $\varrho > 0$ и матрица $k_{\alpha\beta}$ симметрична и невырождена в каждой точке пространства-времени

$$k_{\alpha\beta} = k_{\beta\alpha}, \qquad \mathsf{det}\, k_{\alpha\beta} \neq 0.$$

Мы предполагаем также, что матрицы $g_{\alpha\beta}$ и $k_{\alpha\beta}$ имеют одинаковую лоренцевы сигнатуру

$$\mathsf{sign}\, g_{\alpha\beta} = \mathsf{sign}\, k_{\alpha\beta} = (+\underbrace{-\ldots-}_{n-1}).$$

Размерность нового конфигурационного пространства равна

$$\mathsf{dim}\,\mathcal{N} = \left(\frac{n(n+1)}{2} + 1\right) \times \infty^{(n-1)}.$$

Выделим в $\mathcal{N}$ подпространство $\mathcal{M}'$ с помощью дополнительного условия

$$\mathsf{det}\, k_{\alpha\beta} = \begin{cases} 1, & \mathsf{det}\, g_{\alpha\beta} > 0, \quad \text{нечетные } n, \\ -1, & \mathsf{det}\, g_{\alpha\beta} < 0, \quad \text{четные } n. \end{cases} \tag{21.38}$$

Тогда между точками подпространства $\mathcal{M}' \subset \mathcal{N}$ и исходного конфигурационного пространства $\mathcal{M}$ можно установить взаимно однозначное соответствие

$$g_{\alpha\beta} = \varrho^2 k_{\alpha\beta}. \tag{21.39}$$

Обратное преобразование имеет вид

$$\varrho = |\,\mathsf{det}\, g_{\alpha\beta}|^{1/2n}, \qquad k_{\alpha\beta} = \varrho^{-2} g_{\alpha\beta}. \tag{21.40}$$

Поэтому мы отождествим $\mathcal{M} = \mathcal{M}'$. Представление для обратной метрики $g^{\alpha\beta} = \varrho^{-2} k^{\alpha\beta}$, где $k^{\alpha\beta} k_{\beta\gamma} = \delta^\alpha_\gamma$ следует из уравнения (21.39).

Сделаем два важных замечания. Во-первых, компоненты обратной матрицы $k^{\alpha\beta}$ полиномиальны по $k_{\alpha\beta}$. В общем случае из условия (21.38) следует, что они являются полиномами степени $n-1$ по компонентам $k_{\alpha\beta}$. Во-вторых, компоненты $k_{\alpha\beta}$ являются компонентами не тензора, а тензорной плотности второго ранга. Действительно, потребуем, чтобы равенство (21.39) было выполнено в произвольной системе координат. Поскольку определитель метрики g представляет собой скалярную плотность степени $\mathsf{deg}\, g = -2$, то матрица $k_{\alpha\beta}$ является симметричной тензорной плотностью второго ранга и степени $\mathsf{deg}\, k_{\alpha\beta} = 2/n$, а поле ϱ – скалярной плотностью степени $\mathsf{deg}\, \varrho = -1/n$. То есть при преобразовании координат $x^\alpha \to x^{\alpha'}(x)$ новые переменные преобразуется по-правилу

$$k_{\alpha'\beta'} = \partial_{\alpha'} x^\alpha \partial_{\beta'} x^\beta k_{\alpha\beta} J^{2/n}, \qquad \varrho' = \varrho J^{-1/n},$$

где $J = \det \partial_\alpha x^{\alpha'}$ – якобиан преобразования координат. Это значит, что дополнительное условие (21.38) инвариантно относительно преобразования координат. В связи с этим будем называть поле $k_{\alpha\beta}(x)$ *плотностью метрики.*

Явная формула для компонент обратной плотности метрики имеет вид

$$k^{\alpha\beta} = \frac{1}{(n-1)!}\hat{\varepsilon}^{\alpha\gamma_2\ldots\gamma_n}\hat{\varepsilon}^{\beta\delta_2\ldots\delta_n}k_{\gamma_2\delta_2}\ldots k_{\gamma_n\delta_n},$$

где $\hat{\varepsilon}^{\alpha_1\ldots\alpha_n}$ – полностью антисимметричная тензорная плотность степени -1 с единичными компонентами $|\hat{\varepsilon}^{\alpha_1\ldots\alpha_n}| = 1$. Данная формула показывает, что компоненты $k^{\alpha\beta}$ – это полиномы степени $n-1$ от компонент $k_{\alpha\beta}$ и наоборот.

Перепишем действие Гильберта–Эйнштейна в новых переменных $\varrho, k_{\alpha\beta}$. Для вычислений нам понадобятся следующие тождества, которые следуют после дифференцирования (21.38)

$$\begin{aligned} k^{\gamma\delta}\partial_\alpha k_{\gamma\delta} &= 0,\\ k^{\gamma\delta}\partial^2_{\alpha\beta}k_{\gamma\delta} - k^{\gamma\delta}k^{\epsilon\zeta}\partial_\alpha k_{\gamma\epsilon}\partial_\beta k_{\delta\zeta} &= 0. \end{aligned}$$

Прямые вычисления приводят к равенствам

$$\begin{aligned} \partial_\alpha g_{\beta\gamma} &= \varrho^2\left(\partial_\alpha k_{\beta\gamma} + 2\partial_\alpha \ln\varrho\, k_{\beta\gamma}\right),\\ \partial^2_{\alpha\beta}g_{\gamma\delta} &= \varrho^2\Big(\partial^2_{\alpha\beta}k_{\gamma\delta} + 2\partial^2_{\alpha\beta}\ln\varrho\, k_{\gamma\delta} +\\ &\qquad +2\partial_\alpha\ln\varrho\,\partial_\beta k_{\gamma\delta} + 2\partial_\beta\ln\varrho\,\partial_\alpha k_{\gamma\delta} + 4\partial_\alpha\ln\varrho\,\partial_\beta\ln\varrho\, k_{\gamma\delta}\Big),\\ g^{\beta\delta}\partial_\alpha g_{\delta\gamma} &= k^{\beta\delta}\partial_\alpha k_{\delta\gamma} + 2\partial_\alpha\ln\varrho\,\delta^\beta_\gamma,\\ g^{\epsilon\zeta}\partial_\alpha g_{\gamma\epsilon}\partial_\beta g_{\delta\zeta} &= \varrho^2\Big(k^{\epsilon\zeta}\partial_\alpha k_{\beta\epsilon}\partial_\beta k_{\delta\zeta} + 2\partial_\alpha\ln\varrho\,\partial_\beta k_{\gamma\delta} + 2\partial_\beta\ln\varrho\,\partial_\alpha k_{\gamma\delta} +\\ &\qquad +4\partial_\alpha\ln\varrho\,\partial_\beta\ln\varrho\, k_{\gamma\delta}\Big), \qquad (21.41)\\ g^{\gamma\delta}g^{\epsilon\zeta}\partial_\alpha g_{\gamma\epsilon}\partial_\beta g_{\delta\zeta} &= k^{\gamma\delta}k^{\epsilon\zeta}\partial_\alpha k_{\gamma\epsilon}\partial_\beta k_{\delta\zeta} + 4n\partial_\alpha\ln\varrho\,\partial_\beta\ln\varrho,\\ g^{\gamma\delta}\partial^2_{\alpha\beta}g_{\gamma\delta} &= k^{\gamma\delta}\partial^2_{\alpha\beta}k_{\gamma\delta} + 2n\partial^2_{\alpha\beta}\ln\varrho + 4n\partial_\alpha\ln\varrho\,\partial_\beta\ln\varrho,\\ \partial^2_{\alpha\beta}g_{\gamma\delta} - \frac{1}{2}g^{\epsilon\zeta}\partial_\alpha g_{\gamma\epsilon}\partial_\beta g_{\delta\zeta} - \frac{1}{2}g^{\epsilon\zeta}\partial_\beta g_{\gamma\epsilon}\partial_\alpha g_{\delta\zeta} &=\\ = \varrho^2\left(\partial^2_{\alpha\beta}k_{\gamma\delta} - \frac{1}{2}k^{\epsilon\zeta}\partial_\alpha k_{\gamma\epsilon}\partial_\beta k_{\delta\zeta} - \frac{1}{2}k^{\epsilon\zeta}\partial_\beta k_{\gamma\epsilon}\partial_\alpha k_{\delta\zeta} + 2\partial^2_{\alpha\beta}\ln\varrho\, k_{\gamma\delta}\right). \end{aligned}$$

Теперь нетрудно найти выражение для скалярной кривизны

$$R = \varrho^{-4}\left[\varrho^2 R^{(k)} + 2(n-1)\varrho\partial_\alpha(k^{\alpha\beta}\partial_\beta\varrho) + (n-1)(n-4)\partial\varrho^2\right], \qquad (21.42)$$

где

$$\partial\varrho^2 := k^{\alpha\beta}\partial_\alpha\varrho\partial_\beta\varrho,$$

и "скалярная" кривизна $R^{(k)}$ для плотности метрики $k_{\alpha\beta}$ принимает удивительно простой вид

$$R^{(k)} = \partial_{\alpha\beta}k^{\alpha\beta} + \frac{1}{2}k^{\alpha\beta}\partial_\alpha k^{\gamma\delta}\partial_\gamma k_{\beta\delta} - \frac{1}{4}k^{\alpha\beta}\partial_\alpha k^{\gamma\delta}\partial_\beta k_{\gamma\delta}. \qquad (21.43)$$

Отметим, что это выражение полиномиально и по плотности метрики $k_{\alpha\beta}$, и по ее обратной $k^{\alpha\beta}$.

Заметим, что для заданной плотности метрики $k_{\alpha\beta}$ мы можем формально построить символы Кристоффеля, тензоры кривизны и Риччи, а также скалярную кривизну. Соответствующие "символы Кристоффеля" не определяют связность на

$\mathbb{M}$, а "кривизна" не является тензором, потому что новая переменная $k_{\alpha\beta}$ является не тензором, а тензорной плотностью. Например, скаляром в выражении (21.42) является не просто $R^{(k)}$, а вся правая часть выражения. Вместе с этим, группа общих преобразований координат содержит подгруппу, состоящую из преобразований с единичным якобианом. Относительно этой подгруппы "символы Кристоффеля" $\Gamma^{(k)\gamma}_{\alpha\beta}$ для $k_{\alpha\beta}$ преобразуются, как компоненты связности, а "кривизна" $R^{(k)\ \delta}_{\alpha\beta\gamma}$ является тензором.

Вторые производные $\partial^2_{\alpha\beta}k^{\alpha\beta}$ и $\partial_\alpha(k^{\alpha\beta}\partial_\beta\varrho)$, как и в общей теории относительности, можно исключить из действия Гильберта–Эйнштейна (21.6), вычтя граничный член

$$\partial_\alpha\Big(\varrho^{n-2}\partial_\beta k^{\alpha\beta} + 2(n-1)\varrho^{n-3}k^{\alpha\beta}\partial_\beta\varrho\Big).$$

Здесь, для краткости, мы положили $\kappa = 0$. В результате действие примет вид

$$S_{\text{не}} \overset{\mathsf{div}}{=} \int d^n x L_{\text{не}}, \tag{21.44}$$

где

$$\begin{aligned} L_{\text{не}} = \varrho^{n-4}\Big[&\frac{1}{2}\varrho^2 k^{\alpha\beta}\partial_\alpha k^{\gamma\delta}\partial_\gamma k_{\beta\delta} - \frac{1}{4}k^{\alpha\beta}\partial_\alpha k^{\gamma\delta}\partial_\beta k_{\gamma\delta} - \\ &-(n-2)\varrho\partial_\alpha k^{\alpha\beta}\partial_\beta\varrho - (n-1)(n-2)\partial\varrho^2 - (n-2)\Lambda\varrho^4\ \Big]. \end{aligned} \tag{21.45}$$

При размерности пространства-времени $n \geq 4$ этот лагранжиан полиномиален по полям $\varrho, k_{\alpha\beta}$ со степенями n и $n+1$ соответственно. Общая степень полинома равна $2n-1$. По-построению, с точностью до граничных слагаемых это действие инвариантно относительно общих преобразований координат. Необходимо только помнить, что поля ϱ и $k_{\alpha\beta}$ являются не тензорами, а тензорными плотностями.

Лагранжиан (21.45) имеет вид лагранжиана дилатонной гравитации, где роль дилатона играет определитель метрики. От обычных моделей он отличается тем, что содержит перекрестный член с производными $\partial_\alpha k^{\alpha\beta}\partial_\beta\varrho$. Кроме того, он содержит меньшее число независимых полей, так как на плотность метрики наложено условие (21.38).

Дополнительное условие на плотность метрики (21.38) можно учесть, добавив к лагранжиану связь,

$$L_{\text{не}} \mapsto L_{\text{не}} + \lambda(\det k_{\alpha\beta} \pm 1),$$

где λ – множитель Лагранжа.

Введение множителей Лагранжа необязательно. Из условия (21.38) следует условие на вариации плотности метрики

$$\delta\det k_{\alpha\beta} = \pm k^{\alpha\beta}\delta k_{\alpha\beta} = \mp\delta k^{\alpha\beta}k_{\alpha\beta} = 0. \tag{21.46}$$

Поэтому действие (21.44) можно варьировать по $k_{\alpha\beta}$ или $k^{\alpha\beta}$, рассматривая все компоненты как независимые, а затем взять бесследовую часть получившихся уравнений.

Уравнения Эйлера–Лагранжа для действия (21.44) эквивалентны вакуумным уравнениям Эйнштейна (21.4). Вариация действия по ϱ приводит к уравнению

$$2(n-1)\varrho\Box\varrho + (n-1)(n-4)\partial\varrho^2 + R^{(k)}\varrho^2 = n\Lambda\varrho^4, \tag{21.47}$$

которое эквивалентно следу уравнений Эйнштейна (21.4). Вариация действия (21.44) по плотности метрики дает $n(n+1)/2-1$ уравнение

$$\varrho^2\left(R^{(k)}_{\alpha\beta}-\frac{1}{n}k_{\alpha\beta}R^{(k)}\right)+ \\ +(n-2)(\varrho\nabla_\alpha\nabla_\beta\varrho-\nabla_\alpha\varrho\nabla_\beta\varrho)-\frac{n-2}{n}k_{\alpha\beta}(\varrho\Box\varrho-\partial\varrho^2)=0, \quad (21.48)$$

где

$$\begin{aligned}\nabla_\alpha\varrho&=\partial_\alpha\varrho,\\ \nabla_\alpha\nabla_\beta\varrho&=\partial^2_{\alpha\beta}\varrho-\Gamma^{(k)\gamma}_{\alpha\beta}\partial_\gamma\varrho,\\ \Box\varrho&=\partial_\alpha(k^{\alpha\beta}\partial_\beta\varrho),\end{aligned} \quad (21.49)$$

и символы Кристоффеля $\Gamma^{(k)\gamma}_{\alpha\beta}$ построены по плотности метрики $k_{\alpha\beta}$. Эти уравнения эквивалентны бесследовой части уравнений Эйнштейна, потому что при варьировании необходимо учесть условие на вариации компонент плотности метрики (21.46). Заметим, что "ковариантные" производные тензоров и тензорных плотностей при условии (21.38) совпадают, так как $\Gamma^{(k)\beta}_{\alpha\beta}=0$. "Тензор Риччи" в (21.48) имеет вид

$$R^{(k)}_{\alpha\beta}=\frac{1}{2}k^{\gamma\delta}(\partial^2_{\gamma\delta}k_{\alpha\beta}-\partial^2_{\alpha\gamma}k_{\beta\delta}-\partial^2_{\beta\gamma}k_{\alpha\delta})-\frac{1}{2}\partial_\gamma k^{\gamma\delta}(\partial_\alpha k_{\beta\delta}+\partial_\beta k_{\alpha\delta}-\partial_\delta k_{\alpha\beta}) \\ -\frac{1}{4}\partial_\alpha k_{\gamma\delta}\partial_\beta k^{\gamma\delta}-\frac{1}{2}k^{\gamma\delta}k^{\epsilon\zeta}(\partial_\gamma k_{\alpha\epsilon}\partial_\delta k_{\beta\zeta}-\partial_\gamma k_{\alpha\epsilon}\partial_\zeta k_{\beta\delta}).$$

Уравнения (21.47) и (21.48) представляют собой соответственно след и бесследовую часть уравнений Эйнштейна (21.4). Для любого решения уравнений (21.47) и (21.48) существует единственная метрика (21.39), которая удовлетворяет уравнениям Эйнштейна (21.4). Наоборот, для любого решения уравнений Эйнштейна (21.4) можно построить единственные плотность метрики и поле ϱ (21.40), которые будут удовлетворять уравнениям (21.47) и (21.48).

Таким образом, при $n\geq 4$, действие (21.44), уравнения Эйлера–Лагранжа (21.47), (21.48) и связь (21.38) полиномиальны по полям $\varrho, k_{\alpha\beta}$. Это существенное упрощение теории достигнуто за счет расширения конфигурационного пространства путем введения дополнительной полевой переменной ϱ. Если связь (21.38) явно решить относительно одной из компонент плотности метрики $k_{\alpha\beta}$ и подставить найденное решение в действие (21.44), то полиномиальность теории будет утеряна. Отметим, что введение нового поля и связи не является чем-то, из ряда вон выходящим: исходная метрика $g_{\alpha\beta}$ и так содержит нефизические степени свободы, которые исключаются из теории путем решения калибровочных условий и связей, содержащихся в общей теории относительности. Мы лишь увеличили число переменных и связей, оставив физические степени свободы прежними.

На гамильтоновом языке проведенная процедура означает следующее. Фазовое пространство, соответствующее переменным $\varrho, k_{\alpha\beta}$, также расширится, при этом возникнет дополнительная связь на импульсы (равенство нулю следа импульсов, соответствующих $k_{\alpha\beta}$), которая вместе со связью (21.38) образует пару связей второго рода. Однако полное фазовое пространство уже будет не симплектическим, а только пуассоновым, так как матрица скобок Пуассона координат этого пространства будет вырождена. Подробнее этот вопрос рассмотрен в разделе 22.7.

Добавление скалярного $\varphi(x)$ и электромагнитного $A_\alpha(x)$ полей сохраняет действие и уравнения Эйлера–Лагранжа полиномиальными. При минимальной подстановке получаем следующие лагранжианы:

$$L_{\text{s}} = \frac{1}{2}\sqrt{g}[g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi)] = \frac{1}{2}\varrho^{n-2}[k^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - \varrho^2 V(\varphi)],$$
$$L_{\text{ем}} = -\frac{1}{4}\sqrt{g}g^{\alpha\beta}g^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} = -\frac{1}{4}\varrho^{n-4}k^{\alpha\beta}k^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta},$$

где $V(\varphi)$ – потенциал скалярного поля, включая массовый член, и $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$ – напряженность электромагнитного поля.

В четырехмерном пространстве-времени уравнение (21.47) упрощается

$$\Box\varrho + \frac{1}{6}R^{(k)}\varrho = \frac{2}{3}\Lambda\varrho^3. \tag{21.50}$$

При этом мы разделили его на ϱ, что возможно в силу предположения $\varrho > 0$.

Отбросим на время условие на плотность метрики (21.50) и будем считать $k_{\alpha\beta}$ метрикой, а ϱ – скалярным полем. Тогда уравнение (21.50) ковариантно относительно конформных преобразований

$$\bar{k}_{\alpha\beta} = \Omega^2 k_{\alpha\beta}, \qquad \bar{\varrho} = \Omega^{-1}\varrho, \tag{21.51}$$

где $\Omega(x) > 0$ – дважды дифференцируемая функция. Оно рассматривалось в [?, ?, ?] при $\Lambda = 0$ и в [?] при $\Lambda \neq 0$. Рассматриваемый подход существенно отличается, поскольку на плотность метрики $k_{\alpha\beta}$ наложено дополнительно условие (21.38), которое явно нарушает конформную инвариантность. Однако появление множителя $1/6$ неслучайно, так как параметризация (21.39) по виду совпадает с конформным преобразованием метрики (21.51).

Модели гравитации для метрики с единичным определителем рассматривались в физике неоднократно. В общековариантных теориях, которой является общая теория относительности, существует произвол в выборе системы координат, которым можно воспользоваться. Нетрудно доказать, что в окрестности произвольной точки пространства-времени существует такая система координат, что определитель метрики равен по модулю единице $|\det g_{\alpha\beta}| = 1$. Такие системы координат определены с точностью до преобразований координат с единичным якобианом. Условие $|\det g_{\alpha\beta}| = 1$ существенно упрощает многие формулы, в частности, выражение для скалярной кривизны. Этим обстоятельством пользовался Эйнштейн при создании общей теории относительности [?].

21.8 Точечные частицы в теории гравитации

Пусть задано пространство-время, т.е. многообразие $\mathbb{M}$ с метрикой g лоренцевой сигнатуры. Будем считать, что кручение и неметричность на $\mathbb{M}$ равны нулю, и, для простоты, не будем помечать это обстоятельство знаком тильды. Размерность пространства-времени пока не фиксируем, $\dim\mathbb{M} = n \geq 2$. Точечная частица движется в пространстве-времени $\mathbb{M}$ по некоторой дифференцируемой времениподобной кривой $\{q^\alpha(\tau)\} \in \mathbb{M}$, где $\tau \in \mathbb{R}$ – произвольный параметр вдоль этой кривой. Напомним, что кривая называется времениподобной, если вектор скорости кривой, $u^\alpha = \dot{q}^\alpha = dq^\alpha/d\tau$, времениподобен, т.е. $u^2 = u^\alpha u^\beta g_{\alpha\beta} > 0$. Мы считаем, что частица

движется в будущее, т.е. $u^0 > 0$. Форма кривой определяется рассматриваемой задачей и силами, которые действуют на частицу. В общем случае параметр вдоль кривой произволен, и его выбирают из соображений удобства. Наиболее часто в качестве параметра вдоль траектории частицы выбирают ее длину s. Это всегда возможно, т.к. обыкновенное дифференциальное уравнение $ds = \sqrt{\dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}}\, d\tau$ разрешимо. В дальнейшем точка будет обозначать дифференцирование по параметру s.

Определение. Времениподобная дифференцируемая кривая $\{q^\alpha(\tau)\} \in \mathbb{M}$, вдоль которой движется точечная частица, называется *траекторией* или *мировой линией* частицы. Если параметр вдоль траектории частицы совпадает с ее длиной, $\tau = s$, то он называется *собственным временем*. Векторное поле

$$u^\alpha = \frac{dq^\alpha}{ds}, \tag{21.52}$$

определенное на траектории частицы, называется *собственной скоростью* частицы. Ковариантная производная от скорости частицы вдоль ее траектории

$$w^\alpha := \frac{dq^\beta}{ds}\nabla_\beta u^\alpha = u^\beta \nabla_\beta u^\alpha \tag{21.53}$$

называется *ускорением* частицы. Частица называется *свободной*, если ее ускорение равно нулю. Производная вдоль траектории частицы

$$v^\alpha = \frac{dq^\alpha}{dq^0} = \frac{\dot{q}^\alpha}{\dot{q}^0} \tag{21.54}$$

называется *наблюдаемой скоростью* частицы в системе координат x^α. □

Собственная скорость и ускорение частицы являются n-мерными векторами, определенным вдоль траектории частицы. Собственное время – это то время, которое показывают часы наблюдателя, движущегося вместе с частицей. Когда наблюдатель движется вместе с частицей, то он может измерить свою скорость относительно системы координат x^α, это и будут компоненты собственной скорости. Наблюдаемая скорость, как следует из определения, не является векторным полем. Она привязана к выбранной системе отсчета.

Равенство нулю ускорения частицы

$$u^\beta \nabla_\beta u^\alpha = 0,$$

определяет экстремали (17.18). Это значит, что свободные частицы в теории тяготения движутся вдоль экстремалей пространства времени. Если на частицу действуют негравитационные силы, например, электромагнитные, то в уравнении движения $mw^\alpha = f^\alpha$, где $m = \mathsf{const}$ – масса частицы, появится внешняя сила с компонентами f^α. В этом случае ее траектория будет отличаться от экстремали.

Предложение 21.8.1. *Для любой точечной частицы квадрат собственной скорости равен единице,*

$$u^2 := u^\alpha u^\beta g_{\alpha\beta} = 1. \tag{21.55}$$

При этом ускорение всегда ортогонально скорости

$$u^\alpha w_\alpha = 0.$$

Доказательство. Первое утверждение следует из определения:

$$u^2 = \frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}g_{\alpha\beta} = \frac{ds^2}{ds^2} = 1.$$

Продифференцируем это равенство вдоль траектории

$$u^\beta \nabla_\beta (u^2) = 2u^\beta u^\alpha \nabla_\beta u_\alpha = 2u^\alpha w_\alpha = 0,$$

где мы воспользовались тем, что ковариантная производная от метрики для связности Леви–Чивиты равна нулю. Отсюда вытекает второе утверждение предложения. □

Наблюдаемая скорость является нековариантным объектом. Из определения следует соотношение между компонентами собственной и наблюдаемой скоростями

$$v^\alpha = \frac{u^\alpha}{\dot{q}^0}. \tag{21.56}$$

Более подробно,

$$v^0 = 1, \qquad v^\mu = \frac{dq^\mu}{dq^0}.$$

Возведем равенство $v^\alpha \dot{q}^0 = u^\alpha$ в квадрат и учтем, что $u^2 = 1$. Тогда получим, что компоненты собственной скорости частицы можно записать в виде

$$u^0 := \dot{q}^0 = \frac{1}{\sqrt{v^2}}, \qquad u^\mu := \dot{q}^\mu = \frac{v^\mu}{\sqrt{v^2}}, \tag{21.57}$$

где $v^2 = v^\alpha v^\beta g_{\alpha\beta}$ – квадрат наблюдаемой скорости. Отсюда также следует, что $v^2 > 0$. Понятие наблюдаемой скорости частицы будет использовано в дальнейшем при определении ультрарелятивистского предела для точечной частицы.

Замечание. Определения даны для точечной частицы, находящейся под действием произвольных сил, не только гравитационных. Мы также предполагаем, что, если наблюдатель находится в определенной точке пространства в определенный момент времени, то ему известны значения всех координат x^α, соответствующих данной точке пространства-времени. Мы также предполагаем, что ему известны координаты всех близлежащих точек, что необходимо для вычисления производных. Эти упрощающие предположения сделаны для того, чтобы обойти важный и интересный, но сложный вопрос измерений. □

В общем случае, когда на частицу действуют произвольные силы, она может двигаться по любой времениподобной кривой. В моделях гравитации мы предполагаем, что точечная частица, на которую действуют только гравитационные силы, описывается инвариантным действием

$$S_{\mathrm{m}} = -mc\int_{\mathrm{P}}^{\mathrm{Q}} ds = -mc\int_{\tau_1}^{\tau_2} d\tau \sqrt{\dot{q}^\alpha \dot{q}^\beta g_{\alpha\beta}}, \tag{21.58}$$

где $m = \mathsf{const} > 0$ – масса частицы, c – скорость света и интегрирование проводится вдоль времениподобной кривой $q(\tau)$, $\tau \in \mathbb{R}$, соединяющей точки $\mathrm{P} = q(\tau_1)$ и $\mathrm{Q} = q(\tau_2)$. Действие (21.58) отличается от длины мировой линии частицы (17.16) постоянным множителем $-mc$ и, если метрика задана, варьируется только по траектории частицы $\delta q^\alpha(\tau)$. В дальнейшем положим $c = 1$.

Допустим, что точки $\mathrm{P}, \mathrm{Q} \in \mathbb{M}$ можно соединить времениподобной экстремалью и притом только одной. Ясно, что эти точки всегда можно соединить также ломанными светоподобными кривыми γ. Длина этих кривых равна нулю, потому что каждый сегмент ломанных светоподобен. Если некоторая времениподобная кривая аппроксимирует одну из этих кривых γ, то ее длина будет близка к нулю. Остальные времениподобные кривые, соединяющие точки P и Q будут иметь некоторую положительную длину. Поэтому, если точки P и Q соединяются только одной экстремалью, то это будет времениподобная кривая наибольшей длины. Благодаря знаку минус перед действием (21.58), экстремаль соответствует минимуму функционала (21.58) среди всех времениподобных кривых, соединяющих точки P и Q. Отсюда следует, что точечные частицы движутся вдоль экстремалей пространства-времени, если на них действуют только внешние гравитационные силы.

Если метрика пространства-времени считается заданной, то это означает, что мы не учитываем гравитационное поле, создаваемое самой частицей. Во многих случаях такая постановка задачи вполне допустима, например, если масса частицы мала. Такие частицы называются *пробными*.

Замечание. Мы предполагаем, что точечная частица описывается действием (21.58) независимо от того, задана ли на $\mathbb{M}$ аффинная геометрия общего вида с кручением и неметричностью или задана только метрика. Тогда в общем случае точечная частица будет двигаться по экстремали, а не по геодезической. Если кручение и неметричность пространства-времени равны нулю, то экстремали совпадают с геодезическими. Поэтому в общей теории относительности можно также сказать, что точечные частицы под действием гравитационных сил движутся по геодезическим. Если, помимо гравитационных сил, присутствуют также другие, например, электромагнитные взаимодействия, то вид действия (21.58) изменится. Поэтому при наличии сил негравитационного происхождения траектории частиц, вообще говоря, отличаются от экстремалей. □

Выше было введено понятие собственного времени для произвольной времениподобной линии. Для экстремали собственное время по существу совпадает с каноническим параметром. Действительно, для любой экстремали квадрат вектора скорости сохраняется, $u^2 = \mathsf{const}$, (17.22). Поскольку канонический параметр определен с точностью до умножения на отличную от нуля постоянную, то его всегда можно выбрать таким образом, чтобы квадрат вектора скорости частицы был равен единице (21.55). Другими словами, в качестве канонического параметра выбирается длина экстремали, $\tau = s$.

Предположим, что в пространстве-времени находится N частиц с массами m_i, $i = 1, \ldots, \mathrm{N}$, которые взаимодействуют между собой только посредством гравитационных сил. В общей теории относительности суммарное действие гравитационного поля и совокупности точечных частиц равно сумме действия Гильберта–Эйнштейна (21.6) и действий для каждой частицы,

$$S = S_{\text{не}} + \sum_{i=1}^{\mathrm{N}} S_i = \\ = \kappa \int_{\mathbb{M}} d^n x \sqrt{|g|} R - \sum_i m_i \int_{\tau_{i1}}^{\tau_{i2}} d\tau_i \sqrt{\dot{q}_i^\alpha \dot{q}_i^\beta g_{\alpha\beta}},$$

где мы, для простоты, опустили космологическую постоянную и знак тильды у скалярной кривизны. Во втором слагаемом метрика рассматривается как сложная функ-

ция $g_{\alpha\beta}(\tau_i) = g_{\alpha\beta}\Big(q(\tau_i)\Big)$, и параметры τ_i могут быть выбраны произвольно для каждой частицы. Первый интеграл берется по всему пространству-времени, а последующие – в пределах $\tau_{i1,2}$ (возможно, бесконечных), которые соответствуют пересечению мировых линий частиц с краем $\partial\mathbb{M}$. Для простоты, предположим, что мировые линии частиц нигде не пересекаются, т.е. частицы не сталкиваются между собой.

Замечание. В настоящем разделе нас не будут интересовать граничные эффекты. Поэтому пределы интегрирования, для простоты, мы в дальнейшем опустим. □

Действие (21.59) инвариантно относительно общих преобразований координат и независимой перепараметризации параметров τ_i вдоль каждой траектории. В дальнейшем мы предполагаем, что вдоль каждой траектории параметр совпадает с собственным временем. Более того, поскольку в действие входит сумма интегралов вдоль траекторий, то индекс i у параметров τ_i можно опустить,

$$S = \int d^n x \sqrt{|g|}\kappa R - \int d\tau \sum_i m_i \sqrt{\dot q_i^\alpha \dot q_i^\beta g_{\alpha\beta}}. \tag{21.59}$$

В общем случае пределы интегрирования для различных частиц могут отличаться. Однако, поскольку нас не интересуют граничные эффекты, мы этого указывать не будем.

Для вывода уравнений движения, действие (21.59) необходимо проварьировать по метрике $g_{\alpha\beta}(x)$ и траекториям частиц $q_i^\alpha(\tau)$. Вариация действия для частиц по компонентам метрики $\delta g_{\alpha\beta}(x)$ неопределена, т.к. оно записано только вдоль траекторий. Поэтому мы преобразуем интегралы вдоль траекторий в интегралы по всему пространству-времени. Для этого вставим в подынтегральное выражение единицу,

$$1 = \int d^n x \delta^{(n)}(x - q_i) := \int d^n x \delta(x^0 - q_i^0)\delta(x^1 - q_i^1), \dots \delta(x^{n-1} - q_i^{n-1}),$$

и изменим порядок интегрирования, предположив, что это возможно. Тогда действие примет вид

$$S = \int d^n x \left[\kappa\sqrt{|g|}R - \int d\tau \sum_i m_i \sqrt{\dot q_i^\alpha \dot q_i^\beta g_{\alpha\beta}}\,\delta^{(n)}(x - q_i)\right]. \tag{21.60}$$

Теперь метрику во втором слагаемом можно рассматривать, как функцию от точки пространства-времени, $g_{\alpha\beta} = g_{\alpha\beta}(x)$. Вариационные производные этого действия по траекториям частиц и метрике равны

$$\frac{\delta S}{\delta q_i^\alpha} = m_i \left(\ddot q_i^\beta + \Gamma_{\gamma\delta}{}^\beta \dot q_i^\gamma \dot q_i^\delta\right) g_{\beta\alpha}, \tag{21.61}$$

$$\frac{\delta S}{\delta g_{\alpha\beta}} = -\sqrt{|g|}\kappa\left(R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R\right) - \int d\tau \sum_i \frac{m_i}{2\sqrt{\dot q_i^\gamma \dot q_i^\delta g_{\gamma\delta}}} \dot q_i^\alpha \dot q_i^\beta \delta^{(n)}(x - q_i). \tag{21.62}$$

Вариационная производная по траекториям частиц (21.60) была подробно рассмотрена при изучении экстремалей в разделе (17.2). Вариационная производная действия Гильберта–Эйнштейна также была получена ранее в разделе (21.3). Для любого решения уравнения (21.61) параметр τ можно выбрать так, что $\sqrt{\dot q_i^\gamma \dot q_i^\delta g_{\gamma\delta}} = 1$ для каждой частицы. Поэтому, не ограничивая общности, знаменатель во втором слагаемом (21.62) можно упростить, отбросив квадратный корень. Таким образом,

связанная система уравнений движения гравитационного поля и системы точечных частиц примет вид

$$\kappa\left(R_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}R\right)=-\frac{1}{2}T_{\mathrm{m}\alpha\beta}, \tag{21.63}$$

$$\ddot{q}_i^{\alpha}=-\Gamma_{\beta\gamma}{}^{\alpha}\dot{q}_i^{\beta}\dot{q}_i^{\gamma}, \tag{21.64}$$

где

$$T_{\mathrm{m}\alpha\beta}=\frac{1}{\sqrt{|g|}}\int d\tau\sum_i m_i\dot{q}_{i\alpha}\dot{q}_{i\beta}\delta^{(n)}(x-q_i) \tag{21.65}$$

– тензор энергии-импульса точечных частиц и $\dot{q}_{i\alpha}=\dot{q}_i^{\beta}g_{\beta\alpha}$. Интегрирование по каноническим параметру τ в тензоре энергии-импульса можно снять, использовав одну δ-функцию, а именно $\delta\Big(x^0-q_i^0(\tau)\Big)$. Поскольку $\dot{q}_i^0>0$ (все частицы движутся в будущее), то для тензора энергии-импульса точечных частиц получаем следующее выражение

$$T_{\mathrm{m}}{}^{\alpha\beta}=\frac{1}{\sqrt{|g|}}\sum_i\frac{m_i\dot{q}_i^{\alpha}\dot{q}_i{}^{\beta}}{\dot{q}_i^0}\delta^{(n-1)}(\boldsymbol{x}-\boldsymbol{q}_i), \tag{21.66}$$

где

$$\delta^{(n-1)}(\boldsymbol{x}-\boldsymbol{q}_i):=\delta(x^1-q_i^1)\dots\delta(x^{n-1}-q_i^{n-1})$$

и параметр τ является неявной функцией x^0, заданной уравнением $x^0=q^0(\tau)$.

Замечание. Появление множителя $1/\sqrt{|g|}$ в выражении для тензора энергии-импульса точечных частиц неслучайно. Напомним, что δ-функция является не функцией на многообразии, а скалярной плотностью степени -1, как и элемент объема $\sqrt{|g|}$. □

Таким образом, для точечных частиц, на которые действуют только гравитационные силы, мы имеем связанную систему уравнений (21.63), (21.64). Каждая частица движется по экстремали пространства-времени в соответствии с уравнением (21.64), где метрика определяется уравнениями Эйнштейна (21.63). В свою очередь, метрика зависит от распределения частиц, так как в правой части уравнений Эйнштейна стоит нетривиальный тензор энергии-импульса.

В заключение раздела обсудим с общей точки зрения те трудности, которые возникают при решении уравнений (21.63) и (21.64) в связи с наличием δ-функций.

Уравнение для экстремалей (траекторий точечных частиц) (21.65) хорошо определено, если компоненты метрики – дифференцируемые функции, $g_{\alpha\beta}\in\mathcal{C}^1(\mathbb{U})$, $\mathbb{U}\subset\mathbb{M}$. Однако это условие может не выполняться, т.к. компоненты метрики могут иметь особенности на мировых линиях частиц. В настоящее время общего подхода к этой проблеме не существует. Ниже мы обсудим возможный класс метрик, для которых уравнениям можно придать смысл.

Для уравнений Эйнштейна (21.63) также возникают серьезные проблемы, т.к. в правой части стоят δ-функции. Если компоненты метрики – дважды дифференцируемые функции, то тензор Эйнштейна будет непрерывен и не может быть равен δ-функции. Если компоненты метрики разрывны, то в общем случае символы Кристоффеля будут содержать δ-функции, а тензор кривизны – квадраты δ-функций. Поскольку в классе обобщенных функций нельзя определить ассоциативное умножение [], то для разрывных компонент метрики тензор Эйнштейна в общем случае неопределен. Поэтому решение уравнений Эйнштейна (21.63) для точечных частиц ищется в классе метрик, имеющих непрерывные кусочно дифференцируемые компоненты. Тогда компоненты обратной метрики будут также непрерывны, а символы

Кристоффеля будут иметь разрывы. Допустим, что мы определили значение производных в точках разрыва каким либо образом. Тогда разрывные функции образуют алгебру с поточечным умножением. В этой алгебре нельзя ввести дифференцирование, однако оно нам и не нужно, т.к. тензор кривизны не содержит дифференцирования произведения символов Кристоффеля. Соответствующий тензор кривизны содержит δ-функции. Их нельзя перемножать, однако это также не требуется, так как вторые производные от компонент метрики входят в тензор кривизны линейно.

Мы предполагаем, что производные от компонент метрики имеют разрывы только вдоль направлений, перпендикулярных траекториям частиц. Производная компонент метрики вдоль траекторий предполагается непрерывной. Этого достаточно для получения необходимых δ-функций в тензоре Эйнштейна. При этом уравнения для экстремалей (21.64) также имеют смысл. Действительно, символы Кристоффеля $\Gamma_{\beta\gamma\alpha}$ содержат три слагаемых с производными от компонент метрики по координатам x^β, x^γ и x^α. Поскольку индексы β и γ в уравнении (21.64) свернуты с компонентами скорости $\dot{q}$, то соответствующие производные от компонент метрики берутся вдоль траекторий частиц. Следовательно, они непрерывны по предположению. Остается открытым вопрос о непрерывности производных, соответствующих свободному индексу α. Допустим, что существует дважды дифференцируемое классическое решение $q^\alpha(\tau)$ уравнения (21.64). Тогда проекция уравнения (21.64) на любое направление, перпендикулярное траектории, равно нулю. В самом деле, обозначим через n произвольный вектор в какой либо точке кривой, который перпендикулярен вектору скорости, т.е. $(\dot{q}^\alpha, n) = 0$, и перенесем его параллельно вдоль траектории. По-построению, для него выполнены равенства:

$$\dot{q}^\beta \nabla_\beta n^\alpha = 0. \tag{21.67}$$

Перепишем уравнение для экстремалей (21.64) в ковариантном виде,

$$\dot{q}^\beta \nabla_\beta \dot{q}^\alpha = 0.$$

Проекция этого уравнения на вектор n равна нулю,

$$(\dot{q}^\beta \nabla_\beta \dot{q}^\alpha) n_\alpha = \dot{q}^\beta \nabla_\beta (\dot{q}, n) - \dot{q}^\beta \dot{q}^\alpha \nabla_\beta n_\alpha = 0,$$

т.к. скалярное произведение при параллельном переносе сохраняется, $(\dot{q}, n) = \mathsf{const}$, и выполнено уравнение (21.67). Поскольку вектор n произволен, то это значит, что решение уравнения для экстремалей (21.65) зависит от производных компонент метрики, взятых только вдоль экстремалей. Значит, уравнение для экстремалей имеет смысл для метрики, компоненты которой имеют разрыв производных вдоль направлений, перпендикулярных траекториям. При этом само решение может быть дифференцируемым.

Таким образом, мы ищем дважды дифференцируемые классические решения $q_i^\alpha(\tau)$ уравнения для экстремалей (21.65) и обобщенное решение уравнений Эйнштейна (21.63) в классе метрик, имеющих непрерывные кусочно дифференцируемые компоненты. Вне мировых линий частиц компоненты метрики должны быть дважды непрерывно дифференцируемыми функциями. На мировых линиях частиц производные в направлениях, перпендикулярных мировым линиям, могут иметь разрывы. Производные вдоль мировых линий предполагаются непрерывными. Напомним, что решение уравнения называется обобщенным, если свертка его левой и правой части с произвольной пробной функцией приводит к равенству (см., например, [**?**]).

21.8.1 Нерелятивистский предел для точечной частицы

В настоящем разделе будет показана связь между уравнениями движения для точечных частиц (21.64) и хорошо знакомыми уравнениями движения частиц под действием гравитационного поля в механике Ньютона.

Пусть метрика пространства Минковского $\mathbb{R}^{1,n-1}$ имеет вид

$$\eta_{\alpha\beta} = \mathsf{diag}\,(1, \underbrace{-1, \ldots, -1}_{n-1}). \tag{21.68}$$

Для простоты рассмотрим движение одной частицы. В пространстве-времени $\mathbb{M}$ с нетривиальной метрикой $g_{\alpha\beta}(x)$ функции $\{q^\alpha(\tau)\}$ задают траекторию точечной частицы. Пусть $\tau = t$ – канонический параметр (собственное время). Предположим, что координата x^0 на $\mathbb{M}$ выбрана таким образом, что на траектории частицы $c\tau = ct = x^0$. Здесь мы ввели явно скорость света c для того, чтобы в дальнейшем строить разложение по малому параметру $\boldsymbol{u}^2/c^2 \ll 1$, где $\boldsymbol{u}$ – пространственная часть собственной скорости частицы. Условимся нумеровать, как обычно, пространственные координаты буквами из середины греческого алфавита:

$$\{x^\alpha\} = \{x^0, x^\mu\}, \qquad \mu = 1, \ldots, n-1.$$

Поскольку исходное действие (21.58) инвариантно относительно общих преобразований координат, то у нас имеется возможность дополнительно фиксировать $n-1$ компонент метрики. Положим $g_{0\mu} = 0$. Тогда метрика примет блочно диагональный вид

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & 0 \\ 0 & g_{\mu\nu} \end{pmatrix} \tag{21.69}$$

Другими словами, система координат выбрана таким образом, чтобы времениподобный вектор ∂_0 был ортогонален всем касательным векторам к пространственным сечениям $x^0 = \mathsf{const}$.

Введем два параметра разложения. Во-первых, нерелятивистский предел соответствует скоростям, малым по сравнению со скоростью света,

$$\frac{\boldsymbol{u}^2}{c^2} \sim \epsilon \ll 1, \qquad \boldsymbol{u}^2 := -\eta_{\mu\nu}\dot{q}^\mu\dot{q}^\nu = \delta_{\mu\nu}u^\mu u^\nu \geq 0.$$

Во-вторых, слабому гравитационному полю соответствует метрика, которая мало отличается от метрики Минковского:

$$g_{00} = 1 + h_{00}, \qquad g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad h_{00}, h_{\mu\nu} \sim \epsilon \ll 1.$$

Если гравитационное поле мало, $g_{\alpha\beta} \to \eta_{\alpha\beta}$, то в нерелятивистском пределе $\boldsymbol{u}^2 \to 0$ временна́я компонента собственной скорости стремится к скорости света, $\dot{q}^0 \to c$, т.к. выполнено соотношение (21.55). Если гравитационное поле не очень велико, $0 < g_{00} < \mathsf{const}$, то в нерелятивистском пределе $g_{00}(\dot{q}^0)^2 \to c^2$ и производная $\dot{q}^0$ ограничена снизу некоторой положительной постоянной. Поскольку пространственные компоненты собственной u^μ и наблюдаемой v^μ скорости частицы отличаются на строго положительный множитель $\dot{q}^0$, то пределы $\boldsymbol{u}^2 \to 0$ и $\boldsymbol{v}^2 \to 0$, где $\boldsymbol{v}^2 := -\eta_{\mu\nu}v^\mu v^\nu \to 0$, эквивалентны.

Нерелятивистской частице в слабом гравитационном поле соответствует интервал

$$ds^2 = (c^2 + c^2 h_{00} + \eta_{\mu\nu}u^\mu u^\nu + h_{\mu\nu}u^\mu u^\nu)dt^2.$$

В силу сделанных предположений о малости гравитационного поля и скоростей последним слагаемым в этом представлении можно пренебречь. Тогда в нерелятивистском пределе с учетом первой поправки интервал для точечной частицы примет вид

$$ds^2 \approx \left(c^2 + \frac{2U}{m} - \boldsymbol{u}^2\right) dt^2, \tag{21.70}$$

где введено обозначение

$$h_{00} = \frac{2U}{mc^2}. \tag{21.71}$$

Подставим приближенное выражение для интервала (21.70) в действие для точечной частицы (21.58), умноженное на скорость света, и разложим по степеням ϵ. Тогда в первом порядке по ϵ получим приближенное выражение

$$\begin{aligned} S_{\mathrm{m}} &= -mc \int d\tau \sqrt{c^2 + \frac{2U}{m} - \boldsymbol{u}^2} \approx \\ &\approx \int d\tau \left(-mc^2 + \frac{m\boldsymbol{u}^2}{2} - U\right). \end{aligned} \tag{21.72}$$

С точностью до энергии покоя точечной частицы с обратным знаком $-mc^2$ подынтегральное выражение совпадает с хорошо известным выражением для лагранжиана точечной частицы в нерелятивистской механике (20.17). Тем самым мы показали, что в нерелятивистском пределе поправка к временно́й компоненте метрики, умноженной на mc^2, следует интерпретировать, как потенциальную энергию $U = mc^2 h_{00}/2$ точечной частицы, находящейся во внешнем гравитационном поле.

Отметим, что разумный нерелятивистский предел обусловливает выбор общего знака минус в исходном действии для точечной частицы (21.58).

Интервалу (21.70) соответствует метрика

$$g_{\alpha\beta} = \mathsf{diag}\,(1 + 2U/m, -1, \ldots, -1). \tag{21.73}$$

Предположим, что потенциал U не зависит от времени t, то есть движение частицы происходит в статическом гравитационном поле. Символы Кристоффеля в первом порядке малости имеют только три нетривиальные компоненты:

$$\Gamma_{00}{}^{\mu} = -\frac{\eta^{\mu\nu}\partial_\nu U}{m}, \qquad \Gamma_{0\mu}{}^{0} = \Gamma_{\mu 0}{}^{0} = \frac{\partial_\mu U}{m}.$$

Соответствующие уравнения для экстремалей (17.17) принимают вид

$$\ddot{q}^0 = -2\frac{\partial_\mu U}{m}\dot{q}^\mu, \tag{21.74}$$

$$\ddot{q}^\mu = -\frac{\delta^{\mu\nu}\partial_\nu U}{m}. \tag{21.75}$$

Поскольку в выбранной системе координат $\ddot{q}^0 = 0$ и $\dot{U} = \dot{q}^\mu \partial_\mu U = 0$ (поле статично), то первое уравнение удовлетворяется автоматически. Второе уравнение совпадает с *вторым законом Ньютона* для движения точечной частицы в статическом потенциальном поле

$$m\ddot{q}^\mu = -\delta^{\mu\nu}\partial_\nu U. \tag{21.76}$$

В общей теории относительности ($n = 4$) метрика вдали от точечной массы M дается решением Шварцшильда (**??**). Соответствующее нерелятивистское выражение для потенциальной энергии имеет вид

$$U = -G\frac{mM}{r}, \tag{21.77}$$

где G – гравитационная постоянная, что совпадает с *законом всемирного тяготения*. Таким образом мы показали, что закон всемирного тяготения вытекает из общей теории относительности в нерелятивистском пределе.

21.8.2 Теория гравитации Ньютона

В настоящем разделе мы опишем гравитационное взаимодействие точечных частиц в механике Ньютона. При этом, по возможности, мы будем следовать общей схеме, принятой в общей теории относительности.

Пусть $\mathbb{M}$ – это тривиальное четырехмерное многообразие, $\mathbb{M} \approx \mathbb{R}^4$, с декартовой системой координат $x = \{x^\alpha\} = \{x^0, x^\mu\} = \{t, \boldsymbol{x}\}$, где $\alpha = 0, 1, 2, 3$ и $\mu = 1, 2, 3$. Как и раньше мы отождествим нулевую координату с временем, $x^0 = t$. Пусть в $\mathbb{M}$ движутся N точечных частиц по траекториям $q_i(t) = \{q_i^\mu(t)\}$, $i = 1, \ldots, \mathrm{N}$. В механике Ньютона время имеет абсолютный характер в том смысле, что оно одинаково для всех частиц и играет роль параметра вдоль траектории каждой частицы. Предположим, что между частицами действуют только гравитационные силы. Это значит, что каждая частица движется в гравитационном поле, которое создается другими частицами. В свою очередь каждая частица создает гравитационное поле, которое влияет на движение других частиц. Обозначим суммарный потенциал гравитационного поля через $\varphi(x)$, который, по определению, является функцией (скалярным полем) на $\mathbb{R}^4$.

Действие для точечных частиц, взаимодействующих посредством гравитационного поля, является суммой действия для гравитационного поля

$$S_{\mathrm{g}} = \frac{1}{4\pi G}\int_{\mathbb{R}^4} d^4x\,\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi, \tag{21.78}$$

где G – гравитационная постоянная и $\eta_{\mu\nu} = \mathsf{diag}\,(---)$ – отрицательно определенная пространственная метрика, и действия для точечных частиц

$$S_{\mathrm{m}} = -\sum_{i=1}^{\mathrm{N}}\int_{-\infty}^{\infty} dt\,\frac{1}{2}m_i\dot{q}_i^\mu\dot{q}_i^\nu\eta_{\mu\nu} - \sum_{i=1}^{\mathrm{N}}\int_{\mathbb{R}^4} d^4x\,m_i\delta^{(3)}(\boldsymbol{x} - \boldsymbol{q}_i)\varphi. \tag{21.79}$$

Действие для гравитационного поля (21.78) отрицательно определено и равно потенциальной энергии гравитационного поля, взятой с обратным знаком. Действие для точечных частиц (21.79) как обычно представляет собой разность кинетической и потенциальной энергии. Отметим, что потенциальная энергия взаимодействия точечных частиц с гравитационным полем содержит только трехмерную дельта-функцию. Варьирование суммарного действия, $S = S_{\mathrm{g}} + S_{\mathrm{m}}$, по гравитационному полю и траекториям частиц дает уравнения движения

$$\frac{\delta S}{\delta\varphi} = \frac{1}{4\pi G}\triangle\varphi - \sum_i m_i\delta^{(3)}(\boldsymbol{x} - \boldsymbol{q}_i) = 0, \tag{21.80}$$

$$\frac{\delta S}{\delta q_i} = m_i\left(\ddot{q}_{i\mu} - \partial_\mu\varphi\right) = 0, \tag{21.81}$$

где $\triangle = \partial_1^2 + \partial_2^2 + \partial_3^2$ – лапласиан, $\ddot{q}_{i\mu} = \ddot{q}_i^\nu \eta_{\mu\nu}$ и градиент потенциала гравитационного поля $\partial_\mu \varphi$ во втором уравнении берется в той точке $q_i = \{t, q_i^\mu\} \in \mathbb{M}$, где в данный момент времени расположена частица.

Уравнение для гравитационного поля (21.80) представляет собой уравнение Пуассона

$$\triangle \varphi = 4\pi G \sum_i m_i \delta^{(3)}(\boldsymbol{x} - \boldsymbol{q}_i). \tag{21.82}$$

Мы рассматриваем решения этого уравнения в слабом смысле, т.е. равенство достигается только после свертки левой и правой части с основными функциями. Если ограничить класс рассматриваемых решений только теми решениями, которые равны нулю на бесконечности, то решение единственно и имеет вид (см., например, [**?**])

$$\varphi(t, \boldsymbol{x}) = -G \sum_i \frac{m_i}{|\boldsymbol{x} - \boldsymbol{q}_i(t)|}, \tag{21.83}$$

где

$$|\boldsymbol{x} - \boldsymbol{q}_i| = \sqrt{(x^1 - q_i^1)^2 + (x^2 - q_i^2)^2 + (x^3 - q_i^3)^2} = \sqrt{-\eta_{\mu\nu}(x^\mu - q_i^\mu)(x^\nu - q_i^\nu)}$$

– расстояние от точки пространства $\boldsymbol{x} = \{x^\mu\} \in \mathbb{R}^3$ до i-той частицы в момент времени t.

Таким образом мы нашли потенциал гравитационного поля для произвольного движения частиц, которое описывается функциями $\boldsymbol{q}_i(t)$. Принято говорить, что гравитационное взаимодействие в механике Ньютона является дальнодействующим и распространяется с бесконечной скоростью. Это отражает тот факт, что решение (21.83) в произвольной точке пространства в произвольный момент времени однозначно определяется только массами частиц и их расположением в тот же момент времени. Можно сказать по другому: изменение положения частицы мгновенно приводит к изменению гравитационного поля во всем пространстве.

Теперь подставим решение для потенциала гравитационного поля (21.83) в уравнения движения точечных частиц (21.81),

$$\ddot{q}_{i\mu} = -G \frac{\partial}{\partial q_i^\mu} \sum_j \frac{m_j}{|\boldsymbol{q}_i - \boldsymbol{q}_j|},$$

чтобы полностью исключить потенциал гравитационного поля. Однако на этом этапе возникает серьезная трудность, т.к. правая часть уравнения расходится в точке $\boldsymbol{q}_i = \boldsymbol{q}_j$ и поэтому неопределена. Чтобы устранить эту трудность мы отбросим в сумме слагаемое с $i = j$. Физически это означает, что частица не взаимодействует с собственным гравитационным полем. Таким образом получаем систему, состоящую из 3N обыкновенных дифференциальных уравнений:

$$\ddot{q}_i^\mu = -G \sum_{i \neq j} m_j \frac{q_i^\mu - q_j^\mu}{|\boldsymbol{q}_i - \boldsymbol{q}_j|^3}. \tag{21.84}$$

В механике Ньютона уравнение (21.84) интерпретируется следующим образом. Если имеется всего две частицы с массами m_i и m_j, то между ними возникает притяжение, обусловленное гравитационным взаимодействием. При этом сила $\boldsymbol{F} = \{F^\mu\}$, действующая на частицу m_i со стороны частицы m_j, равна

$$F^\mu = -G m_i m_j \frac{q_i^\mu - q_j^\mu}{|\boldsymbol{q}_i - \boldsymbol{q}_j|^3}. \tag{21.85}$$

Это – *закон всемирного тяготения* Ньютона.

Замечание. В настоящее время закон всемирного тяготения Ньютона подтвержден экспериментально с высокой степенью точности в лабораторных условиях и в небесной механике. С его помощью в первом приближении рассчитывают движение планет в солнечной системе и звезд в галактиках. Численное значение гравитационной постоянной в системе СГС равно

$$G = (6,673 \pm 0,003) \cdot 10^{-8} \frac{\text{см}^3}{\text{г} \cdot \text{сек}^2}.$$

Общая теория относительности в первом приближении приводит к результатам, которые совпадают с результатами, полученными в рамках механики Ньютона. Кроме этого общая теория относительности приводит к поправкам, которые называются *постньютоновскими.* □

Систему уравнений движения для точечных частиц (21.84) можно получить из действия

$$S_{\text{eff}} = \int dt \left(-\sum_{i=1}^{\text{N}} \frac{1}{2} \dot{q}_i^\mu \dot{q}_i^\nu \eta_{\mu\nu} + \frac{1}{2} G \sum_{i=1}^{\text{N}} \sum_{j \neq i} \frac{m_i m_j}{|\boldsymbol{q}_i - \boldsymbol{q}_j|} \right), \tag{21.86}$$

которое варьируется только по траекториям частиц. Это действие получается из исходного действия для точечных частиц (21.79) подстановкой в него общего решения уравнения Эйлера–Лагранжа для потенциала (21.83). Эта процедура подстановки решения части уравнений Эйлера–Лагранжа в исходное действие с целью исключения некоторых динамических переменных была описана в разделе эффективное действие (19.4) в общем виде. Множитель $1/2$ во втором слагаемом в действии (21.86) возникает из-за двойной суммы, где каждое слагаемое встречается дважды.

Таким образом, движение точечных частиц, между которыми действуют только гравитационные силы, сводится к системе обыкновенных дифференциальных (21.84). Это полное описание, и в таком виде оно обычно встречается в курсах классической механики. Как видим, введение гравитационного поля $\varphi(x)$ в механике Ньютона совсем необязательно. Мы проделали более длинный путь с тем, чтобы показать аналогию с общей теорией относительности. К сожалению, решить уравнения Эйнштейна (21.63) для компонент метрики при произвольном движении частиц, как в случае механики Ньютона (21.83), не удается. Поэтому эффективное действие для точечных частиц в общей теории относительности в настоящее время неизвестно.

21.8.3 Свойства тензора энергии-импульса точечных частиц

В настоящем разделе мы обсудим некоторые свойства тензора энергии-импульса точечных частиц: ковариантное сохранение тензора энергии-импульса, аналогию с тензором энергии-импульса сплошной среды, неотрицательность следа тензора энергии-импульса и ультрарелятивистский предел.

Ковариантное сохранение тензора энергии-импульса

Поскольку тензор энергии-импульса (21.65) получен из вариации действия, инвариантного относительно общих преобразований координат, то, в силу второй теоремы Нетер (см. раздел 19.3), между уравнениями движения существует зависимость,

$$2\sqrt{|g|}\nabla_\gamma \left(\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta g_{\beta\gamma}} \right) g_{\alpha\beta} + \sum_i \frac{\delta S}{\delta q_i^\alpha} = 0.$$

На языке уравнений движения (21.63), (21.64) это означает следующее. Возьмем ковариантную дивергенцию от уравнения (21.63). Поскольку тензор Эйнштейна удовлетворяет свернутым тождествам Бианки (19.67),

$$\nabla_\gamma G^{\beta\gamma} = 0,$$

то тензор энергии-импульса точечных частиц ковариантно сохраняется,

$$\nabla_\alpha T_{\rm m}{}^{\alpha\beta} = 0, \tag{21.87}$$

для любого решения системы уравнений движения для точечных частиц (21.64). То есть условия интегрируемости уравнений Эйнштейна выполняются автоматически, если выполнены уравнения движения точечных частиц.

Явная проверка ковариантного сохранения тензора энергии-импульса точечных частиц (21.87) требует осторожности. Поэтому проведем соответствующие вычисления, для простоты, для одной частицы,

$$\begin{aligned}\nabla_\beta T_{\rm m}{}^{\alpha\beta} &= \partial_\alpha T_{\rm m}{}^{\alpha\beta} + \Gamma_\alpha T_{\rm m}{}^{\alpha\beta} + \Gamma_{\alpha\gamma}{}^{\beta} T_{\rm m}{}^{\alpha\gamma} = \\ &= \frac{m}{\sqrt{|g|}}\left[\partial_\alpha \int d\tau \dot q^\alpha \dot q^\beta \delta^{(n)}(x-q) + \Gamma_{\alpha\gamma}{}^{\beta}\int d\tau \dot q^\alpha \dot q^\gamma \delta^{(n)}(x-q)\right],\end{aligned}$$

где мы воспользовались тождеством (7.46)

$$\partial_\alpha \sqrt{|g|} = \sqrt{|g|}\Gamma_\alpha.$$

Во втором слагаемом символы Кристоффеля можно внести под знак интеграла и рассматривать их, как функции от q ввиду наличия δ-функции. Для первого слагаемого в слабом смысле справедливо равенство

$$\partial_\alpha \int d\tau \dot q^\alpha \dot q^\beta \delta^{(n)}(x-q) = \int d\tau \ddot q^\beta \delta^{(n)}(x-q). \tag{21.88}$$

Чтобы доказать это равенство, напомним определения [**?**].

Определение. Назовем функцию φ на многообразии $\mathbb{M}$, $\dim \mathbb{M} = n$, *финитной*, если ее носитель (замыкание множества точек, в которых функция отлична от нуля) компактен. Пространство $\mathcal{D}(\mathbb{M})$, состоящее из всех финитных бесконечно дифференцируемых функций на $\mathbb{M}$ называется пространством *основных* (*пробных*) функций. *Обобщенной функцией* или *распределением* f на многообразии $\mathbb{M}$ называется линейный непрерывный функционал на пространстве основных функций. Равенство двух обобщенных функций $f = g$ называется *слабым*, если для любой пробной функции справедливо равенство

$$\int_{\mathbb{M}} d^n x\, \varphi f = \int_{\mathbb{M}} d^n x\, \varphi g,$$

для всех $\varphi \in \mathcal{D}(\mathbb{M})$. □

Мы не будем останавливаться на определении топологии в пространствах основных и обобщенных функций, отсылая читателя к [**?**].

Вернемся к равенству (21.88). Умножим левую часть на пробную функцию $\varphi \in \mathcal{D}(\mathbb{M})$ и проинтегрируем по $\mathbb{M}$:

$$\begin{aligned}\int_{\mathbb{M}} d^n x\, \varphi \partial_\alpha \left(\int d\tau \dot q^\alpha \dot q^\beta \delta^{(n)}(x-q)\right) &= -\int_{\mathbb{M}} d^n x\, \partial_\alpha \varphi \int d\tau \dot q^\alpha \dot q^\beta \delta^{(n)}(x-q) = \\ &= -\int_{\mathbb{M}} d^n x \int d\tau \frac{d\varphi}{d\tau}\dot q^\beta \delta^{(n)}(x-q) = -\int d\tau \frac{d\varphi}{d\tau}\dot q^\beta = \int d\tau\, \varphi \ddot q^\beta,\end{aligned}$$

где мы два раза проинтегрировали по частям, воспользовались равенством $d\varphi/d\tau = \dot{q}^\alpha \partial\varphi/\partial q^\alpha$ и взяли интеграл по $\mathbb{M}$, используя δ-функцию. Интеграл по $\mathbb{M}$ от правой части (21.88) с пробной функцией приводит, очевидно, к тому же результату. Таким образом, слабое равенство (21.88) доказано.

Теперь равенство (21.87) можно переписать в виде

$$\nabla_\alpha T_{\rm m}{}^{\alpha\beta} = \frac{1}{\sqrt{|g|}} \int d\tau \sum_i m_i \left(\ddot{q}_i^\beta + \Gamma_{\alpha\gamma}{}^\beta \dot{q}_i^\alpha \dot{q}_i^\gamma \right) \delta^{(n)}(x - q), \tag{21.89}$$

где мы вернулись к общему случаю N частиц. Отсюда следует, что тензор энергии-импульса точечных частиц ковариантно сохраняется, если выполнены уравнения движения (21.64). Таким образом мы привели прямое доказательство

Предложение 21.8.2. *Тензор энергии-импульса точечных частиц ковариантно сохраняется (21.87) тогда и только тогда, когда выполнены уравнения движения для точечных частиц.*

Доказательство. То, что из уравнений движения точечных частиц следует ковариантное сохранение тензора энергии-импульса, было доказано выше. Докажем обратное утверждение. Для этого умножим равенство (21.89) на пробную функцию и проинтегрируем по $\mathbb{M}$, используя δ-функцию. Если тензор энергии-импульса ковариантно сохраняется, то должно быть выполнено равенство

$$\int d\tau \varphi \frac{1}{\sqrt{|g|}} \sum_i m_i \left(\ddot{q}_i^\beta + \Gamma_{\alpha\gamma}{}^\beta \dot{q}_i^\alpha \dot{q}_i^\gamma \right) = 0.$$

Поскольку пробная функция φ произвольна, и ее носитель может быть отличен от нуля для одной произвольной частицы, то отсюда следуют уравнения движения для точечных частиц (21.64). □

Таким образом, ковариантный закон сохранения тензора энергии-импульса точечных частиц (21.87), который понимается в слабом смысле, эквивалентен уравнениям движения (21.64). Это означает, что вместо решения полной системы уравнений для точечных частиц и компонент метрики достаточно решить уравнения Эйнштейна (21.63).

Аналогия с тензором энергии-импульса сплошной среды

Введем новую временну́ю координату $\tau = \tau(x^0, x^\mu)$ в пространстве-времени $\mathbb{M}$ таким образом, чтобы вдоль каждой мировой линии частицы она совпадала с собственным временем $\tau(x^\alpha = q_i^\alpha) = \tau$. Это всегда можно сделать, причем не единственным образом, так как траектории всех частиц не пересекаются, времениподобны, а канонический параметр определен с точностью до сдвигов. Тогда в новой системе координат τ, x^μ производные $\dot{q}_i^\alpha$ в выражении (21.65) можно заменить на частные производные $\dot{x}^\alpha = \partial x^\alpha/\partial\tau$ ввиду наличия δ-функций, и вынести за знак интегрирования

$$T_{\rm m}{}^{\alpha\beta} = \frac{\dot{x}^\alpha \dot{x}^\beta}{\sqrt{|g|}} \int d\tau \sum_i m_i \delta^{(n)}(x - q_i) = \frac{\dot{x}^\alpha \dot{x}^\beta}{\sqrt{|g|}} \sum_i \frac{m_i}{\dot{q}_i^0} \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i). \tag{21.90}$$

Полученное выражение для тензора энергии-импульса точечных частиц имеет такой же вид, как и для сплошной среды (21.117). Для точечных частиц давление равно

нулю, $\mathcal{P} = 0$, а плотность энергии принимает вид

$$\mathcal{E} = \frac{1}{\sqrt{|g|}} \sum_i \frac{m_i}{\dot{q}_i^0} \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i).$$

Поскольку временна́я координата x^0 на траекториях частиц совпадает с собственным временем τ, то $\dot{q}_i^0 = 1$ и выражение для плотности энергии приобретает интуитивно ясную форму,

$$\mathcal{E} = \frac{1}{\sqrt{|g|}} \sum_i m_i \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i).$$

То есть энергия сосредоточена в точках расположения частиц, и каждая частица несет энергию, которая равна ее массе. Тензор энергии-импульса точечных частиц соответствует пылевидной материи, поскольку давление равно нулю.

След тензора энергии-импульса

Из формулы (21.66) следует выражение для следа тензора энергии-импульса произвольного распределения точечных частиц

$$T_{\mathrm{m}}{}^{\alpha}{}_{\alpha} = \frac{1}{\sqrt{|g|}} \sum_i \frac{m_i}{\dot{q}_i^0} \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i). \tag{21.91}$$

Поскольку $m_i > 0$ и $\dot{q}_i^0 > 0$, то след тензора энергии-импульса положителен (при этом мы рассматриваем δ-функцию, как положительную).

Замечание. Поскольку след тензора энергии-импульса положителен для *произвольного* распределения точечных частиц, то в моделях математической физики делается предположение о том, что след тензора энергии-импульса для любой обычной (наблюдаемой) материи всегда неотрицателен, $T^{\alpha}{}_{\alpha} \geq 0$. При этом равенство следа тензора энергии-импульса нулю, как будет показано ниже, достигается только для частиц, движущихся со скоростью света, или излучения. □

След тензора энергии-импульса электромагнитного поля 23.141, который соответствует электромагнитному излучению, также равен нулю. Это согласуется с утверждением о том, что след тензора энергии-импульса произвольного распределения ультрарелятивистских частиц равен нулю. Напомним, что в квантовой электродинамике электромагнитное поле описывает безмассовые частицы – фотоны, которые распространяются со скоростью света.

Мы выделили рассмотрение следа тензора энергии-импульса точечных частиц в отдельный пункт именно в свете последнего замечания, т.к. положительность следа тензора энергии-импульса для произвольной материи ниоткуда больше не следует, и в то же время ведет к важным следствиям.

Ультрарелятивистский предел

Рассмотрим ультрарелятивистский предел для тензора энергии-импульса точечных частиц (21.90). Этот предел, прежде всего, требует определения, потому что, глядя на определение собственной скорости частицы (21.52), непонятно что и куда стремить. Тензор энергии-импульса точечных частиц и его след можно выразить через

наблюдаемую скорость

$$T_{\rm m}{}^{\alpha\beta} = \frac{v^\alpha v^\beta}{\sqrt{|g|v^2}} \sum_i m_i \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i), \tag{21.92}$$

$$T_{\rm m}{}^{\alpha}{}_{\alpha} = \frac{\sqrt{v^2}}{\sqrt{|g|}} \sum_i m_i \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{q}_i), \tag{21.93}$$

где мы воспользовались равенствами (21.56) и (21.57).

Определение. Предел, когда квадрат наблюдаемой скорости частицы (21.54) стремится к нулю,

$$v^2 = v^\alpha v^\beta g_{\alpha\beta} \to 0, \tag{21.94}$$

называется *ультрарелятивистским пределом* для точечной частицы. □

Замечание. Если метрика имеет блочно диагональный вид (21.69), то

$$v^2 = g_{00} - \boldsymbol{v}^2, \quad \text{и} \quad \{u^\alpha\} = \left\{ \frac{1}{\sqrt{g_{00} - \boldsymbol{v}^2}}, \frac{v^\mu}{\sqrt{g_{00} - \boldsymbol{v}^2}} \right\}$$

где $\boldsymbol{v}^2 = -v^\mu v^\nu g_{\mu\nu} \geq 0$ – квадрат пространственной наблюдаемой скорости и собственная скорость частицы u^α определена в (21.52). В этом случае ультрарелятивистский предел соответствует пределу $\boldsymbol{v}^2 \to g_{00}$. Для пространства Минковского это означает, что наблюдаемая скорость частицы стремится к скорости света. □

В ультрарелятивистском пределе компоненты самого тензора энергии-импульса точечных частиц (21.92) не определены, однако след тензора энергии-импульса (21.93) стремится к нулю

$$\lim_{v^2 \to 0} T_{\rm m}{}^{\alpha}{}_{\alpha} = 0.$$

Замечание. Это утверждение справедливо для произвольного количества и распределения точечных частиц, находящихся только под действием гравитационных сил. Если присутствуют другие взаимодействия, то действие для точечных частиц (21.59) может измениться и, следовательно, изменится выражение для тензора энергии-импульса. В таком случае требуется дополнительное исследование. □

21.9 Ньютонов предел

Для того, чтобы сказать, что общая теория относительности не противоречит экспериментальным данным, желательно показать, что теория тяготения Ньютона в каком то смысле (приближении) следует из уравнений Эйнштейна. Поскольку гравитация Ньютона находится в хорошем согласии с экспериментом, то в этом случае можно утверждать, что общая теория относительности описывает гравитационные взаимодействия по крайней мере не хуже, чем законы Ньютона. Такое приближение существует, и будет описано в настоящем разделе.

Сначала сделаем общее замечание. Уравнения Эйнштейна существенно нелинейны, в то время как гравитация Ньютона линейна: гравитационные потенциалы различных массивных тел просто складываются. Поэтому естественно ожидать, что закон всемирного тяготения вытекает из уравнений Эйнштейна в линейном приближении.

Рассмотрим вакуумные уравнения Эйнштейна без космологической постоянной в четырехмерном пространстве-времени

$$\kappa\left(R_{\alpha\beta}-\frac{1}{2}R\right)=-\frac{1}{2}T_{\text{м}\alpha\beta}. \tag{21.95}$$

Будем считать, что пространство-время топологически тривиально, и существует глобальная система координат, в которой метрика, удовлетворяющая уравнениям (21.95), мало отличается от метрики Лоренца в пространстве Минковского $\mathbb{R}^{1,3}$:

$$g_{\alpha\beta}=\eta_{\alpha\beta}+\epsilon h_{\alpha\beta},\qquad \epsilon\ll 1. \tag{21.96}$$

При этом мы считаем малыми также все частные производные: $\partial_\alpha g_{\beta\gamma}\sim\epsilon$. Символы Кристоффеля пропорциональны производным $\partial_\alpha g_{\beta\gamma}$, и поэтому их квадраты дают вклад в тензор кривизны порядка ϵ^2. Тем самым вкладом квадратичных слагаемых по символам Кристоффеля в тензор кривизны (**??**) можно пренебречь по сравнению со вторыми производными от компонент метрики, которые дают вклад прядка ϵ. Таким образом, в линейном приближении по ϵ тензор Риччи имеет вид

$$\frac{1}{\epsilon}R_{\alpha\beta}=\frac{1}{2}\eta^{\gamma\delta}(\partial^2_{\alpha\beta}h_{\gamma\delta}-\partial^2_{\alpha\gamma}h_{\beta\delta}-\partial^2_{\beta\gamma}h_{\alpha\delta}+\partial^2_{\gamma\delta}h_{\alpha\beta}).$$

В правой части равенства свертка проводится с помощью метрики Лоренца, т.к. выражение в скобках имеет первый порядок. Скалярная кривизна имеет вид

$$\frac{1}{\epsilon}R=\eta^{\alpha\beta}\eta^{\gamma\delta}(\partial^2_{\alpha\beta}h_{\gamma\delta}-\partial^2_{\alpha\gamma}h_{\beta\delta}).$$

Введем новые переменные

$$\bar h_{\alpha\beta}:=h_{\alpha\beta}-\frac{1}{2}\eta_{\alpha\beta}h, \tag{21.97}$$

где $h:=\eta^{\alpha\beta}h_{\alpha\beta}$ – след возмущения метрики. Обратное преобразование имеет вид

$$h_{\alpha\beta}=\bar h_{\alpha\beta}-\frac{1}{2}\eta_{\alpha\beta}\bar h,\qquad \bar h:=\eta^{\alpha\beta}\bar h_{\alpha\beta}. \tag{21.98}$$

Тогда тензор Эйнштейна в линейном приближении примет вид

$$\frac{1}{\epsilon}\left(R_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}R\right)=\frac{1}{2}\eta^{\gamma\delta}\left(\partial^2_{\gamma\delta}\bar h_{\alpha\beta}-\partial^2_{\alpha\gamma}\bar h_{\beta\delta}-\partial^2_{\beta\gamma}\bar h_{\alpha\delta}+\eta_{\alpha\beta}\eta^{\epsilon\zeta}\partial^2_{\gamma\epsilon}\bar h_{\delta\zeta}\right).$$

Теперь воспользуемся инвариантностью действия Гильберта–Эйнштейна относительно общих преобразований координат. Рассмотрим бесконечно малые преобразования координат, которые генерируются некоторым векторным полем (3.94),

$$x^\alpha\mapsto x^\alpha+\epsilon u^\alpha(x). \tag{21.99}$$

При этом компоненты метрики получат приращение (3.102)

$$h_{\alpha\beta}\mapsto h_{\alpha\beta}+\partial_\alpha u_\beta+\partial_\beta u_\alpha. \tag{21.100}$$

Возьмем в качестве векторного поля u^α произвольное решение уравнения

$$\Box u^\alpha=-\partial_\beta\bar h_\alpha{}^\beta,$$

где $\Box:=\eta^{\alpha\beta}\partial_\alpha\partial_\beta$ – оператор Даламбера. Тогда в новой системе координат возмущение компонент метрики будет удовлетворять уравнению

$$\partial_\beta\bar h_\alpha{}^\beta=0. \tag{21.101}$$

Это есть ни что иное как условие гармоничности координат (17.58) в линейном приближении.

Замечание. Гармонические координаты в общей теории относительности являются аналогом лоренцевой калибровки в электродинамике (см. раздел 23.3). □

С учетом условия гармоничности уравнения Эйнштейна (21.95) принимают вид

$$\epsilon\Box\bar{h}_{\alpha\beta} = -\frac{1}{\kappa}T_{\text{м}\alpha\beta}. \tag{21.102}$$

Замечание. Если поля материи отсутствуют, $T_{\text{м}\alpha\beta} = 0$, то система уравнений (21.101), (21.102) совпадает с уравнениями для безмассового поля со спином 2 в плоском пространстве-времени Минковского [**?**]. Поэтому общую теорию относительности в целом можно рассматривать как теорию безмассового поля со спином 2 и с некоторым самодействием, которое соответствует отброшенным нелинейным членам. Следует однако заметить, что понятие массы и спина требует наличия метрики Лоренца, которая является фоновой метрикой для линейного приближения. В общем случае, без обращения к линейному приближению, утверждению о том, что метрика описывает безмассовое поле спина 2 придать точный смысл нельзя. □

Рассмотрим в качестве источника в уравнениях Эйнштейна одну частицу массы M. Поскольку мы рассматриваем слабые гравитационные поля, то будем считать, что $M \sim \epsilon$. Этой частице соответствует тензор энергии-импульса (21.66)

$$T_{\text{м}\alpha\beta} = \frac{1}{\sqrt{|g|}}M\frac{\dot{q}_\alpha\dot{q}_\beta}{\dot{q}^0}\delta^{(3)}(\boldsymbol{x}-\boldsymbol{q}).$$

В линейном приближении по ϵ можно сделать замену $\sqrt{|g|} \mapsto 1$.

Предположим, что частица покоится в начале координат, т.е. $\{\dot{q}^\alpha\} = (1,0,0,0)$ и $\{q^\alpha\} = (\tau,0,0,0)$. Предположим также, что компоненты метрики не зависят от времени (статическое решение). Тогда полная система уравнений Эйнштейна примет вид

$$\kappa\epsilon\triangle\bar{h}_{00} = M\delta^{(3)}(\boldsymbol{x}), \tag{21.103}$$

$$\triangle\bar{h}_{0\mu} = \triangle\bar{h}_{\mu\nu} = 0, \tag{21.104}$$

где $\triangle := \partial_1^2 + \partial_2^2 + \partial_3^2$ – лапласиан в трехмерном евклидовом пространстве. Если предположить, что компоненты возмущений метрики $\bar{h}_{\alpha\beta}$ стремятся к нулю на бесконечности, то уравнения (21.104) имеют единственное решение

$$\bar{h}_{0\mu} = 0, \qquad \bar{h}_{\mu\nu} = 0.$$

Для сравнения уравнения (21.103) с законом всемирного тяготения, необходимо восстановить размерные постоянные. Во-первых, положим

$$\epsilon\bar{h}_{00} = \frac{4\varphi}{c^2},$$

где φ – потенциал гравитационного поля. Это следует из нерелятивистского предела для точечной частицы (21.71). Кроме того, в правую часть уравнения (21.103) надо вставить множитель c^2: один множитель c следует из опущенного множителя в действии для точечной частицы (21.58), а второй – из равенства $\dot{q}^0 = c$. Если после этого положить

$$\kappa := \frac{c^4}{16\pi G}, \tag{21.105}$$

где G – гравитационная постоянная в законе тяготения Ньютона, то уравнение (21.103) совпадет с уравнением Пуассона для гравитационного поля (21.82). В этом случае решение уравнения Пуассона (21.103) примет вид

$$\varphi = -G\frac{M}{r},$$

где $r = |\boldsymbol{x}|$.

Ясно, что для найденных компонент метрики калибровочное условие (21.101) выполнено, и, следовательно, найдено самосогласованное решение задачи.

Поскольку след $\epsilon\bar{h} = 4\varphi/c^2$, то из формулы для обратного преобразования возмущения компонент метрики(21.98), получаем, что в ньтоновом приближении метрика имеет вид

$$g = \begin{pmatrix} 1 - G\dfrac{2M}{r} & 0 & 0 & 0 \\ 0 & -1 - G\dfrac{2M}{r} & 0 & 0 \\ 0 & 0 & -1 - G\dfrac{2M}{r} & 0 \\ 0 & 0 & 0 & -1 - G\dfrac{2M}{r} \end{pmatrix}. \tag{21.106}$$

Напомним, что метрика имеет такой вид в линейном приближении в гармонической системе координат.

Таким образом, мы показали, что теория тяготения Ньютона согласуется с общей теорией относительности. Она возникает в статическом случае для слабых гравитационных полей. При этом константа связи перед действием Гильберта–Эйнштейна имеет вид (21.105). По этой причине многие авторы записывают действие Гильберта–Эйнштейна именно с такой константой связи, часто полагая $c = 1$.

Замечание. Наш выбор выражения для тензора кривизны, тензора Риччи и скалярной кривизны через метрику и аффинную связность согласован с обычными обозначениями в теории калибровочных полей и приспособлен к анализу общей теории относительности. В то же время он отличается от стандартных определений, принятых в математической литературе. Например, скалярная кривизна сферы радиуса r, вложенной в трехмерное евклидово пространство, $\mathbb{S}^2_r \hookrightarrow \mathbb{R}^3$, равна $-2/r^2$. В этом нет ничего страшного. Просто об этом следует помнить. □

Несмотря на то, что общая теория относительности содержит в себе теорию тяготения Ньютона в качестве предельного случая, отметим принципиальное отличие. В механике Ньютона свободная частица движется по прямой линии. Если она находится в поле другой массивной частицы, то на нее действует сила гравитационного притяжения. Теперь она уже не является свободной и ее траектория отличается от прямой линии в соответствии с законом всемирного тяготения. В общей теории относительности ситуация совершенно иная. Массивная частица искривляет пространство-время в соответствии с уравнениями Эйнштейна. Пробная частица в гравитационном поле остается свободной и движется вдоль экстремали. Однако теперь экстремаль не является прямой линией, поскольку пространство-время перестает быть плоским из-за наличия массивной частицы.

21.10 Гравитационные волны

В механике Ньютона гравитационных волн нет, что вытекает из системы уравнений (21.80), (21.81). При этом изменение положения одного из массивных тел мгновенно приводит к изменению гравитационного поля во всем пространстве. Кроме этого, если массивные тела отсутствуют, то потенциал гравитационного поля равен нулю. В общей теории относительности ситуация другая. Во-первых, гравитационные взаимодействия распространяются с конечной скоростью света c. Во-вторых, даже если материальные тела отсутствуют, уравнения Эйнштейна допускают нетривиальные решения в виде гравитационных волн. В настоящем разделе мы изучим решения уравнений Эйнштейна, описывающие гравитационные волны.

Рассмотрим вакуумные уравнения Эйнштейна без космологической постоянной

$$R_{\alpha\beta} = 0. \tag{21.107}$$

Как и в предыдущем разделе будем считать, что пространство-время топологически тривиально, $\mathbb{M} \approx \mathbb{R}^4$, и существует глобальная система координат, в которой метрика мало отличается от метрики Лоренца (21.96). В нулевом порядке по ϵ вакуумные уравнения Эйнштейна, очевидно, удовлетворяются, т.к. кривизна пространства Минковского равна нулю. Найдем решение уравнений (21.107) в первом порядке по ϵ.

Используя инвариантность действия Гильберта–Эйнштейна, выберем систему координат таким образом, чтобы выполнялось калибровочное условие (21.101) (гармонические координаты). Тогда все компоненты метрики в первом порядке будут удовлетворять волновому уравнению

$$\Box \bar{h}_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \Box h_{\alpha\beta} = 0,$$

где компоненты $\bar{h}_{\alpha\beta}$ определены уравнением (21.97). Однако далеко не все компоненты метрики являются независимыми и описывают физические степени свободы, от которых нельзя избавиться путем выбора соответствующей системы координат.

Покажем это. Калибровочное условие (21.101) не фиксирует систему координат однозначно. Действительно, допустим, что в некоторой системе координат это условие выполнено. Совершим преобразование координат $x^\alpha \mapsto x^\alpha + \epsilon u^\alpha$, где все компоненты векторного поля u удовлетворяют волновому уравнению

$$\Box u^\alpha = 0. \tag{21.108}$$

Нетрудно проверить, что в новой системе координат калибровочное условие (21.101) будет также выполнено. Следовательно, оставшуюся свободу в выборе системы координат можно использовать для того, чтобы обратить в нуль часть компонент метрики.

Чтобы найти подходящие калибровочные условия, необходимо решить уравнения (21.108) с некоторыми начальными условиями. Чтобы показать откуда появляется система уравнений на начальные данные, посмотрим как преобразуется след возмущений метрики $h := h_\alpha{}^\alpha$ и компоненты $h_{0\mu}$, $\mu = 1, 2.3$, при бесконечно малых преобразованиях координат (24.64):

$$\begin{aligned} h &\mapsto h + 2\partial_\alpha u^\alpha, \\ h_{0\mu} &\mapsto h_{0\mu} + \partial_0 u_\mu + \partial_\mu u_0. \end{aligned}$$

Рассмотрим пространственное сечение $x^0 := t = \mathsf{const}$ в качестве поверхности Коши. На этой поверхности найдем какое либо решение системы линейных уравнений

$$2\left(\dot{u}_0 + \partial_\mu u^\mu\right) = -h, \tag{21.109}$$

$$2\left(\triangle u_0 + \partial_\mu \dot{u}^\mu\right) = -\dot{h}, \tag{21.110}$$

$$\dot{u}_\mu + \partial_\mu u_0 = -h_{0\mu}, \tag{21.111}$$

$$\triangle u_\mu + \partial_\mu \dot{u}_0 = -\dot{h}_{0\mu}, \tag{21.112}$$

где точка обозначает дифференцирование по времени t. Из этой системы уравнений определяем компоненты u^α и их производные по времени $\dot{u}^\alpha$. После этого решаем задачу Коши для уравнений (21.108) с найденными начальными условиями в обе стороны и определяем векторное поле u во всем пространстве-времени. Поскольку выполнены уравнения (21.109)–(21.112), то на поверхности Коши справедливы равенства

$$\begin{aligned} h = 0, \quad \dot{h} = 0, \\ h_{0\mu} = 0, \quad \dot{h}_{0\mu} = 0. \end{aligned} \tag{21.113}$$

Это очевидно, если заметить, что уравнения (21.110) и (21.112) возникают после дифференцирования уравнений (21.109) и (21.111) по времени и использования уравнений движения. Поскольку компоненты h и $h_{0\mu}$ удовлетворяют волновому уравнению, то данные условия выполнены также во всем пространстве-времени.

Определение. Система координат, в которой в линейном приближении выполнены условия

$$\partial_\beta h_\alpha{}^\beta = 0, \quad h = 0, \quad h_{0\mu} = 0 \tag{21.114}$$

называется *радиационной калибровкой.* □

Предложение 21.10.1. *При отсутствии полей материи и космологической постоянной радиационная калибровка существует в линейном приближении к метрике Лоренца.*

Доказательство. Было приведено выше. □

Замечание. В электродинамике аналогичная калибровка называется кулоновской. □

Из первого условия (21.114) с учетом того, что $h_{0\mu} = 0$, получаем уравнение $\dot{h}_{00} = 0$, которое должно быть выполнено во всем пространстве-времени. Тогда уравнение движения для временно́й компоненты сводится к уравнению Лапласа $\triangle h_{00} = 0$. С учетом нулевых граничных условий на бесконечности получаем дополнительное условие на компоненты метрики: $h_{00} = 0$.

Рассмотрим плоскую волну, которая распространяется в направлении волнового вектора $k = \{k^\alpha\}$:

$$h_{\alpha\beta} = H_{\alpha\beta}\, \mathsf{e}^{ik_\gamma x^\gamma}, \tag{21.115}$$

где $H_{\alpha\beta}$ – некоторая постоянная матрица и $k^2 := k^\alpha k_\alpha = 0$. Радиационная калибровка (21.114) для этого решения уравнений движения задается следующими восемью условиями:

$$k^\beta H_{\alpha\beta} = 0, \quad H_\alpha{}^\alpha = 0, \quad H_{0\mu} = 0.$$

Из первого и третьего условия вытекает равенство $k^0 H_{00}$. Для нетривиального решения $k^0 \neq 0$, и поэтому $H_{00} = 0$. Ввиду симметрии по индексам матрица $H_{\alpha\beta}$ имеет

10 независимых элементов. Радиационная калибровка накладывает на них 8 независимых условий. Отсюда вытекает, что в выбранной системе координат только 2 компоненты возмущения метрики являются независимыми.

Допустим, что гравитационная волна распространяется вдоль оси x^1, т.е. нормированный волновой вектор имеет только две отличные от нуля компоненты $k = (1, 1, 0, 0)$. Тогда матрица $H_{\alpha\beta}$ в радиационной калибровке имеет вид

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & A & B \\ 0 & 0 & B & -A \end{pmatrix},$$

где $H_{22} = -H_{33} = A$ и $H_{23} = H_{32} = B$ – два произвольных числа (амплитуды волн).

Введем новое понятие спиральности плоской волны. Для этого рассмотрим вращение пространства $\mathbb{R}^3 \subset \mathbb{R}^{1,3}$ на угол ω вокруг оси x^1. Матрица вращения задана в явном виде (2.60), и ясно, что такое вращение не меняет волнового вектора k. При преобразовании координат компоненты метрики преобразуются по тензорному закону

$$g_{\alpha\beta} \mapsto g'_{\alpha\beta} = S_\alpha{}^\gamma S_\beta{}^\delta g_{\gamma\delta}.$$

Простые вычисления приводят к равенствам:

$$\begin{aligned} A' &= \quad \cos 2\omega A + \sin 2\omega B, \\ B' &= -\sin 2\omega A + \cos 2\omega B. \end{aligned}$$

Это означает, что при повороте системы координат на угол ω амплитуда волны поворачивается на удвоенный угол 2ω. В физике часто рассматривают комплексные амплитуды

$$H_\pm := H_{22} \mp iH_{23} = A \mp iB.$$

При вращении они преобразуются по правилу

$$H'_\pm = \mathrm{e}^{\pm 2i\omega} H_\pm.$$

Определение. Если амплитуда плоской поперечной волны при повороте на угол ω вокруг направления распространения волны поворачивается на угол $h\omega$, то говорят, что волна имеет *спиральность* h. □

Таким образом, плоские гравитационные волны описывают поперечные волны спиральности два.

Тензор Риччи и скалярная кривизна для данного решения вакуумных уравнений Эйнштейна в линейном приближении равны, конечно, нулю. Тем не менее полный тензор кривизны отличен от нуля. В линейном приближении тензор кривизны имеет вид (см. выражение (**??**))

$$\frac{1}{\epsilon} R_{\alpha\beta\gamma\delta} = \frac{1}{2}(\partial^2_{\alpha\gamma} h_{\beta\delta} - \partial^2_{\alpha\delta} h_{\beta\gamma} - \partial^2_{\beta\gamma} h_{\alpha\delta} + \partial^2_{\beta\delta} h_{\alpha\gamma}). \tag{21.116}$$

Простые вычисления показывают, что среди 20 независимых компонент тензора кривизны только 9 отличны от нуля:

$$\begin{aligned} &R_{0303} = -R_{0202} = R_{1313} = -R_{1212} = R_{0212} = -R_{0313} = \frac{1}{2}\epsilon A\, \mathrm{e}^{i(t-x^1)}, \\ &R_{0203} = R_{1213} = -R_{0213} = -\frac{1}{2}\epsilon B\, \mathrm{e}^{i(t-x^1)}. \end{aligned}$$

При преобразовании координат компоненты тензора кривизны ведут себя ковариантным образом – на то он и тензор. Однако в линейном приближении они не ковариантны, а инвариантны. Нетрудно проверить, что выражение (21.116) действительно инвариантно относительно преобразований (21.100) с произвольным вектором u.

Из явного выражения для нетривиальных компонент тензора кривизны вытекает, что амплитуды волн A и B нельзя обратить в нуль никаким преобразованием координат. Следовательно, они описывают физические распространяющиеся степени свободы.

Поскольку вакуумные уравнения Эйнштейна в линейном приближении линейны, то им будет удовлетворять произвольная суперпозиция плоских волн. В частности, поправки к метрике вида

$$h_{\alpha\beta} = H_{\alpha\beta}\Big[f(x^1 - t) + g(x^1 + t)\Big],$$

где f и g – произвольные функции, описывающие распространение волн вдоль оси x^1 в положительную и отрицательную стороны, также удовлетворяют линеаризованным уравнениям Эйнштейна в радиационной калибровке. Отсюда, в частности, следует, что для однозначного задания волнового решения уравнений Эйнштейна необходимо задать четыре функции на поверхности Коши: по две для каждой волны. При этом функции можно задать произвольным образом. Таким образом, вакуумные уравнения Эйнштейна без космологической постоянной описывают распространение двух физических степеней свободы, которые порождают нетривиальную кривизну пространства-времени и не устраняются никаким преобразованием координат. Данный подсчет степеней свободы приводит к тому же результату, что и общий подход, основанный на гамильтоновом формализме (см. главу 22).

21.11 Сплошная среда в общей теории относительности

В правой части уравнений Эйнштейна (6.27) находится тензор энергии-импульса материи $T_{\text{м}\alpha\beta}$. В случае скалярного, электромагнитного и других полей, уравнения движения которых следуют из вариационного принципа, правая часть уравнений Эйнштейна определяется вариацией соответствующего действия по метрике. В этом случае вопросов с определением тензора энергии-импульса материи не возникает. Некоторые из этих тензоров будут рассмотрены далее.

В то же время в общей теории относительности существует ряд важных моделей (особенно в космологии), для которых тензор энергии-импульса материи не следует из вариационного принципа. В настоящем разделе мы определим тензор энергии-импульса материи $T^{\alpha\beta}_{\text{м}}$, которая рассматривается, как сплошная среда, и изучим некоторые из его свойств. При этом мы не будем опираться на вариационный принцип.

Пусть пространство-время $(\mathbb{M}, g)$ топологически тривиально $\mathbb{M} \approx \mathbb{R}^{1,3}$ и покрыто одной картой. Мы предполагаем, что координаты $\{x^\alpha\} = \{x^0, x^\mu\}$ выбраны таким образом, что координата x^0 является временем, т.е. $g_{00} > 0$. Кроме того, мы считаем, что все сечения $x^0 = \mathsf{const}$ – пространственноподобны.

Можно привести ряд физических аргументов [**?**] в пользу того, что тензор энергии-импульса материи, которая рассматривается как сплошная среда, имеет вид

$$T^{\alpha\beta}_{\text{м}} = (\mathcal{E} + \mathcal{P})u^\alpha u^\beta - \mathcal{P} g^{\alpha\beta}, \tag{21.117}$$

где $\mathcal{E}(x)$ и $\mathcal{P}(x)$ – плотность энергии и давление материи в точке $x \in \mathbb{M}$, и

$$u^\alpha := dx^\alpha/ds, \qquad ds := \sqrt{|ds^2|}, \qquad ds^2 := g_{\alpha\beta}dx^\alpha dx^\beta,$$

– четырехмерная скорость материи в точке $x \in \mathbb{M}$, которая удовлетворяет тождеству $u^\alpha u_\alpha = 1$. Здесь мы предполагаем, что каждая точка материи движется вдоль времениподобной мировой линии $x^\alpha(s)$ в будущее, $u^0 > 0$. Ясно, что мировые линии точек материи – это интегральные кривые векторного поля u.

Поскольку u^α и $g^{\alpha\beta}$ являются соответственно компонентами вектора и тензора относительно преобразований координат, то мы считаем, что плотность энергии $\mathcal{E}$ и давление материи $\mathcal{P}$ являются достаточно гладкими скалярными полями на пространстве-времени $\mathbb{M}$. В этом случае правая часть равенства (21.117) представляет собой ковариантный симметричный тензор второго ранга. Для обычной (наблюдаемой) материи плотность энергии предполагается положительной, $\mathcal{E} > 0$, а давление – неотрицательным, $\mathcal{P} \geq 0$[1].

Мы не обсуждаем физических аргументов, приводящих к тензору энергии-импульса (21.117), отсылая читателя к монографии [**?**]. В настоящем разделе мы рассматриваем, в основном, математические свойства данного определения.

Замечание. В гидро- и газодинамике все уравнения определены так, что в них входит не сама энергия и давление, а только их градиенты. Это означает, что энергия и давление определены с точностью до добавления произвольной постоянной. В общей теории относительности ситуация отличается принципиально, т.к. уравнения меняются, если к $\mathcal{E}$ или $\mathcal{P}$ добавить постоянную. В частности, наличие космологической постоянной можно интерпретировать как среду с постоянной плотностью энергии $\mathcal{E} = \mathsf{const}$ и постоянным давлением $\mathcal{P} = -\mathcal{E}$. Если $\mathcal{E} > 0$, то давление отрицательно, $\mathcal{P} < 0$. Поэтому космологическую постоянную можно интерпретировать, как некоторую среду, заполняющую все пространство-время со свойствами обыкновенной резины. □

Из общих физических представлений следует, что след тензора энергии-импульса для обычной материи должен быть неотрицательным [**?**]

$$T^\alpha_{\text{м}\,\alpha} = \mathcal{E} - 3\mathcal{P} \geq 0. \tag{21.118}$$

Этим свойством обладает, в частности, тензор энергии-импульса для произвольного распределения точечных частиц (21.91). Отсюда вытекает ограничение на давление

$$\mathcal{P} \leq \frac{\mathcal{E}}{3}. \tag{21.119}$$

Поскольку давление, по предположению, неотрицательно, то с учетом равенства (21.119) существует два крайних случая. Если материя, которой заполнена вселенная, настолько разрежена, что давление можно считать равным нулю, то говорят, что материя пылевидна. Максимальное возможное давление, $\mathcal{P} = \mathcal{E}/3$, соответствует газу ультрарелятивистских частиц, скорости которых близки к скорости света (см.

[1] Давление, в принципе, может быть отрицательным. Примером является резина. Для нее увеличение объема по сравнению с состоянием равновесия приводит к увеличению давления. Поскольку выбор точки отсчета давления в классической механике сплошных сред является условным, то нельзя утверждать, что ограничение $\mathcal{P} \geq 0$ обосновано с физической точки зрения.

раздел 21.8.3). В этом случае говорят, что вселенная заполнена газом излучения или, просто, излучением.

$$\begin{aligned} &\mathcal{P} = 0, && \text{– } \textit{пыль}, \\ &\mathcal{P} = \frac{\mathcal{E}}{3}, && \text{– } \textit{излучение}. \end{aligned}$$

Пример 21.11.1 (Нерелятивистская гидродинамика). Рассмотрим пространство Минковского $\mathbb{R}^{1,3}$ в декартовой системе координат с метрикой Лоренца $\eta_{\alpha\beta} = \mathsf{diag}\,(+---)$. Пусть пространство-время заполнено идеальной жидкостью. Течение жидкости описывается плотностью ρ, давлением $\mathcal{P}$ и трехмерной скоростью u^μ, $\mu = 1, 2, 3$. Тензор энергии-импульса идеальной жидкости, по-определению, имеет вид

$$T^{\alpha\beta}_{\text{м}} = (\mathcal{E} + \mathcal{P})u^\alpha u^\beta - \mathcal{P}\eta^{\alpha\beta}. \tag{21.120}$$

Покажем, что уравнения движения нерелятивистской идеальной жидкости (если не считать уравнения состояния) следуют из закона сохранения четырехмерного тензора энергии-импульса, $\partial_\beta T^{\alpha\beta}_{\text{м}} = 0$.

В нерелятивистском приближении мы считаем, что пространственные компоненты скорости малы: $u^0 \simeq 1$, $u^\mu \ll 1$, где $\{u^\alpha\} = \{u^0, u^\mu\}$, $\alpha = 0, 1, 2, 3$. Кроме того, давление мало, $\mathcal{P} \ll \mathcal{E}$, и в нулевом приближении плотность энергии совпадает с плотностью жидкости, $\mathcal{E} \simeq \rho$. Тогда в низшем приближении компоненты тензора энергии-импульса равны:

$$\begin{aligned} T^{00}_{\text{м}} &= (\mathcal{E} + \mathcal{P})u^0 u^0 - \mathcal{P} && \simeq \rho, \\ T^{0\mu}_{\text{м}} = T^{\mu 0}_{\text{м}} &= (\mathcal{E} + \mathcal{P})u^0 u^\mu && \simeq \rho u^\mu, \\ T^{\mu\nu}_{\text{м}} &= (\mathcal{E} + \mathcal{P})u^\mu u^\nu - \mathcal{P}\eta^{\mu\nu} && \simeq \rho u^\mu u^\nu - \mathcal{P}\eta^{\mu\nu}. \end{aligned} \tag{21.121}$$

Рассмотрим закон сохранения энергии-импульса $\partial_\beta T^{\alpha\beta}_{\text{м}} = 0$. Нулевая компонента этого равенства в главном приближении имеет вид

$$\partial_\beta T^{0\beta}_{\text{м}} = \partial_0 T^{00}_{\text{м}} + \partial_\mu T^{0\mu}_{\text{м}} \simeq \partial_0 \rho + \partial_\mu(\rho u^\mu) = 0. \tag{21.122}$$

Полученное уравнение совпадает с *уравнением непрерывности*. Пространственные компоненты закона сохранения энергии-импульса в главном приближении приводят к равенству

$$\partial_\beta T^{\mu\beta}_{\text{м}} = \partial_0 T^{\mu 0}_{\text{м}} + \partial_\nu T^{\mu\nu}_{\text{м}} \simeq \rho\partial_0 u^\mu + u^\mu\left[\partial_0\rho + \partial_\nu(\rho u^\nu)\right] + \rho u^\nu \partial_\nu v^\mu - \partial^\mu \mathcal{P} = 0,$$

что, с учетом уравнения непрерывности (21.122), дает *уравнение Эйлера*

$$\partial_0 u^\mu + u^\nu \partial_\nu u^\mu = \frac{1}{\rho}\eta^{\mu\nu}\partial_\nu \mathcal{P}. \tag{21.123}$$

Если дополнить уравнение непрерывности и уравнение Эйлера уравнением состояния идеальной жидкости $\mathcal{P} = \mathcal{P}(\rho)$, связывающим давление и плотность, то получим полную систему уравнений для идеальной жидкости. Таким образом, уравнения движения нерелятивистской идеальной жидкости следуют из закона сохранения четырехмерного тензора энергии-импульса (21.120), дополненного уравнениям состояния.

Эта же система уравнений (21.122), (21.123) описывает движение идеального газа. Разница заключается только в уравнении состояния. Для идеального газа уравнение состояния имеет вид

$$p = \frac{\rho}{\mu}RT, \tag{21.124}$$

где μ, R и T есть, соответственно, молекулярный вес, универсальная газовая постоянная и абсолютная температура. При постоянной температуре $T = \mathsf{const}$ давление идеального газа прямо пропорционально плотности. □

Наличие в пространстве-времени метрики и времениподобного векторного поля u, $u^2 = 1$, позволяет определить проекционные операторы (5.25):

$$\Pi_\alpha^{\text{L}\,\beta} := u_\alpha u^\beta, \qquad \Pi_\alpha^{\text{T}\,\beta} := \delta_\alpha^\beta - u_\alpha u^\beta.$$

В каждой точке $x \in \mathbb{M}$ эти операторы проектируют тензорные поля, соответственно, на направление вектора скорости u и перпендикулярную гиперплоскость в касательном пространстве $\mathbb{T}_x(\mathbb{M})$. Например, проекция метрики имеет вид

$$g^{\text{L}}_{\alpha\beta} := \Pi_\alpha^{\text{L}\,\gamma}\Pi_\beta^{\text{L}\,\delta} g_{\gamma\delta} = u_\alpha u_\beta,$$
$$g^{\text{T}}_{\alpha\beta} := \Pi_\alpha^{\text{T}\,\gamma}\Pi_\beta^{\text{T}\,\delta} g_{\gamma\delta} = g_{\alpha\beta} - u_\alpha u_\beta.$$

Ясно также, что

$$u^{\text{L}\alpha} := u^\beta \Pi_\beta^{\text{L}\,\alpha} = u^\alpha, \qquad u^{\text{T}} := u^\beta \Pi_\beta^{\text{T}\,\alpha} = 0.$$

Поскольку тензор энергии-импульса сплошной среды (21.117) не был получен из вариационного принципа, то на него необходимо наложить дополнительное условие

$$\nabla_\beta T_{\text{м}}^{\beta\alpha} = 0, \tag{21.125}$$

которое является условием совместности уравнений Эйнштейна. Более подробно

$$(\mathcal{E} + \mathcal{P})u^\beta \nabla_\beta u^\alpha + u^\alpha \nabla_\beta \Big[(\mathcal{E} + \mathcal{P})u^\beta\Big] - g^{\alpha\beta}\partial_\beta \mathcal{P} = 0,$$

где мы воспользовались условием метричности связности Леви-Чивиты $\nabla_\beta g^{\gamma\alpha} = 0$. Проекции этого уравнения на вектор u и перпендикулярную гиперплоскость имеют следующий вид

$$(\mathcal{E} + \mathcal{P})\nabla_\alpha u^\alpha + u^\alpha \partial_\alpha \mathcal{E} = 0, \tag{21.126}$$
$$(\mathcal{E} + \mathcal{P})u^\beta \nabla_\beta u^\alpha - (g^{\alpha\beta} - u^\alpha u^\beta)\partial_\beta \mathcal{P} = 0, \tag{21.127}$$

где мы воспользовались уравнением $u_\alpha \nabla_\beta u^\alpha = 0$, которое следует из условия $u^2 = 0$ после дифференцирования. Легко проверить, что свертка уравнений (21.127) с ковектором u_α тождественно обращается в нуль. Следовательно, только четыре уравнения из (21.126), (21.127) являются независимыми, и они эквивалентны условию ковариантного сохранения тензора энергии-импульса $\nabla_\beta T_{\text{м}}^{\beta\alpha} = 0$.

Уравнение (21.126) является ковариантным обобщением уравнения непрерывности (21.122), а уравнение (21.127) – ковариантным обобщением уравнения Эйлера (21.123). Эти уравнения представляют собой систему уравнений *релятивистской гидродинамики.*

Система уравнений (21.126), (21.127) вместе с уравнениями Эйнштейна не образует полной системы уравнений релятивистской гидродинамики. Ее необходимо дополнить уравнением состояния. Широкий класс моделей описывается уравнением состояния $\mathcal{P} = \mathcal{P}(\mathcal{E})$, связывающим давление с плотностью энергии в каждой точке пространства-времени. Такие жидкости называются *баротропными.*

Второе слагаемое в уравнении Эйлера (21.127) после опускания индекса имеет вид

$$\Pi_\alpha^{\text{T}\,\beta}\partial_\beta \mathcal{P} = (\delta_\alpha^\beta - u_\alpha u^\beta)\partial_\beta \mathcal{P}.$$

Если оно равно нулю, т.е. градиент давления параллелен вектору скорости, то уравнение Эйлера упрощается $u^\beta \nabla_\beta u^\alpha = 0$. Это есть уравнение экстремалей. В этом случае точки жидкости движутся так же, как и точечные частицы.

Для пылевидной материи давление равно нулю и система уравнений релятивистской гидродинамики существенно упрощается:

$$\nabla_\alpha(\mathcal{E}u^\alpha) = 0, \qquad u^\beta \nabla_\beta u^\alpha = 0. \tag{21.128}$$

Мы видим, что пылевидная материя движется вдоль экстремалей, как множество точечных частиц.

При анализе общих свойств решений уравнений Эйнштейна на тензор энергии-импульса материи накладываются определенные *энергетические условия.* Перечислим три наиболее распространенных.

Определение. Говорят, что источники (поля материи) удовлетворяют:

1) *слабому энергетическому условию*, если для любого времениподобного векторного поля X выполнено неравенство

$$T_{\text{м}\alpha\beta} X^\alpha X^\beta \geq 0; \tag{21.129}$$

2) *сильному энергетическому условию*, если для любого времениподобного векторного поля X выполнено неравенство

$$\rho_{\alpha\beta} X^\alpha X^\beta \geq 0, \qquad \rho_{\alpha\beta} := T_{\text{м}\alpha\beta} - \frac{1}{n-2} g_{\alpha\beta} T_{\text{м}\gamma}{}^\gamma, \tag{21.130}$$

где n – размерность пространства-времени. Это условие называется также *условием положительности Риччи* (см. уравнение (21.3);

3) *доминантному энергетическому условию*, если для произвольного времениподобного векторного поля X, направленного в будущее, векторное поле с компонентами $T_{\text{м}\beta}{}^\alpha X^\beta$ также времениподобно и направлено в будущее.

□

Замечание. В данном определении векторное поле X времениподобно, т.е. $X^2 := X^\alpha X_\alpha > 0$. Поэтому, не ограничивая общности, можно считать, что оно нормировано на единицу, $X^2 = 1$. Кроме того, всегда можно выбрать такую систему координат, в которой произвольное времениподобное векторное поле имеет только одну нетривиальную компоненту $X = (1, 0, \ldots, 0)$. Поэтому слабое и сильное энергетические условия можно переформулировать. В произвольной системе координат, в которой координата x^0 является временем, должны выполняться равенства: $T_{\text{м}00} \geq 0$ или $\rho_{00} \geq 0$, соответственно. □

Ясно, что если выполнено доминантное энергетическое условие, то $T_{\text{м}\alpha\beta} X^\alpha X^\beta > 0$. Следовательно из 3) следует 1). Кроме того, если след тензора энергии-импульса материи положителен, как, например, для точечных частиц (см. раздел 21.8.3), то из сильного энергетического условия вытекает слабое, причем получается строгое неравенство $T_{\text{м}\alpha\beta} X^\alpha X^\beta > 0$. Если же след тензора энергии-импульса материи отрицателен, то, наоборот, слабое энергетическое условие сильнее, чем сильное энергетическое условие.

Предложение 21.11.1. *Если выполнены условия $\mathcal{E} > 0$, $\mathcal{P} \geq 0$, то тензор энергии-импульса сплошной среды удовлетворяет слабому и сильному энергетическим условиям.*

Доказательство. Зафиксируем систему координат, в которой времениподобное векторное поле имеет вид $X = (1, 0, 0, 0)$. Тогда слабое энергетическое условие эквивалентно неравенству

$$T_{\text{м}00} = \mathcal{E}u_0^2 + \mathcal{P}(u_0^2 - g_{00}) \geq 0.$$

Первое слагаемое в этом выражении, очевидно, неотрицательно (в принципе оно может обратиться в нуль, т.к. условия $u^0 > 0$ недостаточно для положительности нулевой ковариантной компоненты скорости u_0). Поскольку

$$u_0 = g_{00}u^0 + g_{0\mu}u^\mu,$$

то второе слагаемое принимает вид

$$\mathcal{P}\left(g_{00}^2(u^0)^2 + 2g_{00}g_{0\mu}u^0u^\mu + (g_{0\mu}u^\mu)^2 - g_{00}\right) = -\mathcal{P}g_{00}\left(g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}\right)u^\mu u^\nu,$$

где мы воспользовались тождеством $u^2 = 1$. Согласно теореме 5.2.1 на лоренцевом многообразии матрица в скобках в правой части равенства отрицательно определена, и поэтому слабое энергетическое условие выполнено.

Для тензора энергии-импульса сплошной среды справедливо равенство

$$\rho_{\alpha\beta} = (\mathcal{E} + \mathcal{P})u_\alpha u_\beta - \mathcal{P}g_{\alpha\beta} - \frac{1}{2}(\mathcal{E} - 3\mathcal{P})g_{\alpha\beta}.$$

Отсюда следует выражение для нулевой компоненты

$$\rho_{00} = \mathcal{E}(u_0^2 - g_{00}) + \mathcal{P}(u_0^2 - g_{00}) + \frac{1}{2}(\mathcal{E} + 3\mathcal{P})g_{00}.$$

Первые два слагаемых неотрицательны в силу предыдущих рассуждений. Последнее слагаемое положительно, т.к. координата x^0 является временем и, следовательно, $g_{00} > 0$. □

Замечание. В доказательстве предложения мы не использовали неравенство $\mathcal{P} \leq \mathcal{E}/3$, обеспечивающее неотрицательность следа тензора энергии-импульса. □

Предложение 21.11.2. *Если выполнено условие $|\mathcal{P}| < |\mathcal{E}|$, то тензор энергии-импульса сплошной среды удовлетворяет условию энергодоминантности.*

Доказательство. В системе координат, где $X = (1, 0, 0, 0)$, условие энергодоминантности имеет вид

$$T_{\text{м}0}{}^\alpha T_{\text{м}0\alpha} > 0.$$

Подставляя в это неравенство выражение для тензора энергии-импульса (21.117), получим соотношение

$$(\mathcal{E}^2 - \mathcal{P}^2)u_0^2 + \mathcal{P}^2 g_{00} > 0.$$

Очевидно, что неравенство $|\mathcal{P}| < |\mathcal{E}|$ достаточно для выполнения условия энергодоминантности. □

21.11.1 Акустические фононы в нерелятивистской гидродинамике

В физике всегда большой интерес вызывали аналогии между явлениями из разных областей. В настоящем разделе мы покажем, что некоторые явления в нерелятивистской гидродинамике и общей теории относительности описываются уравнениями, которые обладают рядом одинаковых свойств.

Уравнения движения для акустических фононов следуют из классических нерелятивистских уравнений гидродинамики следующим образом. Рассмотрим некоторое точное решение уравнений гидродинамики. Тогда уравнение для малых возбуждений (фононов) вблизи этого решения сводится к уравнению Даламбера с нетривиальной эффективной четырехмерной метрикой лоренцевой сигнатуры, в которой роль скорости света играет скорость звука в жидкости. Отличие от общей теории относительности сводится к тому, что эффективная метрика для фононов определяется уравнениями гидродинамики, а не уравнениями Эйнштейна. Тем не менее уравнение для фононов задается нетривиальной четырехмерной метрикой, для которой тензор кривизны отличен от нуля. Другими словами, фононы двигаются на многообразии с нетривиальной геометрией. При этом возможно возникновение горизонтов, когда скорость течения жидкости превышает скорость звука, и, следовательно, образование акустических аналогов черных дыр.

В настоящем разделе мы следуем выводу уравнения для фононов, предложенному в [**?**, **?**] (см. также [**?**]).

Рассмотрим 4-мерное галилеево пространство-время с декартовой системой координат $\{x^\alpha\}$, $\alpha = 0, 1, 2, 3$, которые мы будем обозначать индексами из начала греческого алфавита $\alpha, \beta, \ldots$. Координату $x^0 \in \mathbb{R}$ мы отождествляем с временем, $x^0 = t$. Пространственные координаты $\{x^\mu\} \in \mathbb{R}^3$ мы будем обозначать индексами из середины греческого алфавита $\mu, \nu, \ldots$. Жидкость без вязкости называется *идеальной* и описывается плотностью $\rho(x)$, давлением $\mathcal{P}(x)$ и вектором скорости $\boldsymbol{u} = \{u^\mu(x)\}$, где $x = \{x^\alpha\}$. Движение идеальной жидкости или идеального газа в пространстве определяется следующей замкнутой системой из пяти нелинейных уравнений для пяти переменных (см., например, [**?**])

$$\rho\dot{\boldsymbol{u}} + \rho(\boldsymbol{u}\nabla)\boldsymbol{u} = -\nabla\mathcal{P} + \boldsymbol{f}, \tag{21.131}$$

$$\dot{\rho} + \operatorname{div}(\rho\boldsymbol{u}) = 0, \tag{21.132}$$

$$\mathcal{P} = \mathcal{P}(\rho), \tag{21.133}$$

где точка обозначает дифференцирование по времени и ∇ – градиент. Уравнение (21.131) называется уравнением Эйлера (21.123) и представляет собой второй закон Ньютона для элемента объема жидкости. Здесь $\boldsymbol{f}(x)$ – плотность внешних сил. Например, в гравитационном поле $\boldsymbol{f} = -\rho\nabla\varphi$, где $\varphi(x)$ – потенциал гравитационного поля. В дальнейшем мы ограничимся только этим случаем. Уравнение (21.132) является уравнением непрерывности (21.122). Последнее уравнение (21.133) является уравнением состояния жидкости, которое характеризует саму жидкость и считается заданным. Здесь мы предполагаем, что давление жидкости зависит только от ее плотности, т.е. она является баротропной.

Для несжимаемой жидкости $\rho = \mathsf{const}$ и уравнение непрерывности принимает вид $\operatorname{div}\boldsymbol{u} = 0$.

Уравнения (21.131)–(21.133) записаны в стандартном для гидродинамики виде, где не делается различие между верхними и нижними индексами.

Если уравнение состояния задано, то давление $\mathcal{P}$ (или плотность ρ) можно исключить из системы уравнений движения. Для этого заметим, что

$$\nabla\mathcal{P} = \frac{d\mathcal{P}}{d\rho}\nabla\rho = c^2\nabla\rho,$$

где введена скорость звука $c(\rho)$

$$c^2 := \frac{d\mathcal{P}}{d\rho}. \tag{21.134}$$

Для обычных жидкостей с увеличением плотности давление увеличивается, и поэтому $c^2 > 0$. Тогда уравнение Эйлера можно переписать в виде

$$\dot{\boldsymbol{u}} + (\boldsymbol{u}\nabla)\boldsymbol{u} = -c^2\frac{\nabla\rho}{\rho} - \nabla\varphi, \tag{21.135}$$

где скорость звука $c = c(\rho)$ рассматривается, как заданная функция от плотности жидкости. Уравнение Эйлера (21.135) вместе с уравнением непрерывности (21.132) однозначно описывают движение такой жидкости при заданных начальных и граничных условиях.

Воспользовавшись тождеством

$$\frac{1}{2}\nabla\boldsymbol{u}^2 = [\boldsymbol{u},\, \mathsf{rot}\,\boldsymbol{u}] + (\boldsymbol{u}\nabla)\boldsymbol{u},$$

где квадратные скобки обозначают векторное произведение, уравнение Эйлера (21.135) можно переписать в виде

$$\dot{\boldsymbol{u}} - [\boldsymbol{u},\, \mathsf{rot}\,\boldsymbol{u}] = -\nabla\left(h + \frac{\boldsymbol{u}^2}{2} + \varphi\right),$$

где введена *энтальпия* жидкости

$$h(\mathcal{P}) = \int_0^{\mathcal{P}} \frac{d\mathcal{P}'}{\rho(\mathcal{P}')}.$$

Предположим, что движение жидкости является безвихревым:

$$\mathsf{rot}\,\boldsymbol{u} = 0 \quad\Leftrightarrow\quad \partial_\mu u_\nu - \partial_\nu u_\mu = 0. \tag{21.136}$$

На геометрическом языке это означает, что 1-форма $dx^\mu u_\mu$ является замкнутой на пространственном сечении $x^0 = \mathsf{const}$. Тогда локально (во всем $\mathbb{R}^3$) существует потенциальное поле $\psi(x)$ (потенциал) такое, что

$$u_\mu = -\partial_\mu\psi. \tag{21.137}$$

Для безвихревой жидкости уравнение Эйлера эквивалентно *уравнению Бернулли*

$$-\dot{\psi} + h + \frac{(\nabla\psi)^2}{2} + \varphi = F(t), \tag{21.138}$$

где $F(t)$ – произвольная функция времени. Поскольку потенциал ψ определен с точностью до добавления произвольной функции времени, то, без ограничения общности, положим $F(t) = 0$.

Теперь можно приступить к изучению фононов в нерелятивистской гидродинамике. Допустим, что нам известно точное решение уравнений гидродинамики $\rho_0(x)$, $\mathcal{P}_0(x)$ и $\boldsymbol{u}_0(x) = -\nabla\psi_0$. Получим уравнение, описывающее распространение акустических возбуждений (фононов) вблизи этого решения. Пусть

$$\begin{aligned} \rho &\approx \rho_0 + \epsilon\rho_1, \\ \mathcal{P} &\approx \mathcal{P}_0 + \epsilon\mathcal{P}_1, \\ u^\mu &\approx u_0^\mu + \epsilon u_1^\mu, \\ \psi &\approx \psi_0 + \epsilon\psi_1. \end{aligned} \tag{21.139}$$

где $\epsilon \ll 1$ – малый параметр разложения. При этом мы считаем внешние силы заданными $\varphi = \varphi_0$.

В дальнейшем мы будем использовать обозначения, принятые в дифференциальной геометрии, и различать верхние и нижние индексы. Пространственные индексы в декартовой системе координат поднимаются и опускаются с помощью метрики $\eta_{\mu\nu} = -\delta_{\mu\nu}$ и ее обратной, которая отличается от евклидовой метрики знаком. В наших обозначениях $u_\mu = \partial_\mu\psi$, $u^\mu = \eta^{\mu\nu}\partial_\nu\psi = -\partial_\mu\psi$.

Уравнение Бернулли (21.138) в нулевом и первом порядке по ϵ имеет вид

$$\epsilon^0: \quad -\partial_0\psi_0 + h_0 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\psi_0\partial_\nu\psi_0 + \varphi_0 = 0, \tag{21.140}$$

$$\epsilon^1: \quad -\partial_0\psi_1 + \frac{\mathcal{P}_1}{\rho_0} - u_0^\mu\partial_\mu\psi_1 = 0, \tag{21.141}$$

где учтено разложение для энтальпии

$$h(\mathcal{P}) \approx h(\mathcal{P}_0) + \epsilon\frac{\mathcal{P}_1}{\rho_0}.$$

Учтем, что

$$\rho_1 = \frac{d\rho}{d\mathcal{P}}\mathcal{P}_1 = \frac{\mathcal{P}_1}{c^2},$$

и найдем поправку к плотности ρ_1 из уравнения (21.141)

$$\rho_1 = \frac{\rho_0}{c^2}(\partial_0\psi_1 + u_0^\mu\partial_\mu\psi_1). \tag{21.142}$$

Уравнение непрерывности в нулевом и первом порядке по ϵ имеет вид

$$\epsilon^0: \quad \partial_0\rho_0 + \partial_\mu(\rho_0 u_0^\mu) = 0, \tag{21.143}$$

$$\epsilon^1: \quad \partial_1\rho_1 + \partial_\mu(\rho_0 u_1^\mu + \rho_1 u_0^\mu) = 0. \tag{21.144}$$

Подставим во второе уравнение решение для поправки к плотности (21.142). В результате получим уравнение для поправки к потенциалу скорости:

$$\partial_0\left[\frac{\rho_0}{c^2}(\partial_0\psi_1 + u_0^\mu\partial_\mu\psi_1)\right] + \partial_\mu\left[\rho_0\partial^\mu\psi_1 + \frac{\rho_0}{c^2}u_0^\mu(\partial_0\psi_1 + u_0^\nu\partial_\nu\psi_1)\right] = 0. \tag{21.145}$$

Это волновое уравнение для $\psi_1(x)$ полностью определяет распространение акустических колебаний в движущейся жидкости, описываемой плотностью $\rho_0(x)$ и полем скоростей $u_0^\mu(x)$, с заданным уравнением состояния $c = c(\rho)$. В случае, когда скорость жидкости равна нулю, $u_0 = 0$, а плотность ρ_0 и величина c постоянны, уравнение

(21.145) сводится к уравнению Даламбера, которое описывает распространение акустических возбуждений со скоростью c. Это оправдывает введенное выше понятие скорости звука (21.134).

Если решение для $\psi_1(x)$ известно, то поправка к плотности ρ_1 однозначно определяется формулой (21.142).

Уравнение (21.145), как легко проверить, можно записать в матричных обозначениях

$$\partial_\alpha(f^{\alpha\beta}\partial_\beta\psi_1) = 0,$$

где

$$f^{\alpha\beta} = \frac{\rho_0}{c^2}\begin{pmatrix} 1 & u_0^\nu \\ u_0^\mu & c^2\eta^{\mu\nu} + u_0^\mu u_0^\nu. \end{pmatrix}$$

Введем метрику в галилеевом пространстве

$$g_{\alpha\beta} = \frac{\rho_0}{c}\begin{pmatrix} c^2 + u_0^\mu u_{0\mu} & -u_{0\nu} \\ -u_{0\mu} & \eta_{\mu\nu}, \end{pmatrix} \tag{21.146}$$

и ее обратную

$$g^{\alpha\beta} = \frac{1}{\rho_0 c}\begin{pmatrix} 1 & u_0^\nu \\ u_0^\mu & c^2\eta^{\mu\nu} + u_0^\mu u_0^\nu. \end{pmatrix} \tag{21.147}$$

Интервал, соответствующий метрике (21.146), можно записать в виде

$$ds^2 = \frac{\rho_0}{c}\left[c^2dt^2 + \eta_{\mu\nu}(dx^\mu - u_0^\mu dt)(dx^\nu - u_0^\nu dt)\right] \tag{21.148}$$

"Эффективная" метрика (21.146) имеет лоренцеву сигнатуру $(+---)$, и ее определитель равен

$$g := \det g_{\alpha\beta} = -\frac{\rho_0^4}{c^2}.$$

Сравнение метрики (21.146) с АДМ параметризацией произвольной псевдоримановой метрики (22.5) дает следующее выражение для функций хода и сдвига

$$N = \sqrt{\rho_0 c}, \qquad N^\mu = -\frac{\rho_0 u_0^\mu}{c}.$$

Вернемся к уравнению для фононов. Обратная метрика $g^{\alpha\beta}$ отличается от матрицы $f^{\alpha\beta}$ простым множителем

$$g^{\alpha\beta} = \frac{c}{\rho_0^2} f^{\alpha\beta}.$$

Теперь уравнение для акустических фононов можно переписать в инвариантном относительно общих преобразований координат виде

$$\frac{1}{\sqrt{|g|}}\partial_\alpha\left(\sqrt{|g|}g^{\alpha\beta}\partial_\beta\psi\right) = 0, \tag{21.149}$$

где мы, для простоты обозначений, отбросили индекс у поправки к потенциалу скорости.

Таким образом, распространение фононов в движущейся жидкости описывается инвариантным волновым уравнением в четырехмерном пространстве-времени с

нетривиальной метрикой лоренцевой сигнатуры (21.146). Эта метрика определяется плотностью ρ_0, скоростью звука c и полем скоростей $\boldsymbol{u}_0$, которые удовлетворяют исходным уравнениям (21.140), (21.144). Подчеркнем, что движение самой жидкости происходит в плоском галилеевом пространстве-времени, а распространение акустических возбуждений в этой движущейся жидкости описывается волновым уравнением на псевдоримановом пространстве-времени с нетривиальной "эффективной" метрикой.

Мы считаем, что уравнение состояния жидкости задано, и, следовательно, задана скорость звука в жидкости, как функция плотности. Тогда эффективная метрика определяется четырьмя функциями $\rho_0(x)$ и $\boldsymbol{u}_0(x)$, которые удовлетворяют уравнениям гидродинамики. При постановке задачи Коши для однозначного определения этих функций необходимо задать четыре произвольные функции в качестве начальных условий. В общей теории относительности метрика удовлетворяет уравнениям Эйнштейна и имеет две распространяющихся степени свободы. При постановке задачи Коши для уравнений Эйнштейна также необходимо задать четыре произвольные функции на пространственноподобном сечении: по две на каждую степень свободы, так как уравнения движения второго порядка.

Нулевая компонента метрики g_{00} в (21.146) меняет знак в тех точках пространства-времени, где течение жидкости становится сверхзвуковым: $c^2 = \boldsymbol{u}^2 := -u^\mu u_\mu$. Эти поверхности в пространстве соответствуют горизонтам черных дыр. Действительно, поскольку скорость фононов ограничена скоростью звука в жидкости, то они не могут покинуть область сверхзвукового течения. Следовательно, в быстро текущей жидкости для фононов могут образовываться аналоги черных дыр в общей теории относительности. На рис.21.1 показана качественная картинка образования

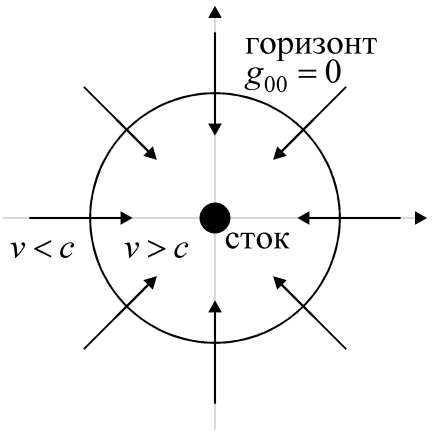


Рис. 21.1: Черная дыра для акустических фононов. Горизонт определяется равенством модуля скорости жидкости $v := \sqrt{-u_0^\mu u_{0\mu}}$ скорости звука c.

черной дыры в жидкости. Представим, что все трехмерное пространство $\mathbb{R}^3$ заполнено несжимаемой жидкостью, и в центре декартовой системы координат находится точечный сток, в который жидкость засасывается. Предположим, что жидкость на бесконечности покоится, и ее движение радиально. Поскольку скорость жидкости стремится к бесконечности при приближении к стоку, то существует сфера некоторого радиуса, на которой скорость жидкости равна скорости звука. Эта сфера называется горизонтом. Если фонон испущен в некоторой точке, лежащей вне горизонта, то он может либо попасть в сток, либо уйти на бесконечность. Если же фонон испущен из точки, лежащей внутри горизонта, то он неминуемо попадет в сток, т.к. его

скорости недостаточно для прохождения через горизонт. Описанная ситуация является аналогом черных дыр в общей теории относительности. Надо только скорость звука заменить на скорость света. Аналогом фононов является электромагнитное излучение – фотоны.

21.12 Выбор системы координат

Уравнения общей теории относительности ковариантны относительно общих преобразований координат. Эту свободу можно использовать для выбора подходящей системы отсчета, которая может упростить уравнения Эйнштейна. В настоящем разделе будут описаны несколько широко распространенных способа фиксирования системы координат.

21.12.1 Сопутствующая система координат

Рассмотрим уравнения Эйнштейна

$$\kappa\left(R_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}R\right)=-\frac{1}{2}T_{\text{м}\alpha\beta} \tag{21.150}$$

для сплошной среды

$$T_{\text{м}}^{\alpha\beta}=(\mathcal{E}+\mathcal{P})u^{\alpha}u^{\beta}-\mathcal{P}g^{\alpha\beta}. \tag{21.151}$$

Для получения замкнутой системы уравнений уравнения Эйнштейна необходимо дополнить законом сохранения (уравнениями релятивистской гидродинамики)

$$\nabla_{\beta}T_{\text{м}}^{\beta\alpha}=0 \tag{21.152}$$

и уравнением состояния среды

$$\mathcal{P}=\mathcal{P}(\mathcal{E}), \tag{21.153}$$

предполагая среду баротропной. Система уравнений (21.150), (21.152) и (21.153) образуют полную систему для неизвестных функций: $g_{\alpha\beta}$, u^{α}, $\mathcal{P}$ и $\mathcal{E}$. Нетрудно проверить, что число уравнений равно числу неизвестных (напомним, что на вектор скорости u наложено условие $u^2=1$ и число его независимых компонент на единицу меньше размерности пространства-времени).

По-построению, все уравнения ковариантны. Поэтому преобразования координат можно использовать для упрощения системы уравнений. Обычно преобразования координат используют для фиксирования части компонент метрики. Однако для системы уравнений (21.150), (21.152) и (21.153) существует другая естественная возможность. Если размерность пространства-времени равна n, то в нашем распоряжение имеется n функций, которых достаточно для фиксирования векторного поля скорости. Тем самым число неизвестных функций уменьшится, и задача упростится. Такой подход часто используется в космологии.

Опишем этот способ задания системы координат. Рассмотрим псевдоримановo многообразие $\mathbb{M}$, $\dim\mathbb{M}=n$, с метрикой g лоренцевой сигнатуры. Пусть на нем задано произвольное достаточно гладкое времениподобное векторное поле $u=u^{\alpha}(x)\partial_{\alpha}$, всюду отличное от нуля, $u^2\neq 0$. Не ограничивая общности, будем считать, что

$$u^2=g_{\alpha\beta}u^{\alpha}u^{\beta}=1.$$

В противном случае можно просто заменить $u^{\alpha}\mapsto u^{\alpha}/u^2$.

Пример 21.12.1. Пусть пространство-время $\mathbb{M}$ заполнено сплошной средой. Тогда задано векторное поле скорости каждой точки среды u, для которого $u^2 = 1$. □

Если на $\mathbb{M}$ задано единичное времениподобное векторное поле, то определены проекционные операторы:

$$\Pi_\alpha^{\text{L}\,\beta} := u_\alpha u^\beta, \qquad \Pi_\alpha^{\text{T}\,\beta} := \delta_\alpha^\beta - u_\alpha u^\beta. \tag{21.154}$$

Произвольное векторное поле $X^\alpha \partial_\alpha$ можно спроектировать на векторное поле u и ортогональное дополнение

$$X^{\text{L}\alpha} = (Xu)u^\alpha, \qquad X^{\text{T}\alpha} = X^\alpha - (Xu)u^\alpha.$$

Аналогично проектируются тензоры произвольного ранга.

Согласно общему рассмотрению в разделе 3.6.5 в окрестности произвольной точки существует такая система координат, в которой все компоненты векторного поля, кроме одной, равны нулю. Для единичного времениподобного векторного поля нетривиальная компонента равна единице

$$\{u^\alpha\} = \{1, 0, \ldots, 0\}\,. \tag{21.155}$$

Определение. Система координат, в которой для единичного времениподобного векторного поля выполнено условие (21.155), называется *сопутствующей векторному полю* u. □

В этой системе координат ковариантная производная векторного поля u равна символам Кристоффеля

$$\nabla_\alpha u^\beta = \partial_\alpha u^\beta + \Gamma_{\alpha\gamma}{}^\beta u^g = \Gamma_{\alpha 0}{}^\beta.$$

В сопутствующей системе координат $\partial_\alpha u^\beta = 0$, поэтому при бесконечно малых преобразованиях координат вариация формы компонент вектора (3.98) и ковектора (3.99) содержит только одно слагаемое. Аналогично преобразуются компоненты произвольного тензора. Отсюда вытекает, что производная Ли от любого тензора Y типа (r, s) вдоль векторного поля u совпадает с производной по времени x^0:

$$\mathsf{L}_u Y^{\alpha_1 \ldots \alpha_r}{}_{\beta_1 \ldots \beta_s} = \partial_0 Y^{\alpha_1 \ldots \alpha_r}{}_{\beta_1 \ldots \beta_s}$$

и не зависит от связности, как, впрочем, любая производная Ли.

С физической точки зрения сопутствующую систему координат можно представить следующим образом. Допустим, что некоторая среда заполняет все пространство-время. Тогда с каждой точкой среды связана мировая линия $x(s)$ (линия тока). Мы предполагаем, что касательные векторы к мировым линиям образуют достаточно гладкое времениподобное векторное поле (вектор скорости)

$$u = u^\alpha \partial_\alpha, \qquad u^\alpha = \frac{dx^\alpha}{ds}, \qquad ds = \sqrt{g_{\alpha\beta} dx^\alpha dx^\beta},$$

на многообразии $\mathbb{M}$. Выберем произвольное сечение $\mathbb{S}$, которое пересекает все линии тока один раз, и зададим произвольную систему координат x^μ, $\mu = 1, \ldots, n-1$, на $\mathbb{S}$. Это сечение совсем не обязано быть пространственноподобным. Тогда сопутствующими координатами произвольной точки $y \in \mathbb{M}$ является набор чисел $\{x^0 := s, x^\mu\}$, где x^μ – координаты точки пересечения поверхности $\mathbb{S}$ с кривой $x(s)$, проходящей через точку y.

Замечание. В предыдущем разделе мы установили, что пылевидная материя движется вдоль экстремалей (21.128). Это значит, что в общем случае при наличии давления или других негравитационных сил линии тока среды отличаются от экстремалей. □

Если производная Ли от некоторого тензора вдоль векторного поля скорости u равна нулю, то в сопутствующей системе координат компоненты этого тензора могут зависеть только от координат x^μ, $\mu = 1, \ldots, n-1$ на сечении $\mathbb{S}$. Это значит, что соответствующий тензор жестко связан с движущейся средой и движется вместе с ней.

Сопутствующая векторному полю система координат определена неоднозначно. Действительно, совершим преобразование координат $x^\alpha \mapsto x^{\alpha'}(x)$. Тогда компоненты скорости преобразуются по тензорному закону:

$$u^\alpha \mapsto u^{\alpha'} := \frac{\partial x^{\alpha'}}{\partial x^\alpha} u^\alpha.$$

Если до и после преобразования координат система координат является сопутствующей, то функции преобразования координат должны удовлетворять следующей системе уравнений

$$1 = \frac{\partial x^{0'}}{\partial x^0}, \qquad 0 = \frac{\partial x^{\mu'}}{\partial x^0}.$$

Общее решение данной системы уравнений имеет вид

$$x^0 \mapsto x^0 + f(\boldsymbol{x}), \qquad x^\mu \mapsto x^{\mu'} = x^\mu + f^\mu(\boldsymbol{x}), \tag{21.156}$$

где f, f^μ – n произвольных функций координат на сечении $\mathbb{S} \subset \mathbb{M}$ и $\boldsymbol{x} = \{x^\mu\}$. Функция f соответствует произволу в выборе сечения $x^0 = \mathsf{const}$, и функции f^μ – свободе в выборе координат x^μ на данных сечениях.

Таким образом мы устранили $n-1$ неизвестную функцию в полной системе уравнений (21.150), (21.152) и (21.153). В этой системе координат тензор энергии-импульса принимает вид

$$T^{00}_{\text{м}} = (\mathcal{E} + \mathcal{P}) - \mathcal{P} g^{00}, \qquad T^{0\mu}_{\text{м}} = -\mathcal{P} g^{0\mu}, \qquad T^{\mu\nu}_{\text{м}} = -\mathcal{P} g^{\mu\nu}.$$

В общем случае ни он, ни тензор энергии-импульса с одним опущенным индексом индексом $T^\alpha_{\text{м}\beta}$ не будут диагональны.

Если задано единичное времениподобное векторное поле u, то в каждой точке пространства-времени $x \in \mathbb{M}$ в касательном пространстве $\mathbb{T}_x(\mathbb{M})$ его можно дополнить $n-1$ линейно независимыми векторами e_μ, $\mu = 1, \ldots, n-1$, которые перпендикулярны вектору u. Тогда совокупность векторов $\{u, e_\mu\}$ образует в каждой точке репер. Ясно, что векторы e_μ пространственноподобны, и их можно выбрать достаточно гладкими. Тогда они задают $n-1$ мерное распределение векторных полей на $\mathbb{M}$ (см. раздел 3.11). Согласно теореме Фробениуса для этого распределения существуют интегральные подмногообразия тогда и только тогда, когда векторные поля e_μ находятся в инволюции. В общем случае это не так (это зависит от метрики). Отсюда следует, что остаточного произвола в выборе сопутствующей системы координат (21.156) недостаточно для того, чтобы выбрать секущую поверхность $\mathbb{S}$ таким образом, чтобы вектор u был к ней всюду ортогонален.

Теперь введем другое понятие сопутствующей системы координат. Рассмотрим систему локальных координат x^α на $\mathbb{M}$, где координата x^0 является временем, и

все сечения $x^0 = \mathsf{const}$ пространственноподобны. По-предположению на $\mathbb{M}$ задана метрика лоренцевой сигнатуры, для которой мы будем использовать АДМ параметризацию (22.5). Выберем базис n, e_μ в касательном пространстве, состоящий из пространственных векторов $e_\mu = e_\mu{}^\alpha \partial_\alpha = \partial_\mu$ (символ $e_\mu{}^\alpha = \delta_\mu^\alpha$ введен для удобства последующих выкладок), касательных к гиперповерхностям $x^0 = \mathsf{const}$, и векторное поле

$$n = \frac{1}{N}(\partial_0 - N^\mu \partial_\mu), \tag{21.157}$$

где N – функция хода и N^μ – функции сдвига. Это векторное поле перпендикулярно пространственным гиперповерхностям:

$$(n, e_\mu) := n^\alpha e_\mu{}^\beta g_{\alpha\beta} = n^0 e_\mu{}^\nu g_{0\nu} + n^\rho e_\mu{}^\nu g_{\rho\nu} = 0,$$

и имеет единичную длину, $n^2 = 1$.

В общем случае коммутатор векторных полей $[n, e_\mu]$ отличен от нуля. Поэтому пространственные координаты x^μ нельзя дополнить времениподобной координатой $x^{0'}$ так, чтобы вектор n был касателен к соответствующей координатной линии: $n = \partial_{0'}$.

Нормальному вектору соответствует ортонормальная 1-форма

$$n = dx^\alpha n_\alpha = dx^0 N, \tag{21.158}$$

для которой мы будем использовать то же обозначение.

Для векторов и 1-форм справедливы разложения:

$$X^\alpha = X^\perp n^\alpha + \tilde{X}^\mu e_\mu{}^\alpha, \qquad X_\alpha = X_\perp n_\alpha + \tilde{X}_\mu e^\mu{}_\alpha,$$

где

$$X^\perp = X^0 N, \qquad \tilde{X}^\mu = X^0 N^\mu + X^\mu,$$
$$X_\perp = \frac{1}{N}(X_0 - N^\mu X_\mu), \qquad \tilde{X}_\mu = X_\mu.$$

Поскольку $X_0 := X^\alpha g_{\alpha 0}$ и $X_\mu := X^\alpha g_{\alpha\mu}$, то нетрудно проверить, что $X_\perp = X^\perp$ и $\tilde{X}_\mu = \tilde{X}^\nu g_{\nu\mu}$.

Аналогично раскладываются тензоры произвольного ранга.

По-построению, векторное поле n времениподобно и имеет единичную длину. Поэтому его можно (как мы увидим, не всегда) отождествить с полем скоростей материи u.

Определение. Система координат называется *сопутствующей*, если в этой системе координат скорость $\{u^\alpha\}$ каждой точки материи имеет вид

$$\{u^\alpha\} = \left\{\sqrt{g^{00}}, \frac{g^{0\mu}}{\sqrt{g^{00}}}\right\} = \left\{\frac{1}{N}, -\frac{N^\mu}{N}\right\}. \quad \square \tag{21.159}$$

В правой части равенства (21.159), определяющего сопутствующую систему координат, стоят определенные компоненты метрики, которые не образуют компонент вектора. Следовательно, равенство (21.159) нековариантно и действительно фиксирует систему координат.

Название сопутствующая система координат оправдано следующим образом. Поскольку вектор нормали n к пространственному сечению $x^0 = \mathsf{const}$ совпадает с вектором скорости u, то репер n, e_μ привязан к среде и движется вместе с ней.

Если метрика имеет блочно диагональный вид 5.20, т.е. $N^\mu = 0$ и $N = 1$, и, следовательно, выбрана временна́я калибровка (см. следующий раздел), то вектор скорости материи в каждой точке пространства-времени имеет только временну́ю составляющую $\{u^\alpha\} = \{1, 0, 0, 0\}$. В этом случае сопутствующая система координат совпадает с системой координат, сопутствующей векторному полю скорости u, которая была введена в начале раздела. Это значит, что в сопутствующей системе координат каждая точка материи покоится. Другими словами, система координат движется вместе с материей. При этом каждая точка среды движется по времениподобной экстремали (геодезической).

При наличии сил негравитационного происхождения, например, давления, точки материи могут двигаться не по экстремалям, и метрика в общем случае не будет блочно-диагональна. Математически это означает, что в общем случае нельзя одновременно удовлетворить условиям (21.159) и (5.20). Поэтому данные выше определения сопутствующей системы координат неэквивалентны.

Обозначим тензорные индексы по отношению к базису $\{e_a\} = \{n, e_\mu\}$ латинскими буквами $a, b, \ldots$. Тогда они принимают значения $\{a\} = \{\perp, \mu\} = \{\perp, 1, \ldots, n-1\}$. Отличие этого базиса от координатного базиса в касательном пространстве заключается в том, что он неголономен, т.е. в общем случае не существует такой системы координат $y^a(x)$, в которой были бы выполнены условия:

$$\frac{\partial x^\alpha}{\partial y^0} = n^\alpha, \qquad \frac{\partial x^\nu}{\partial y^\mu} = \delta^\nu_\mu$$

(см. раздел 7.9). В этом базисе метрика имеет блочно диагональный вид

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \tag{21.160}$$

поскольку базисные векторы n были выбраны единичными и перпендикулярными к гиперповерхностям. В каждой точке пространства-времени векторы четырехмерной скорости (21.159) в базисе n, ∂_μ имеют только одну отличную от нуля компоненту,

$$\{u^a\} = \{1, 0, 0, 0\}.$$

Это значит, что в рассматриваемом базисе материя покоится, что оправдывает название сопутствующая.

Понятие сопутствующей системы координат полезно, т.к. позволяет исключить из тензора энергии-импульса сплошной среды компоненты скорости u^α, заменив их на компоненты метрики следующим образом. Тензор энергии-импульса материи с одним контравариантным и одним ковариантным индексом в сопутствующей системе координат (21.159), и, следовательно, относительно базиса n, ∂_μ, по-предположению, является диагональным

$$T^a_{\text{м}b} = \begin{pmatrix} \mathcal{E} & 0 & 0 & 0 \\ 0 & -\mathcal{P} & 0 & 0 \\ 0 & 0 & -\mathcal{P} & 0 \\ 0 & 0 & 0 & -\mathcal{P} \end{pmatrix}. \tag{21.161}$$

Тогда в координатном базисе его контравариантные компоненты равны

$$T^{00}_{\text{м}} = \mathcal{E}\frac{1}{N^2}, \qquad T^{0\mu}_{\text{м}} = T^{\mu 0}_{\text{м}} = -\mathcal{E}\frac{N^\mu}{N^2}, \qquad T^{\mu\nu}_{\text{м}} = \mathcal{E}\frac{N^\mu N^\nu}{N^2} - \mathcal{P}\hat{g}^{\mu\nu}. \tag{21.162}$$

Предложение 21.12.1. *Функции $\mathcal{E}$ и $\mathcal{P}$ инвариантны относительно масштабного преобразования (растяжки, гомотетии) пространственных координат:*

$$x^0 \mapsto x^0, \qquad x^\mu \mapsto x'^\mu = k(t)x^\mu, \tag{21.163}$$

где $k(t) \neq 0$ – достаточно гладкая функция времени $x^0 := t$.

Доказательство. Координатный базис преобразуется по-правилам:

$$\begin{aligned} \partial_0 &= \partial'_0 + \dot{k}x^\mu \partial'_\mu, & dx^0 &= dx'^0, \\ \partial_\mu &= k\partial'_\mu, & dx^\mu &= \frac{dx'^\mu}{k} - \frac{x'^\mu \dot{k} dt}{k^2}, \end{aligned}$$

где $\dot{k} = dk/dt$. Тогда в новой системе координат компоненты метрики примут вид

$$\begin{aligned} N' &= N, \\ N'_\mu &= \frac{N_\mu}{k} - \frac{\dot{k}x_\mu}{k^2}, & N'^\mu &= kN^\mu - \dot{k}x^\mu, \\ g'_{\mu\nu} &= \frac{g_{\mu\nu}}{k^2}, & \hat{g}'^{\mu\nu} &= k^2 \hat{g}^{\mu\nu}. \end{aligned}$$

Теперь из формул (21.162) следует, что плотность энергии и давление в сопутствующей системе координат не меняются при масштабном преобразовании (21.163)

$$\mathcal{E}' = \mathcal{E}, \qquad \mathcal{P}' = \mathcal{P}.$$

При этом компоненты вектора скорости преобразуются по-правилам

$$u'^0 = u^0, \qquad u'^\mu = ku^\mu + \dot{k}u^0 x^\mu. \qquad \square$$

Замечание. Инвариантность плотности энергии и давления относительно масштабного преобразования следовало ожидать, т.к. плотность энергии и давление являются скалярными полями и инвариантны относительно любых преобразований координат пространства-времени x^α и, в частности, растяжений (21.163). Нетривиальность проведенного рассмотрения заключается в том, что в правой части равенства (21.159), определяющего сопутствующую систему координат, стоят определенные компоненты метрики, которые в общем случае не совпадают с компонентами никакого вектора. Другими словами, сопутствующая система координат (21.159) определена по крайней мере с точностью до масштабных преобразований (21.163). $\square$

Уравнения релятивистской гидродинамики (21.152), которые были выписаны в начале раздела, можно записать в сопутствующей системе координат (21.159). Прямые вычисления приводят к следующей системе уравнений:

$$\begin{aligned} \dot{\mathcal{E}} - N^\mu \partial_\mu \mathcal{E} - NK(\mathcal{E} + \mathcal{P}) &= 0, \\ \partial_\mu \mathcal{P} + \frac{\partial_\mu N}{N}(\mathcal{E} + \mathcal{P}) &= 0, \end{aligned} \tag{21.164}$$

где точка обозначает дифференцирование по времени $\dot{\mathcal{E}} = \partial\mathcal{E}/\partial t$ и

$$K = \frac{1}{2N}(2\hat{\nabla}^\mu N_\mu - \hat{g}^{\mu\nu}\dot{g}_{\mu\nu}) \tag{21.165}$$

– внешняя скалярная кривизна пространственной гиперповерхности.

21.12.2 Временна́я калибровка

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, на котором задана метрика лоренцевой сигнатуры $g_{\alpha\beta}(x)$, $\operatorname{sign} g_{\alpha\beta} = (+ - \ldots -)$.

Определение. Система координат, в которой метрика имеет блочно диагональный вид (5.20)

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix}, \tag{21.166}$$

где $g_{\mu\nu}$ – отрицательно определенная риманова метрика на пространственноподобных сечениях $x^0 = \mathsf{const}$, называется *временно́й калибровка*. Эту систему координат называют также *синхронной, гауссовой или полугеодезической.* □

В синхронной системе отсчета координата x^0 является временем и явно выделена. Напомним, что греческие буквы из начала алфавита пробегают все значения индексов, $\alpha, \beta, \ldots = 0, 1, \ldots, n-1$, а из середины – только пространственные, $\mu, \nu, \ldots = 1, 2, \ldots, n-1$.

При переходе в синхронную систему отсчета n произвольных функций, параметризующих диффеоморфизмы, используются для фиксирования n компонент метрики:

$$g_{00} = 1, \qquad g_{0\mu} = 0.$$

В АДМ параметризации метрики (см. раздел 22.2) временна́я калибровка соответствует условиям $N = 1$, $N_\mu = 0$.

Замечание. Названия гауссова или полугеодезическая система координат распространены в математической литературе, когда рассматриваются римановы пространства с положительно определенной метрикой. В физической литературе, где преимущественно рассматриваются многообразия с метрикой лоренцевой сигнатуры, чаще употребляют термины временна́я калибровка или синхронная система координат, потому что в этой системе отсчета координата x^0 действительно играет роль наблюдаемого времени. □

Название синхронная система координат для метрики (21.166) оправдана следующим обстоятельством.

Синхронизация часов

В общем случае интервал между двумя близкими событиями имеет вид

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta.$$

Предположим, что координата x^0 является (наблюдаемым) временем, т.е. $g_{00} > 0$, и все сечения $x^0 = \mathsf{const}$ пространственноподобны. Если два события C и D произошли в данной системе координат в одной и той же точке пространства, то они имеют координаты $C = \{x^0_C, x^\mu\}$ и $D = \{x^0_D, x^\mu\}$. При этом данные события разделены интервалом собственного времени

$$\triangle s = \int_{x^0_C}^{x^0_D} dx^0 \sqrt{g_{00}}. \tag{21.167}$$

Этот интеграл равен длине времениподобной кривой

$$x^0 = x^0_C + (x^0_D - x^0_C)t, \qquad x^\mu = \mathsf{const}, \qquad t \in [0, 1].$$

Конечно, в другой системе координат события могут произойти не только в разное время, но и в разных точках пространства.

Таким образом, если два события, произошедшие в одной точке пространства в данной системе координат, разделены наблюдаемым временем $x^0_D - x^0_C$, то они разделены интервалом собственного времени (21.167). При этом нулевая компонента метрики g_{00} определяет различие собственного и наблюдаемого времени для событий, произошедших в одной точке.

Теперь определим понятие одновременности для событий, которые произошли в двух разных, но близких точках пространства в данной фиксированной системе координат. Пусть событие A имеет пространственные координаты x^μ, а событие B – близкие координаты $x^\mu + dx^\mu$. На рис.21.2 показаны временны́е оси, проходящие через точки A и B. Возникает следующий вопрос одновременности. Допустим, что событие A имеет координаты $\{x^0, x^\mu\}$. Какова временна́я координата $x^0 + \triangle x^0$ события, произошедшего в точке B, которое можно назвать одновременным с событием A ?

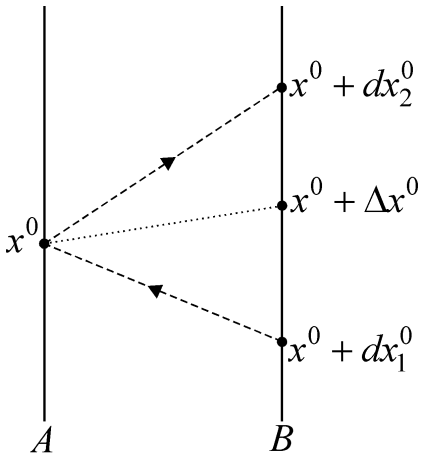


Рис. 21.2: Одновременность близких событий A и B.

Чтобы определить одновременность, испустим свет в точке B в некоторый момент времени $x^0 + dx^0_1$ (величина dx^0_1 отрицательна). Как только свет попадет в точку A, сразу отразим его. Допустим, что свет вернулся в точку B в момент времени $x^0 + dx^0_2$. Поскольку для света $ds^2 = 0$, то изменение наблюдаемого времени в обоих случаях должно удовлетворять уравнению

$$g_{00}\Big(dx^0_{1,2}\Big)^2 + 2g_{0\mu}dx^0_{1,2}dx^\mu + g_{\mu\nu}dx^\mu dx^\nu = 0.$$

Это квадратное уравнение имеет два решения:

$$dx^0_1 = \frac{1}{g_{00}}\left[-g_{0\mu}dx^\mu - \sqrt{(g_{0\mu}g_{0\nu} - g_{\alpha\beta}g_{00})dx^\mu dx^\nu}\right],$$
$$dx^0_2 = \frac{1}{g_{00}}\left[-g_{0\mu}dx^\mu + \sqrt{(g_{0\mu}g_{0\nu} - g_{\alpha\beta}g_{00})dx^\mu dx^\nu}\right].$$

Поскольку мы предположили, что $g_{00} > 0$ и метрика $g_{\mu\nu}$ отрицательно определена, то отсюда вытекает, что $dx^0_2 > 0$, а $dx^0_1 < 0$.

Назовем событие в точке B *одновременным* событию $A = \{x^0, x^\mu\}$, если его временная координата равна $x^0 + \triangle x^0$, где

$$\triangle x^0 := \frac{dx^0_1 + dx^0_2}{2} = -\frac{g_{0\mu}dx^\mu}{g_{00}},$$

т.е. лежит посередине между $x^0 + dx^0_2$ и $x^0 + dx^0_1$.

Таким образом можно синхронизировать часы, расположенные в различных, но близко расположенных точек в пространстве. Этот процесс можно продолжить вдоль произвольной кривой в пространстве. Конечно, данная процедура синхронизации часов зависит от выбора системы координат (нековариантна) и зависит также от выбора кривой, соединяющей две точки в пространстве.

Рассмотрим замкнутую кривую γ в пространстве с началом и концом в точке A. Произведем синхронизацию часов вдоль кривой γ описанным выше способом. Тогда после возвращения в точку A временна́я координата получит приращение

$$\triangle x^0 := -\oint_\gamma \frac{g_{0\mu} dx^\mu}{g_{00}}.$$

Отсюда следует, что синхронизация часов в общем случае невозможна, т.к. приращение $\triangle x^0$ может быть отлично от нуля. Кроме того, если мы хотим синхронизировать часы во всей области пространства-времени $\mathbb{U} \subset \mathbb{M}$, которая покрывается данной системой координат, то равенство $\triangle x^0 = 0$ должно также выполняться для любой замкнутой кривой γ, целиком лежащей в $\mathbb{U}$. Отсюда вытекает

Предложение 21.12.2. *Для того, чтобы в выбранной системе координат x^α, где x^0 – время и все сечения $x^0 = \mathsf{const}$ пространственноподобны, покрывающей некоторую область $\mathbb{U} \subset \mathbb{M}$, можно было синхронизировать часы во всей области $\mathbb{U}$ необходимо и достаточно чтобы $g_{0\mu} = 0$.* □

Вернемся к рассмотрению синхронной системы координат и докажем теорему существования.

Предложение 21.12.3. *Пусть задано пространство-время $\mathbb{M}$ с метрикой лоренцевой сигнатуры. Тогда в некоторой окрестности каждой точки $x \in \mathbb{M}$ существует система координат, в которой метрика имеет блочно диагональный вид (21.166).*

Доказательство. Выберем в многообразии $\mathbb{M}$ произвольную достаточно гладкую пространственноподобную гиперповерхность, содержащую точку $x \in \mathbb{N} \hookrightarrow \mathbb{M}$. Пусть y^μ – некоторая система координат на гиперповерхности в окрестности точки x. Построим на $\mathbb{N}$ векторное поле n, перпендикулярное к гиперповерхности. Через каждую точку $y \in \mathbb{N}$ в направлении n проведем экстремаль в обоих направлениях. Мы уже знаем, что такая экстремаль существует и единственна (см. раздел 17). Поскольку гиперповерхность пространственноподобна, то векторное поле и экстремали времениподобны. Выберем в качестве канонического параметра вдоль каждой экстремали ее длину t таким образом, чтобы гиперповерхность $\mathbb{N}$ задавалась уравнением $t = 0$. Тогда в некоторой окрестности $\mathbb{U}$ гиперповерхности, $\mathbb{U} \subset \mathbb{M}$, будет определена система координат $y = \{y^0 := t, y^\mu\} \in \mathbb{U}$. Это и есть искомая синхронная система координат.

Покажем это. По-построению, координатная кривая $\{y^0 = t, y^\mu = \mathsf{const}\}$, $t \in \mathbb{R}$, является экстремалью. Ее вектор скорости в построенной системе координат имеет одну отличную от нуля компоненту $\dot{y}^\alpha = \delta^\alpha_0$. Поскольку экстремаль удовлетворяет уравнению

$$\ddot{y}^\alpha = -\Gamma_{\beta\gamma}{}^\alpha \dot{y}^\beta \dot{y}^\gamma,$$

то в построенной системе координат на метрику наложены условия $\Gamma_{00}{}^\alpha = 0$. Опустив индекс α, получим уравнения на компоненты метрики:

$$\partial_0 g_{0\alpha} - \frac{1}{2}\partial_\alpha g_{00} = 0. \tag{21.168}$$

Поскольку в качестве параметра вдоль экстремали выбрана ее длина, то касательный вектор ∂_0 имеет единичную длину. Следовательно, в построенной системе координат $g_{00} = 1$. Тогда уравнения (21.168) примут вид $\partial_0 g_{0\alpha} = 0$, т.е. компоненты $g_{0\mu}$ не зависят от времени. Кроме того, вектор скорости, по-построению, перпендикулярен гиперповерхности на $\mathbb{N}$. Это значит, что в начальный момент времени t пространственно-временные компоненты метрики равны нулю, $g_{0\mu} = 0$. Поскольку они не зависят от времени, то это равенство выполнено всюду в $\mathbb{U}$. Тем самым построенная система координат является синхронной. □

Ниже мы докажем обратное утверждение: если метрика имеет блочно диагональный вид (21.166), то координатные линии, соответствующие времени, являются экстремалями. Это значит, что единственный произвол при построении синхронной системы отсчета – это выбор пространственного сечения $\mathbb{N}$, которое может быть произвольно, и выбор пространственных координат на $\mathbb{N}$.

Замечание. Если в качестве исходной гиперповерхности $\mathbb{N}$ для построения координат выбрать времениподобную гиперповерхность, то метрика примет вид

$$g_{\alpha\beta} = \begin{pmatrix} -1 & 0 \\ 0 & g_{\mu\nu} \end{pmatrix},$$

где $g_{\mu\nu}$ – метрика лоренцевой сигнатуры на сечениях $x^0 = \mathsf{const}$. Вообще говоря, построенная система координат не зависит от сигнатуры метрики. Аналогичные системы координат рассматривались еще Гауссом в римановой геометрии. □

Замечание. Построение синхронной системы координат начиналось с выбора гиперповерхности $\mathbb{N} \hookrightarrow \mathbb{M}$, которая имеет размерность $n - 1$. Аналогичные системы координат можно строить, стартуя с подмногообразия $\mathbb{N}$ произвольной размерности, $0 \leq \mathsf{dim}\,\mathbb{N} < n$. Детали построения таких систем координат, которые названы полугеодезическими, можно найти в [?]. В частном случае, когда все экстремали стартуют из одной точки, $\mathsf{dim}\,\mathbb{N} = 0$, получаем нормальные координаты, рассмотренные в разделе 17.9. □

Определение. Сечения пространства-времени в синхронной системе координат $x^0 = t = \mathsf{const}$ образуют семейство гиперповерхностей, которые называются *параллельными* друг другу. □

Для римановых многообразий понятие экстремали и геодезической совпадают. Поэтому построенная система координат иногда называется *полугеодезической*. Мы будем использовать принятый в физике термин временна́я калибровка для метрики.

Обратная метрика во временно́й калибровке имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & g^{\mu\nu} \end{pmatrix}, \tag{21.169}$$

где $g^{\mu\nu}$ – метрика, обратная к $g_{\mu\nu}$. Ясно, что требование блочной диагональности метрики (21.166) эквивалентно требованию блочной диагональности обратной метрики.

Чтобы найти явный вид преобразований координат $x^\alpha \mapsto y^\alpha(x)$, необходимо проинтегрировать уравнения экстремалей с заданными начальными условиями на гиперповерхности $\mathbb{N} \hookrightarrow \mathbb{M}$. Это обычная механическая задача, которую можно решить,

интегрируя уравнение Гамильтона–Якоби (см. раздел 20.1.7). Обрисуем кратко этот подход.

При преобразовании к синхронной системе координат $x^\alpha \mapsto y^\alpha(x)$ временна́я компонента обратной метрики преобразуется по тензорному закону. Следовательно,

$$1 = g^{\alpha\beta}\frac{\partial y^0}{\partial x^\alpha}\frac{\partial y^0}{\partial x^\beta}. \tag{21.170}$$

Это уравнение на функцию $y^0(x)$ совпадает с укороченным уравнением Гамильтона–Якоби (17.48). Поэтому временную координату y^0 можно отождествить с функцией действия точечной частицы. Общий интеграл укороченного уравнения Гамильтона–Якоби зависит от n постоянных. Поскольку функция действия определена с точностью до добавления постоянной, то одна из постоянных интегрирования аддитивна. Поэтому общий интеграл уравнения Гамильтона–Якоби можно представить в виде

$$y^0 = f(x, c) + A(c),$$

где f – некоторая функция от точки многообразия и $n-1$ постоянных интегрирования $c = \{c^\mu\} = (c^1, \ldots, c^{n-1})$, а $A(c)$ – аддитивная постоянная, которую можно рассматривать как некоторую функцию постоянных c. Уравнения $y^0 = \mathsf{const}$ определяют гиперповерхность $\mathbb{N} \hookrightarrow \mathbb{M}$. Локально решение этого уравнения можно представить в виде $x^\alpha = x^\alpha(c)$. То есть постоянные c образуют систему координат на $\mathbb{N}$. Временны́е координатные линии y^0 определяются уравнениями $\partial y^0/\partial c^\mu = 0$ или

$$\frac{\partial f}{\partial c^\mu} = -\frac{\partial A}{\partial c^\mu}.$$

Если зафиксировать точку на гиперповерхности $\mathbb{N}$ (постоянные c), то в правой части данной системы уравнений будут стоять некоторые константы. Соответствующее решение системы $x^\mu = x^\mu(x^0)$ определит кривую, проходящую через данную точку $c \in \mathbb{N}$. Эти кривые являются траекториями частиц и, следовательно, экстремалями.

Из вида обратной метрики (21.169) следует уравнение

$$\delta^{0\gamma} = g^{\alpha\beta}\frac{\partial y^0}{\partial x^\alpha}\frac{\partial y^\gamma}{\partial x^\beta}, \qquad \gamma = 0, \ldots, n-1,$$

которое обобщает уравнение (21.170). Умножив его на матрицу Якоби преобразования координат, получим равенство

$$g^{\alpha\beta}\frac{\partial y^0}{\partial x^\alpha} = \frac{\partial x^\beta}{\partial y^0}. \tag{21.171}$$

Вектор нормали к гиперповерхностям имеет вид

$$n = g^{\alpha\beta}\frac{\partial y^0}{\partial x^\beta}\partial_\alpha.$$

С учетом равенства (21.171) отсюда следует, что вектор нормали касателен к координатным линиям y^0:

$$\frac{\partial}{\partial y^0} = \frac{\partial x^\alpha}{\partial y^0}\partial_\alpha = g^{\alpha\beta}\frac{\partial y^0}{\partial x^\beta}\partial_\alpha.$$

Таким образом, координатные линии y^0 являются экстремалями, перпендикулярными к гиперповерхностям $y^0 = \mathsf{const}$.

Перейдем к вычислению явного вида основных геометрических объектов в синхронной системе координат. Прямые вычисления приводят к следующим выражениям для символов Кристоффеля: (7.22)

$$\begin{aligned}
&\Gamma_{00}{}^{0} = \Gamma_{00}{}^{\mu} = \Gamma_{0\mu}{}^{\nu} = \Gamma_{\mu 0}{}^{\nu} = 0,\\
&\Gamma_{0\mu}{}^{\nu} = \Gamma_{\mu 0}{}^{\nu} = \frac{1}{2} g^{\nu\rho}\partial_0 g_{\mu\rho},\\
&\Gamma_{\mu\nu}{}^{0} = -\frac{1}{2}\partial_0 g_{\mu\nu},\\
&\Gamma_{\mu\nu}{}^{\rho} = \hat{\Gamma}_{\mu\nu}{}^{\rho},
\end{aligned} \tag{21.172}$$

где $\hat{\Gamma}_{\mu\nu}{}^{\rho}$ – символы Кристоффеля на пространственноподобном сечении $x^0 = \mathsf{const}$, построенные только по метрике $g_{\mu\nu}$. В настоящем разделе знак тильды, который мы используем для обозначения геометрических объектов римановой геометрии при нулевых тензорах кручения и неметричности, для простоты, опущен.

Несложные вычисления приводят к следующим выражениям для компонент тензора кривизны со всеми опущенными индексами (7.70)

$$\begin{aligned}
R_{0\mu 0\nu} &= \frac{1}{2}\partial_{00}^2 g_{\mu\nu} - \frac{1}{4} g^{\rho\sigma}\partial_0 g_{\mu\rho}\partial_0 g_{\nu\sigma},\\
R_{0\mu\nu\rho} &= -R_{\mu 0\nu\rho} = \frac{1}{2}(\hat{\nabla}_\nu \partial_0 g_{\mu\rho} - \hat{\nabla}_\rho \partial_0 g_{\mu\nu}),\\
R_{\mu\nu\rho\sigma} &= \hat{R}_{\mu\nu\rho\sigma} + \frac{1}{4}(\partial_0 g_{\mu\rho}\partial_0 g_{\nu\sigma} - \partial_0 g_{\nu\rho}\partial_0 g_{\mu\sigma}),
\end{aligned} \tag{21.173}$$

где $\hat{\nabla}_\nu$ обозначает ковариантную производную на пространственноподобном сечении

$$\hat{\nabla}_\nu \partial_0 g_{\mu\rho} = \partial_\nu \partial_0 g_{\mu\rho} - \hat{\Gamma}_{\nu\mu}{}^{\sigma}\partial_0 g_{\sigma\rho} - \hat{\Gamma}_{\nu\rho}{}^{\sigma}\partial_0 g_{\mu\sigma},$$

и $\hat{R}_{\mu\nu\rho\sigma}$ – тензор кривизны пространственно подобного сечения $t = \mathsf{const}$, построенный только по метрике $g_{\mu\nu}$. Свертка с обратной метрикой дает соответствующие тензор Риччи и скалярную кривизну:

$$\begin{aligned}
R_{00} &= \frac{1}{2} g^{\mu\nu}\partial_{00}^2 g_{\mu\nu} - \frac{1}{4} g^{\mu\nu} g^{\rho\sigma}\partial_0 g_{\mu\rho}\partial_0 g_{\nu\sigma},\\
R_{0\mu} &= R_{\mu 0} = \frac{1}{2} g^{\nu\rho}(\hat{\nabla}_\mu \partial_0 g_{\nu\rho} - \hat{\nabla}_\rho \partial_0 g_{\nu\mu}),\\
R_{\mu\nu} &= \hat{R}_{\mu\nu} + \frac{1}{2}\partial_{00}^2 g_{\mu\nu} - \frac{1}{2} g^{\rho\sigma}\partial_0 g_{\mu\rho}\partial_0 g_{\nu\sigma} + \frac{1}{4}\partial_0 g_{\mu\nu} g^{\rho\sigma}\partial_0 g_{\rho\sigma},\\
R &= \hat{R} + g^{\mu\nu}\partial_{00}^2 g_{\mu\nu} - \frac{3}{4} g^{\mu\nu} g^{\rho\sigma}\partial_0 g_{\mu\rho}\partial_0 g_{\nu\sigma} + \frac{1}{4}(g^{\mu\nu}\partial_0 g_{\mu\nu})^2.
\end{aligned} \tag{21.174}$$

Уравнения для экстремалей $x^\alpha(t)$ во временно́й калибровке имеют вид

$$\ddot{x}^0 = \frac{1}{2}\partial_0 g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu, \tag{21.175}$$

$$\ddot{x}^\mu = -g^{\mu\nu}\partial_0 g_{\nu\rho}\dot{x}^0\dot{x}^\rho - \hat{\Gamma}_{\nu\rho}{}^{\mu}\dot{x}^\nu\dot{x}^\rho, \tag{21.176}$$

где точка обозначает дифференцирование по каноническому параметру t. Из вида уравнений сразу следует

Предложение 21.12.4. *Если выбрана синхронная система координат, то временны́е координатные линии* $\{x^0 = t, x^\mu = \mathsf{const}\}$ *являются экстремалями.*

Из уравнений для экстремалей можно исключить временну́ю компоненту скорости. Для этого воспользуемся законом сохранения (17.22)

$$(\dot{x}^0)^2 - \dot{\boldsymbol{x}}^2 = C_0 = \mathsf{const}, \tag{21.177}$$

где $\dot{\boldsymbol{x}}^2 = -g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$, и исключим производную $\dot{x}^0$ из уравнения (21.176). В результате получим замкнутую систему уравнений только для пространственных координат экстремали:

$$\ddot{x}^\mu = -g^{\mu\nu}\partial_0 g_{\nu\rho}\dot{x}^\rho\sqrt{|C_0 + \dot{\boldsymbol{x}}^2|} - \hat{\Gamma}_{\nu\rho}{}^\mu\dot{x}^\nu\dot{x}^\rho.$$

Неоднозначность при извлечении корня несущественна, так как соответствует замене $x^0 \mapsto -x^0$. Отсюда следует, что во временно́й калибровке пространственные компоненты экстремали $\{x^\mu(t)\}$ в общем случае не являются экстремалями для пространственной части метрики $g_{\mu\nu}$. В частном случае, когда пространственная метрика $g_{\mu\nu}$ не зависит от времени x^0, проекция экстремали $\{x^\alpha(t)\} \mapsto \{c^0, x^\mu(t)\}$ на пространственное сечение $x^0 = c^0 = \mathsf{const}$ является экстремалью для метрики $g_{\mu\nu}$ на этом сечении.

Пусть $\dot{x}^0 \neq 0$. Тогда уравнения для экстремалей во временно́й калибровке (21.175), (21.176) можно переписать в эквивалентном виде, заменив канонический параметр t на наблюдаемое время x^0. Другими словами, рассмотрим траектории $x^\mu(x^0)$. Поскольку

$$\partial_0 x^\mu = \frac{\dot{x}^\mu}{\dot{x}^0},$$
$$\partial_{00}^2 x^\mu = \frac{\ddot{x}^\mu\dot{x}^0 - \dot{x}^\mu\ddot{x}^0}{(\dot{x}^0)^3},$$

то, воспользовавшись уравнениями для экстремалей, получим

$$\partial_{00}^2 x^\mu = -\hat{\Gamma}_{\nu\rho}{}^\mu\partial_0 x^\nu\partial_0 x^\rho - g^{\mu\nu}\partial_0 g_{\nu\rho}\partial_0 x^\rho - \frac{1}{2}\partial_0 x^\mu\partial_0 g_{\nu\rho}\partial_0 x^\nu\partial_0 x^\rho. \tag{21.178}$$

Зависимость канонического параметра t от наблюдаемого времени x^0 можно определить из закона сохранения (21.177), который можно переписать в виде

$$(\dot{x}^0)^2(1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu) = C_0.$$

Отсюда для ненулевых экстремалей, $C_0 \neq 0$, получаем выражение для производной

$$\dot{x}^0 = \pm\sqrt{\frac{C_0}{1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu}}.$$

Это уравнение позволяет определить зависимость $x^0(t)$.

Для нулевых экстремалей $C_0 = 0$ и $1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu = 0$. Последнее условие согласуется с уравнениями (21.178). Действительно, прямые вычисления приводят к равенству

$$\partial_0(1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu) = -(1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu)\partial_0 g_{\rho\sigma}\partial_0 x^\rho\partial_0 x^\sigma.$$

То есть, если условие $1 + g_{\mu\nu}\partial_0 x^\mu\partial_0 x^\nu = 0$ выполнено в начальный момент времени, то оно будет также выполнено во все последующие моменты. Это значит, что для нулевых экстремалей зависимость $x^0(t)$ не определяется уравнениями (21.178) и может быть произвольна. В этом случае канонический параметр можно просто отождествить с временем: $x^0 = t$.

Теперь покажем, что синхронная система координат в общей теории относительности при наличии материи не может быть статична. Для определенности рассмотрим четырехмерное пространство-время. С этой целью введем обозначение для производных по времени от компонент метрики

$$p_{\mu\nu} := \partial_0 g_{\mu\nu}.$$

Это нечто, напоминающее импульсы, сопряженные к компонентам метрики. След новых переменных равен производной от определителя метрики

$$p := g^{\mu\nu} p_{\mu\nu} = \partial_0[\,\mathsf{ln}(-g)], \qquad g := \mathsf{det}\, g_{\alpha\beta},$$

т.к. во временно́й калибровке $\mathsf{det}\, g_{\alpha\beta} = \mathsf{det}\, g_{\mu\nu}$. В новых переменных компоненты тензора Риччи (21.174) примут вид:

$$\begin{aligned} R_{00} &= \frac{1}{2}\partial_0 p + \frac{1}{4} p_{\mu\nu} p^{\mu\nu}, \\ R_{0\mu} &= \frac{1}{2}\hat{\nabla}_\mu p - \frac{1}{2}\hat{\nabla}_\nu p_\mu{}^\nu, \\ R_{\mu\nu} &= \hat{R}_{\mu\nu} + \frac{1}{2}\partial_0 p_{\mu\nu} - \frac{1}{2} p_\mu{}^\rho p_{\nu\rho} + \frac{1}{4} p_{\mu\nu} p. \end{aligned} \tag{21.179}$$

Запишем уравнения Эйнштейна (21.3) без космологической постоянной

$$R_{00} = -\frac{1}{2\kappa}\left(T_{\text{м}00} - \frac{1}{2} T_{\text{м}}\right), \tag{21.180}$$

$$R_{0\mu} = -\frac{1}{2\kappa} T_{\text{м}0\mu}, \tag{21.181}$$

$$R_{\mu\nu} = -\frac{1}{2\kappa}\left(T_{\text{м}\mu\nu} - \frac{1}{2} g_{\mu\nu} T_{\text{м}}\right). \tag{21.182}$$

Эти уравнения для непрерывной среды с тензором энергии-импульса (21.117) принимают вид

$$R_{00} = -\frac{1}{2\kappa}\left((\mathcal{E}+\mathcal{P})(u_0)^2 - \frac{1}{2}\mathcal{E} + \frac{1}{2}\mathcal{P}\right), \tag{21.183}$$

$$R_{0\mu} = -\frac{1}{2\kappa}(\mathcal{E}+\mathcal{P}) u_0 u_\mu, \tag{21.184}$$

$$R_{\mu\nu} = -\frac{1}{2\kappa}\left((\mathcal{E}+\mathcal{P}) u_\mu u_\nu - \frac{1}{2} g_{\mu\nu}\mathcal{E} + \frac{1}{2} g_{\mu\nu}\mathcal{P}\right). \tag{21.185}$$

Предложение 21.12.5. *Пусть пространство-время заполнено непрерывной средой с положительной плотностью энергии и давлением, $\mathcal{E} > 0$ и $\mathcal{P} > 0$. Тогда уравнения Эйнштейна (21.183)–(21.185) без космологической постоянной в синхронной системе координат не имеют статических решений.*

Доказательство. В синхронной системе координат $u^2 = (u_0)^2 + g^{\mu\nu} u_\mu u_\nu = 1$, и поэтому $u_0 \geq 1$. Статичность метрики означает, что $p_{\mu\nu} = 0$. Тогда из уравнения (21.184) следует равенство $u_\mu = 0$. Следовательно, $u_0 = 1$. Тогда уравнение (21.183) примет вид

$$0 = \frac{1}{2}\mathcal{E} + \frac{3}{2}\mathcal{P},$$

что противоречит сделанным предположениям. □

Предложение 21.12.6. *Вакуумные уравнения Эйнштейна без космологической постоянной в синхронной системе координат имеют только тривиальные статические решения.*

Доказательство. При нулевом тензоре энергии-импульса материи, $T_{\text{м}\alpha\beta} = 0$, уравнения (21.180) и (21.181) для $p_{\mu\nu} = 0$ удовлетворяются автоматически. Уравнение (21.185) сводится к уравнению

$$\hat{R}_{\mu\nu} = 0.$$

Если пространство трехмерно, то полный тензор кривизны взаимно однозначно определяется тензором Риччи и, следовательно, полный тензор кривизны пространства равен нулю, $\hat{R}_{\mu\nu\rho\sigma} = 0$. Теперь из выражения для компонент тензора кривизны (21.173) следует, что полный тензор кривизны пространства-времени также обращается в нуль, $R_{\alpha\beta\gamma\delta} = 0$. В свою очередь, это значит, что локально существует такая система координат, в которой метрика четырехмерного пространства-времени является локально лоренцевой, $g_{\alpha\beta} = \eta_{\alpha\beta}$. Это и означает тривиальность решений уравнений Эйнштейна. Глобально соответствующее пространство-время представляет либо пространство Минковского, либо все возможные торы или цилиндры, получающиеся из пространства Минковского путем отображения всех или части декартовых координат на окружность. □

Теперь предположим, что выполнено сильное энергетическое условие (21.130). Тогда из уравнения (21.183) следует неравенство

$$\frac{1}{2}\partial_0 p + \frac{1}{4} p_\mu{}^\nu p_\nu{}^\mu \le 0.$$

Справедливо алгебраическое неравенство

$$p_\mu{}^\nu p_\nu{}^\mu \ge \frac{1}{3} p^2.$$

Его нетрудно доказать путем диагонализации матрицы $p_\mu{}^\nu$. Поэтому должно быть выполнено неравенство

$$\partial_0 p + \frac{1}{6} p^2 \le 0.$$

Перепишем его в виде

$$\partial_0 \frac{1}{p} \ge \frac{1}{6}. \tag{21.186}$$

Теперь предположим, что в некоторый момент времени $x^0 := t$ величина $1/p$ положительна. Тогда при уменьшении времени t за конечное время функция $1/p$ обратится в нуль, поскольку производная больше нуля. Это означает, что модуль определитель метрики $|g|$ обращается в нуль. Допустим, что это происходит в момент времени t_0. Вблизи этой точки положим

$$-g = C(t - t_0)^k, \qquad C = \mathsf{const},$$

с некоторым показателем степени $k \in \mathbb{R}_+$. Тогда неравенство (21.186) ограничивает показатель степени $k \le 6$. Отсюда следует, что модуль определителя метрики обращается в нуль не быстрее, чем $(t - t_0)^6$.

Если в начальный момент времени $t = 0$ величина $1/p$ отрицательна, то то же самое происходит при положительных временах.

Обращение в нуль определителя метрики отнюдь не означает, что возникает сингулярность в пространстве-времени. Например, определитель евклидовой метрики в сферической системе координат равен нулю в начале координат. Как правило, обращение в нуль определителя метрики связано с выбором системы отсчета. Выше было показано, что синхронная система координат возникает при построении семейства экстремалей, которые ортогональны некоторой пространственноподобной гиперповерхности. В общем случае на конечном расстоянии эти экстремали начинают пересекаться, образуя каустические поверхности. В точках пересечения экстремалей система координат вырождается, что может приводить к обращению в нуль определителя метрики.

21.12.3 Калибровка светового конуса

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, на котором задана метрика $g_{\alpha\beta}$ лоренцевой сигнатуры.

Определение. Система координат, в которой метрика имеет вид

$$g_{\alpha\beta} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & & & & \\ 0 & & & & \\ \vdots & & & g_{\mu\nu} & \\ 0 & & & & \end{pmatrix}, \qquad \begin{matrix} \alpha, \beta = 0, 1, \dots, n-1, \\ \mu, \nu = 1, 2, \dots, n-1, \end{matrix} \tag{21.187}$$

называется калибровкой *светового конуса*. □

Ниже мы докажем, что локально калибровка светового конуса существует. При этом n произвольных функций, параметризующих преобразования координат, используются для фиксирования n компонент метрики:

$$g_{00} = g_{02} = g_{03} = \dots = g_{0\,n-1} = 0, \qquad g_{01} = 1.$$

Для построения, системы координат, соответствующей калибровке светового конуса, введем несколько определений. Допустим, что гиперповерхность $\mathbb{S} \hookrightarrow \mathbb{M}$ задана параметрически $x^\alpha = x^\alpha(y^\mu)$, где y^μ – координаты на гиперповерхности. Вложение $\mathbb{S} \hookrightarrow \mathbb{M}$ индуцирует на гиперповерхности метрику

$$ds^2 = h_{\mu\nu} dy^\mu dy^\nu = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^\mu} \frac{\partial x^\beta}{\partial y^\nu} dy^\mu dy^\nu. \tag{21.188}$$

В этом выражении индуцированная метрика $h_{\mu\nu}$ в общем случае не совпадает с блоком $g_{\mu\nu}$, входящим в выражение (21.187).

Определение. Если индуцированная метрика $h_{\mu\nu}$ на гиперповерхности невырождена, то гиперповерхность называется *неизотропной*. В противном случае, если $\det h_{\mu\nu} = 0$, гиперповерхность называется *изотропной*. □

В дальнейшем мы увидим, что индуцированная метрика может быть вырождена только тогда, когда метрика не является положительно или отрицательно определенной.

Набор $n-1$ векторных полей на $\mathbb{S}$, которые нумеруются индексом μ,

$$e_\mu := \partial_\mu x^\alpha \partial_\alpha := \frac{\partial x^\alpha}{\partial y^\mu} \partial_\alpha$$

образуют базис касательного пространства к гиперповерхности $\mathbb{S} \hookrightarrow \mathbb{M}$. Вектор $m = m^\alpha \partial_\alpha$ ортогонален к гиперповерхности, если его скалярное произведение со всеми касательными векторами равно нулю:

$$(m, e_\mu) := g_{\alpha\beta} m^\alpha \partial_\mu x^\beta = m_\alpha \partial_\mu x^\alpha = 0, \qquad \forall \mu = 1, 2, \ldots, n-1, \tag{21.189}$$

где скобки означают скалярное произведение. Поскольку ранги матриц $g_{\alpha\beta}$ и $\partial_\mu x^\alpha$ равны, соответственно, n и $n-1$, то ранг матрицы $g_{\alpha\beta}\partial_\mu x^\beta$ равен $n-1$. Это означает, что система линейных однородных уравнений на компоненты нормального вектора m^α имеет единственное нетривиальное решение с точностью до умножения на произвольную отличную от нуля постоянную.

Важную роль в дальнейшем будет играть следующее

Предложение 21.12.7. *Нормаль к гиперповерхности изотропна тогда и только тогда, когда поверхность изотропна.*

Доказательство. Индуцированную метрику $h_{\mu\nu}$ (21.188) можно рассматривать, как произведение двух прямоугольных матриц $\partial_\mu x^\alpha$ и $g_{\alpha\beta}\partial_\mu x^\beta$ размеров, соответственно, $(n-1)\times n$ и $n \times (n-1)$. Тогда из формулы Бине–Коши (28.9) следует, что

$$\det h_{\mu\nu} = \sum_{\gamma=1}^{n} \det{}_\gamma(\partial_\mu x^\alpha) \det{}_\gamma(g_{\alpha\beta}\partial_\mu x^\beta),$$

где $\det_\gamma$ обозначает определитель квадратной матрицы, полученной из матрицы размера $(n-1)\times n$ вычеркиванием γ-того столбца. Поскольку $m_\alpha \partial_\mu x^\alpha = 0$, то между столбцами матрицы $\partial_\mu x^\alpha = 0$ имеется линейная зависимость. Предположим, что некоторая компонента ковектора $\{m_\alpha\}$ отлична от нуля, $m_{\delta_1} \neq 0$. Тогда, в силу свойств определителя и линейной зависимости столбцов, справедливо равенство

$$\det{}_\gamma(\partial_\mu x^\alpha) = m_\gamma \frac{(-1)^{\delta_1-\gamma} \det{}_{\delta_1}(\partial_\mu x^\alpha)}{m_{\delta_1}}.$$

Аналогичная формула имеет место и для второго определителя под знаком суммы

$$\det{}_\gamma(g_{\alpha\beta}\partial_\mu x^\beta) = m^\gamma \frac{(-1)^{\delta_2-\gamma} \det{}_{\delta_2}(g_{\alpha\beta}\partial_\mu x^\beta)}{m^{\delta_2}}$$

для некоторого индекса δ_2. Таким образом, определитель индуцированной метрики пропорционален длине нормального вектора к гиперповерхности

$$\det h_{\mu\nu} = (-1)^{\delta_1+\delta_2} \frac{\det{}_{\delta_1}(\partial_\mu x^\alpha) \det{}_{\delta_2}(g_{\alpha\beta}\partial_\mu x^\beta)}{m_{\delta_1} m^{\delta_2}} m_\gamma m^\gamma,$$

где в правой части предполагается суммирование только по индексу γ. Индексы δ_1 и δ_2 фиксированы и выбраны таким образом, что $m_{\delta_1} \neq 0$ и $m^{\delta_2} \neq 0$. Поскольку коэффициент пропорциональности между $\det h_{\mu\nu}$ и $m_\gamma m^\gamma$ отличен от нуля, то из этого равенства следует утверждение предложения. □

Рассмотрим некоторые свойства изотропных гиперповерхностей.

Допустим, что изотропная гиперповерхность $\mathbb{S}$ задана уравнением

$$W(x) = \text{const}. \tag{21.190}$$

При этом мы считаем, что различным значениям постоянной соответствуют различные непересекающиеся изотропные гиперповерхности. Обозначим, как и ранее, координаты на $\mathbb{S}$ через y^μ. То есть функции $x^\alpha(y)$ удовлетворяют уравнению (21.190) для всех y. Продифференцируем это уравнение:

$$\frac{\partial W}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial y^\mu} = \frac{\partial W}{\partial x^\alpha} e_\mu{}^\alpha = 0.$$

Сравнение полученного равенства с определением нормали (21.189) показывает, что вектор нормали к гиперповерхности имеет вид

$$m = m^\alpha \partial_\alpha = g^{\alpha\beta}\frac{\partial W}{\partial x^\beta}\partial_\alpha. \tag{21.191}$$

Поскольку вектор нормали к изотропной гиперповерхности имеет нулевую длину, то выполнено равенство

$$m^2 = g^{\alpha\beta}\frac{\partial W}{\partial x^\alpha}\frac{\partial W}{\partial x^\beta} = 0. \tag{21.192}$$

Таким образом доказано

Предложение 21.12.8. *Поверхность $\mathbb{S}$, определяемая уравнением (21.190), изотропна тогда и только тогда, когда выполнено равенство (21.192).*

Следующее утверждение не имеет аналога в римановой геометрии.

Предложение 21.12.9. *Нормальный вектор к изотропной поверхности $\mathbb{S}$ является также и касательным к ней.*

Доказательство. Дополним координаты на гиперповерхности до системы координат в $\mathbb{M}$ еще одной координатой $y^0 := W$. В новой системе координат гиперповерхности определяются условиями $y^0 = \mathsf{const}$, и уравнение (21.192) примет вид $g^{00} = 0$. Тогда нулевая компонента нормального вектора (21.191) обратится в нуль, $m^0 = 0$, т.к. $\partial W/\partial y^\mu = 0$. Это значит, что нормальный вектор является также и касательным. □

На самом деле сделанное утверждение очевидно. Вектор k будет касательным к гиперповерхности $\mathbb{S}$ тогда и только тогда, когда его скалярное произведение с вектором нормали равно нулю, $(k, m) = 0$. Это и означает, что вектор нормали m является касательным к $\mathbb{S}$, т.к. $m^2 = 0$.

Рассмотрим интегральные кривые $x(u)$ векторного поля m:

$$\dot{x}^\alpha = m^\alpha, \qquad m^\alpha := g^{\alpha\beta}\frac{\partial W}{\partial x^\beta}, \tag{21.193}$$

где точка обозначает дифференцирование по параметру $u \in \mathbb{R}$. Эти кривые изотропны, т.к. выполнено равенство (21.192). Кроме того, они перпендикулярны изотропной поверхности $\mathbb{S}$. С другой стороны, вдоль каждой кривой выполнено равенство

$$\frac{dW}{du} = \frac{\partial W}{\partial x^\alpha}\frac{dx^\alpha}{du} = g^{\alpha\beta}\frac{\partial W}{\partial x^\alpha}\frac{\partial W}{\partial x^\beta} = 0.$$

Отсюда следует, что, если интегральная кривая $x(u)$ начинается на изотропной поверхности $\mathbb{S}$, то она целиком лежит на этой поверхности.

Можно сказать по-другому. Поскольку векторное поле m касательно к $\mathbb{S}$, то его интегральные кривые лежат на $\mathbb{S}$.

Вычислим теперь ускорение данной кривой:

$$\dot{x}^\alpha \nabla_\alpha \dot{x}^\beta = \dot{x}^\alpha g^{\beta\gamma} \nabla_\alpha \frac{\partial W}{\partial x^\gamma} = g^{\alpha\delta} \frac{\partial W}{\partial x^\delta} g^{\beta\gamma} \nabla_\alpha \frac{\partial W}{\partial x^\gamma}.$$

С другой стороны, продифференцируем определяющее уравнение (21.192):

$$\nabla_\alpha \left(g^{\beta\gamma} \frac{\partial W}{\partial x^\beta} \frac{\partial W}{\partial x^\gamma} \right) = 2 g^{\beta\gamma} \frac{\partial W}{\partial x^\beta} \nabla_\alpha \frac{\partial W}{\partial x^\gamma} = 0.$$

Поскольку $\nabla_\alpha \nabla_\gamma W = \nabla_\gamma \nabla_\alpha W$, то отсюда вытекает, что ускорение кривой (21.193) равно нулю, $\dot{x}^\alpha \nabla_\alpha \dot{x}^\beta = 0$. Таким образом, интегральные кривые векторного поля m являются экстремалями. Отсюда следует

Предложение 21.12.10. *Все изотропные (нулевые) экстремали, касательные к изотропной поверхности $\mathbb{S}$, целиком лежат на этой поверхности.*

Теперь докажем существование калибровки светового конуса.

Предложение 21.12.11. *Пусть задано пространство-время $\mathbb{M}$ с метрикой лоренцевой сигнатуры. Тогда в некоторой окрестности каждой точки $x \in \mathbb{M}$ существует система координат, в которой метрика имеет вид (21.187).*

Доказательство. Доказательство аналогично доказательству предложения 21.12.3. Выберем произвольную пространственноподобную поверхность $\mathbb{N} \hookrightarrow \mathbb{M}$ с координатами y^μ. В каждой точке $y \in \mathbb{N}$ строим нормальный вектор n. Он времениподобен. Выберем координаты y^μ на $\mathbb{N}$ таким образом, чтобы вектор

$$e_1 := \frac{\partial x^\alpha}{\partial y^1} \partial_\alpha$$

имел единичную длину, $e_1^2 = -1$, и был перпендикулярен всем остальным касательным векторам:

$$(e_1, e_{\text{A}}) = 0, \qquad \text{A} = 2, \ldots, n-1.$$

Это всегда возможно в силу предложения 21.12.3. То есть мы фиксируем временну́ю калибровку на поверхности $\mathbb{N}$. Затем строим вектор

$$m = n - e_1.$$

По построению, этот вектор определен на поверхности $\mathbb{N}$, имеет нулевую длину, $m^2 = 0$, перпендикулярен всем касательным векторам e_{A}, и выполнено равенство $(m, e_1) = 1$. Другими словами, в каждом световом конусе в точке $y \in \mathbb{N}$ мы выбираем единственный вектор m, скалярное произведение которого с касательным вектором e_1 равно единице, а с остальными касательными векторами равно нулю.

Затем через каждую точку $y \in \mathbb{N}$ в направлении m проводим экстремаль. Выбираем канонический параметр y^0 вдоль экстремалей таким образом, чтобы гиперповерхность $\mathbb{N}$ определялась уравнением $y^0 = 0$. Тогда в некоторой окрестности гиперповерхности $\mathbb{N}$ определена система координат $\{y^\alpha\} = \{y^0, y^\mu\}$.

По-построению, координатные линии y^0 являются экстремалями. Поэтому на компоненты метрики в новой системе координат имеются ограничения (21.168). В рассматриваемом случае касательный вектор к координатной линии y^0 изотропен, и, следовательно, $g_{00} = 0$. Поэтому на компоненты метрики возникает уравнение (21.168), которое сводится к условию

$$\partial_0 g_{0\mu} = 0. \tag{21.194}$$

По-построению, координаты y^μ на $\mathbb{N}$ и изотропный вектор m выбраны таким образом, что на поверхности $\mathbb{N}$ выполнены условия:

$$g_{00} = 0, \quad g_{01} = 1, \quad g_{02} = g_{03} = \ldots = g_{0\,n-1} = 0.$$

Тогда из уравнений (21.194) следует, что эти условия будут выполнены во все последующие моменты времени. Следовательно, система координат, в которой выполнена калибровка светового конуса построена.

Аналогичное построение можно провести, если в качестве исходной поверхности $\mathbb{N}$ выбрать поверхность с метрикой лоренцевой сигнатуры. □

Посмотрим теперь на построение калибровки светового конуса с точки зрения уравнения Гамильтона–Якоби. Предположим, что изотропная гиперповерхность $\mathbb{S}$ задана уравнением $W(x) = 0$. Тогда условие изотропности поверхности примет вид (21.192), которое должно быть выполнено на изотропной гиперповерхности $\mathbb{S}$. Изотропная гиперповерхность, задаваемая уравнением $W = 0$, является ни чем иным, как характеристикой волнового уравнения (17.49). Уравнение (21.192) представляет собой укороченное уравнение Гамильтона–Якоби для нулевых экстремалей, где $W(x)$ – укороченная функция действия.

Рассмотрим систему линейных дифференциальных уравнений первого порядка

$$g^{\alpha\beta}\partial_\alpha W \partial_\beta \varphi = 0$$

на функцию $\varphi(x)$. Известно, что эта система имеет $n-1$ функционально независимых решений, которые обозначим через $\varphi^\mu(x)$, $\mu = 1, \ldots, n-1$. Среди этих решений содержится, очевидно, и $W(x)$. Не ограничивая общности, положим $\varphi^1(x) := W(x)$. Введем новую систему координат

$$y^0 := \psi(x), \quad y^\mu := \varphi^\mu(x),$$

где $\psi(x)$ – некоторая функция, функционально независимая от $\varphi^\mu(x)$, т.е.

$$\frac{\partial y^\alpha}{\partial x^\beta} \neq 0$$

где $\{y^\alpha\} = \{y^0, y^\mu\}$. В новой системе координат компоненты обратной метрики примут вид

$$g'^{\alpha\beta} = g^{\gamma\delta}\frac{\partial y^\alpha}{\partial x^\gamma}\frac{\partial y^\beta}{\partial x^\delta}.$$

По-построению

$$g'^{01} = g'^{10} = g^{\alpha\beta}\partial_\alpha W \partial_\beta \psi \neq 0, \quad g'^{1\mu} = g'^{\mu 1} = 0,$$

причем $g'^{01} \neq 0$, так как ψ функционально независима от φ^μ. Эту функцию всегда можно подобрать таким образом, что $g'^{01} = 1$. Таким образом, в новой системе координат обратная метрика примет вид

$$g'^{\alpha\beta} = \begin{pmatrix} g'^{00} & 1 & g'^{0\text{в}} \\ 1 & 0 & 0 \\ g'^{\text{а}0} & 0 & g'^{\text{ав}} \end{pmatrix}, \qquad \text{а, в} = 2, 3, \ldots, n-1.$$

Нетрудно проверить, что метрика $g'_{\alpha\beta}$, которая обратна к матрице $g'^{\alpha\beta}$, в новой системе координат имеет вид (21.187). Тем самым локальное существование калибровки

светового конуса доказано. В этой калибровке изотропные гиперповерхности $\mathbb{S}$ задаются условиями $y^1 = \mathsf{const}$. Векторы ∂_0, касательные к координатным линиям y^0 и, следовательно, касательные также к изотропной гиперповерхности, имеют нулевую длину $(\partial_0, \partial_0) = g_{00} = 0$. Метрика, индуцированная на гиперповерхностях $y^1 = \mathsf{const}$, имеет вид

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & g'_{\text{AB}} & \\ 0 & & & \end{pmatrix}, \qquad \text{A, B} = 2, 3, \dots, n-1,$$

и является, очевидно, вырожденной.

Вернемся к прежним обозначениям координат через x^α вместо y^α. То есть будем считать, что в координатах x^α зафиксирована калибровка светового конуса, и метрика имеет вид (21.187)

$$ds^2 = 2dx^0dx^1 + g_{\mu\nu}dx^\mu dx^\nu. \tag{21.195}$$

Наша интуиция основана на таких системах координат, в которых метрика диагональна. Поскольку в калибровке светового конуса метрика недиагональна, то полезно рассмотреть простой пример плоского пространства Минковского.

Пример 21.12.2. Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$. В инерциальной системе отсчета t, x, y метрика Лоренца имеет вид

$$ds^2 = dt^2 - dx^2 - dy^2.$$

В плоскости t, x можно ввести координаты светового конуса (2.82)

$$u = \frac{1}{\sqrt{2}}(t + x), \qquad v = \frac{1}{\sqrt{2}}(t - x).$$

Рассмотрим более общее преобразование координат

$$\xi = \frac{1}{\sqrt{2}}\Big((2-a)t + ax\Big), \qquad \eta = \frac{1}{\sqrt{2}}(t - x), \qquad a \in \mathbb{R}, \tag{21.196}$$

которое параметризуется одним вещественным параметром a. Эти координаты в плоскости t, x показаны на рис. 21.3. На рисунке показаны конусные оси v и u, которые определяются уравнениями, соответственно, $u = 0$ и $v = 0$. Ось ξ совпадает с осью u. Показана также ось η при $0 < a < 1$. В рассматриваемой системе координат, координата ξ по-прежнему является конусной, а координата η совпадает с конусной координатой v только при $a = 1$. Легко найти обратное преобразование координат

$$t = \frac{1}{\sqrt{2}}(\xi + a\eta), \qquad x = \frac{1}{\sqrt{2}}\Big(\xi + (a-2)\eta\Big).$$

Метрика Лоренца в координатах $(x^0, x^1, x^2) = (\xi, \eta, y)$ имеет вид

$$ds^2 = 2d\xi d\eta + 2(a-1)d\eta^2 - dy^2.$$

Компоненты метрики и ее обратной в координатах ξ, η, y равны

$$g_{\alpha\beta} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2(a-1) & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad g^{\alpha\beta} = \begin{pmatrix} -2(a-1) & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

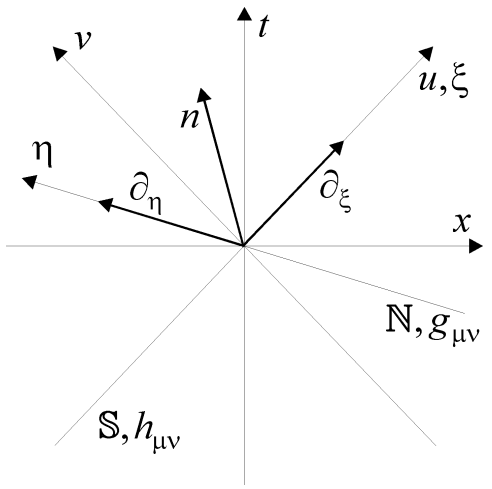


Рис. 21.3: Координаты ξ, η в плоскости t, x.

Следовательно, координаты $(x^0, x^1, x^2) = (\xi, \eta, y)$ задают калибровку светового конуса.

Отметим, что метрика $g_{\mu\nu}$ на плоскостях $\xi = \mathsf{const}$ является отрицательно определенной при $a < 1$ и имеет лоренцеву сигнатуру при $a > 1$.

В трехмерном пространстве Минковского существует два типа изотропных гиперповерхностей: плоскости и конусы (характеристики в примере 17.7.1). Изотропные плоскости – это все плоскости, имеющие угол $\pi/4$ с осью t.

Сравним координаты светового конуса данного примера при $0 < a < 1$ с системой координат, построенной в предложении (21.12.11) (см. рис.21.3). Пространственноподобная поверхность $\mathbb{N}$ натянута на векторы $\partial_\eta, \partial_y$ и определяется уравнением $x^0 = \xi = 0$. Изотропные поверхности $\mathbb{S}$ натянуты на векторы ∂_ξ, ∂_y и определяются уравнениями $x^1 = \eta = \mathsf{const}$. Вектор

$$\partial_\xi = \frac{1}{\sqrt{2}}\partial_t + \frac{1}{\sqrt{2}}\partial_x$$

касается изотропной поверхности $\mathbb{S}$. Вектор

$$\partial_\eta = \frac{a}{\sqrt{2}}\partial_t - \frac{2-a}{\sqrt{2}}\partial_x$$

касается поверхности $\mathbb{N}$. Нормированный вектор e_1 пропорционален ∂_η:

$$e_1 = \frac{1}{\sqrt{2(1-a)}}\partial_\eta = \frac{a}{2\sqrt{1-a}}\partial_t - \frac{2-a}{2\sqrt{1-a}}\partial_x.$$

Нормированный вектор нормали n к пространственноподобной поверхности $\mathbb{N}$ имеет вид

$$n = \frac{2-a}{2\sqrt{1-a}}\partial_t - \frac{a}{2\sqrt{1-a}}\partial_x.$$

Вычислим изотропный вектор m:

$$m := n - e_1 = \sqrt{1-a}\,\partial_t + \sqrt{1-a}\,\partial_x = \sqrt{2(1-a)}\,\partial_\xi.$$

Как и следовало ожидать, он пропорционален вектору ∂_ξ. □

Рассмотренный пример показывает, что калибровка светового конуса (21.195) определена неоднозначно и может быть зафиксирована таким образом, что метрика $g_{\mu\nu}$ на гиперповерхностях $x^0 = \mathsf{const}$ будет либо отрицательно определена, либо иметь лоренцеву сигнатуру в произвольной конечной области $\mathbb{U}$.

21.13 О постановке задач в теории гравитации

Отметим специфику задач, возникающих при рассмотрении произвольного функционала действия, инвариантного относительно общих преобразований координат и содержащего метрику. Уравнения Эйлера–Лагранжа записываются и решаются в произвольной, но фиксированной системе координат, т.е. локальны. Допустим, что мы нашли какое-то решение уравнений Эйлера–Лагранжа в системе координат, совпадающей со всем $\mathbb{R}^n$. Поскольку исходное действие инвариантно относительно общих преобразований координат, то без потери общности можно отобразить все $\mathbb{R}^n$, и тем самым найденное решение, на ограниченную область, например, в шар. В связи с этим возникает вопрос нельзя ли найденное решение продолжить, т.е. существует ли решение в большей области, сужение которого совпадает с уже найденным решением в шаре. Для ответа на этот вопрос необходимо инвариантное определение глобальности решения, которое дается через полноту геодезических и экстремалей.

В теории гравитации, основанной на аффинной геометрии, также, как и в общей теории относительности, понятие геодезической и экстремали является чрезвычайно важным, поскольку позволяет установить связь между локальными решениями уравнений движения и глобальными свойствами многообразий.

Определение. Назовем многообразие $\mathbb{M}$ с заданной аффинной геометрией, т.е. тройку $(\mathbb{M}, g, \Gamma)$, *полным*, если любую геодезическую и экстремаль в $\mathbb{M}$ можно продолжить до бесконечного значения канонического параметра в обе стороны. На практике часто локальное решение уравнений движения нельзя продолжить до полного решения, поскольку возможно существование сингулярностей. Назовем точку многообразия $\mathbb{M}$ *сингулярной*, если в этой точке по крайней мере одно скалярное поле, построенное из метрики и (или) аффинной связности обращается в бесконечность. □

В этом определении важно, чтобы в сингулярной точке именно скалярная комбинация геометрических объектов обращалась в бесконечность, поскольку наряду с истинными сингулярностями могут существовать координатные сингулярности, связанные с неудачным выбором системы координат. Например, хорошо известна координатная особенность метрики на горизонте событий для решения Шварцшильда, от которой можно избавиться, перейдя, например, к системе координат Эддингтона–Финкельстейна или Крускала–Секереша. Простейшими функциями, определяющими положение сингулярностей являются скалярная кривизна и квадрат тензора кручения, имеющие одинаковую размерность. Примером истинной сингулярности может служить черная дыра в решении Шварцшильда, в которой квадрат тензора кривизны обращается в бесконечность.

В связи с возможным существованием сингулярностей оказывается полезным понятие максимально продолженного многообразия.

Определение. Многообразие с заданной аффинной геометрией назовем *максимально продолженным*, если любую геодезическую и экстремаль можно либо продолжить

до бесконечного значения канонического параметра, либо они продолжаются до сингулярной точки при конечном значении канонического параметра. Соответствующую тройку $(\mathbb{M}, g, \Gamma)$ назовем *глобальным решением* в теории гравитации. □

Слово "продолжение" в этом определении связано с тем, что в моделях математической физики, как правило, метрика и кручение сначала находятся только локально в какой либо области, как решение уравнений движения в заданной системе координат, а затем, если необходимо, это решение гладко продолжается в соседние области до тех пор, пока дальнейшее продолжение становится невозможным либо вследствие сингулярностей, либо в связи с тем, что полученное решение станет полным.

Важно отметить, что определение максимально продолженного многообразия является инвариантным и не зависит от выбора локальной системы координат. Это связано с тем, что канонический параметр определяется однозначно с точностью до линейного преобразования и не зависит от системы координат.

Пример 21.13.1. Рассмотрим евклидову плоскость $\mathbb{R}^2$ в полярных координатах r, φ. Будем считать, что полярный угол меняется в бесконечном интервале, $0 < \varphi < \infty$.

Определение. *Гиперболической спиралью* называется кривая на евклидовой плоскости, которая в полярных координатах заданная уравнением

$$r = \frac{1}{\varphi}, \quad 0 < \varphi < \infty. \quad \square \tag{21.197}$$

Гиперболическая спираль соединяет бесконечно удаленную точку с декартовыми координатами $(\infty, 0)$ и начало координат $(0, 0)$, вокруг которого она наматывается бесконечное число раз. Если полярный угол меняется в интервале $\varphi \in (\varphi_1, \varphi_2) \subset (0, \infty)$, то ее длина определяется интегралом

$$l(\varphi_1, \varphi_2) = \int_{\varphi_1}^{\varphi_2} d\varphi \sqrt{dr^2 + r^2 d\varphi^2} = \int_{\varphi_1}^{\varphi_2} d\varphi \frac{\sqrt{1+\varphi^2}}{\varphi^2}.$$

Этот интеграл расходится на обоих концах гиперболической спирали, как при $\varphi \to 0$, так и при $\varphi \to \infty$. Расходимость интеграла при $\varphi \to 0$ естественна, т.к. любая кривая, уходящая в бесконечность, имеет бесконечную длину. Начало координат является конечной точкой, т.к. любая экстремаль, проходящая через эту точку, проходит ее при конечном значении канонического параметра. В то же время не любая кривая, подходящая к этой точке, имеет конечную длину. Это связано с тем, что гиперболическая спираль достаточно быстро наматывается вокруг начала координат. □

Данный пример показывает, что в определении полноты многообразий требование полноты экстремалей, а не произвольных кривых является существенным.

Для простоты, хотелось бы, чтобы конечность канонического параметра для одной экстремали, проходящей через заданную точку $x \in \mathbb{M}$, влекла конечность канонических параметров для всех других экстремалей, проходящих через эту точку. В случае риманова многообразия $(\mathbb{M}, g)$ это, по-видимому, так. Однако для псевдоримановых многообразий это не так. В разделе **??** будет построен пример лоренцева двумерного многообразия, у которого существуют точки, которые достигаются всеми пространственно- и времениподобными геодезическими при конечном значении канонического параметра, в то время как светоподобные геодезические достигают этих точек при бесконечном значении канонического параметра.

С физической точки зрения требование полноты или максимального продолжения пространства-времени $\mathbb{M}$ является естественным. Действительно, если рассмотреть движение точечной частицы в сопутствующей системе отсчета, в которой время является каноническим параметром, то естественно предположить, что эволюция продолжается либо бесконечно долго, либо обрывается в сингулярной точке.

Отметим, что полное многообразие не может иметь края. При этом существуют две возможности: либо $\mathbb{M}$ некомпактно, либо пространство-время компактно и без края (замкнутая вселенная). В первом случае геометрические инварианты в бесконечности могут стремиться как к конечным, так и бесконечным значениям. В космологии принята следующая терминология. Если все пространственные сечения пространства-времени $\mathbb{M}$ некомпактны, то вселенная *бесконечна*. Если все пространственные сечения компактны и без края, то говорят, что вселенная *замкнута*. При наличии сингулярностей, соответствующих конечному значению канонического параметра, сингулярные точки образуют край пространства-времени, находящийся на конечном расстоянии (при конечных значениях канонического параметра).

Чтобы найти максимально продолженной решение необходимо пройти несколько этапов: 1) решить уравнения Эйлера–Лагранжа в некоторой области, 2) найти и проанализировать полноту всех геодезических и экстремалей для соответствующей метрики и связности, 3) если область, где найдено решение, оказалась неполной, то продолжить решение. Первые два этапа чрезвычайно сложны, поскольку предполагают решение нелинейных систем дифференциальных уравнений. Последний этап также сложен. Его можно осуществить по крайней мере двумя способами. Либо перейти в новую систему координат, охватывающую бóльшую область, либо найти решение в соседней области, а затем доказать гладкость склейки.

В общей теории относительности существуют только отдельные примеры максимально продолженных многообразий. Например, расширение Крускала–Секереша решения Шварцшильда.

Глава 22

Гамильтонова формулировка общей теории относительности

В настоящей главе мы перепишем вакуумные (без полей материи) уравнения Эйнштейна в гамильтоновом виде. Эта задача важна как для исследования классических уравнений движения, в частности, при рассмотрении задачи Коши, так и для построения квантовой теории гравитации. Сначала будет развит метрический гамильтонов формализм, в котором независимыми динамическими переменными являются компоненты метрики и сопряженные им импульсы. Затем мы решим эту же задачу, рассматривая в качестве независимой переменной репер.

Каноническая формулировка общей теории относительности была впервые дана Дираком в формализме второго порядка [**?**]. Он показал, что гамильтониан гравитационного поля равен линейной комбинации связей и нашел его явное выражение. Позже Арновитт, Дезер и Мизнер в серии статей, завершившихся обзором [**?**] (без ссылки на Дирака), существенно упростили вычисления и прояснили геометрический смысл канонических импульсов, выразив их через внешнюю кривизну пространственной гиперповерхности, вложенной в четырехмерное пространство-время. Выражение для гамильтониана было найдено ими в формализме первого порядка, когда метрика $g_{\alpha\beta}$ и симметричная аффинная связность $\Gamma_{\{\alpha\beta\}}{}^{\gamma}$ рассматриваются в качестве независимых переменных. По сути дела этот подход и упростил вычисления. В пионерских работах [**?**, **?**] не была посчитана явно алгебра связей, но показана их согласованность с уравнениями движения (связи первого рода). В статье [**?**] была также проанализирована роль граничных членов и дано определение полной энергии гравитационного поля в виде поверхностного интеграла для асимптотически плоского пространства-времени. Для получения выражения для гамильтониана для физических степеней свободы была использована калибровка, явно зависящая от времени. Численное значение гамильтониана на уравнениях движения при выбранном калибровочном условии выражается через поверхностный интеграл. Если границы нет, то энергия равна нулю. В линейном приближении доказано, что граничный член дает правильные уравнения движения для физических степеней свободы. Для Шварцшильда $E = M$.

В статье [**?**] дана каноническая формулировка общей теории относительности и подробно рассмотрено обобщение уравнения Шредингера для волновой функции Вселенной, которое в дальнейшем получило название уравнения Уилера–ДеВитта. Здесь впервые посчитана вся алгебра связей общей теории относительности. Коммутаторы $[H_\mu, H'_\nu]$ и $[H_\mu, H'_\perp]$ (обозначения будут объяснены ниже) найдены из соображений симметрии, а коммутатор $[H_\perp, H'_\perp]$ впервые вычислен явно.

22.1 Лагранжиан Гильберта–Эйнштейна

Рассмотрим пространство-время $\mathbb{M}$ произвольной размерности, $\dim \mathbb{M} = n \geq 3$. Локальные координаты, как и ранее, мы нумеруем греческими буквами из начала алфавита x^α, $\alpha = 0, 1, \ldots, n-1$. Мы предполагаем, что на $\mathbb{M}$ задана метрика $g_{\alpha\beta}$ лоренцевой сигнатуры, а кручение и тензор неметричности тождественно равны нулю (псевдориманова геометрия). Как и ранее мы предполагаем, что координата x^0 является временем, и все сечения $x^0 = \mathsf{const}$ пространственноподобны. Уравнения движения для метрики в общей теории относительности при отсутствии полей материи следуют из вариационного принципа для действия Гильберта–Эйнштейна (21.6)

$$S_{\text{НЕ}} = \int d^n x \sqrt{|g|} R, \qquad \sqrt{|g|} = \mathsf{det}\, e_\alpha{}^a = \sqrt{|\,\mathsf{det}\, g_{\alpha\beta}|}, \tag{22.1}$$

где R – псевдориманова скалярная кривизна, построенная по метрике $g_{\alpha\beta}$. В настоящей главе мы опускаем знак тильды для геометрических объектов псевдримановой геометрии, чтобы упростить обозначения. Гравитационную постоянную перед действием Гильберта–Эйнштейна мы положили равной единице, $\kappa = 1$, поскольку уравнения не меняются при умножении полного действия на постоянную. Космологическую постоянную мы пока положим равной нулю, $\Lambda = 0$, так как ее учет приводит только к изменению потенциала для метрики и не вызывает трудностей при канонической формулировке общей теории относительности.

Действие Гильберта–Эйнштейна в виде (22.1) содержит вторые производные от метрики, что создает определенные трудности для перехода к каноническому формализму. Однако, из явного вида тензора кривизны (**??**) следует, что вторые производные входят линейно. Поэтому от них можно избавиться, добавив к подынтегральному выражению полную производную. При этом мы потеряем явную ковариантность, зато действие будет зависеть только от первых производных метрики. С этой целью представим скалярную кривизну в виде двух слагаемых, что проверяется прямой проверкой,

$$R = \partial\Gamma - L_{\text{НЕ}}, \tag{22.2}$$

где

$$L_{\text{НЕ}} = g^{\alpha\beta}(\Gamma_{\alpha\beta}{}^\gamma \Gamma_\gamma - \Gamma_{\alpha\delta}{}^\gamma \Gamma_{\beta\gamma}{}^\delta), \tag{22.3}$$

$$\partial\Gamma = g^{\alpha\beta}(\partial_\alpha \Gamma_\beta - \partial_\gamma \Gamma_{\alpha\beta}{}^\gamma) = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \Gamma_\beta - \sqrt{|g|} g^{\beta\gamma} \Gamma_{\beta\gamma}{}^\alpha) + 2 L_{\text{НЕ}}. \tag{22.4}$$

Отсюда следует, что лагранжиан Гильберта–Эйнштейна $\sqrt{|g|} L_{\text{НЕ}}$ отличается от скалярной кривизны, умноженной на определитель репера, на полную производную

$$\sqrt{|g|} R - \partial_\alpha (\sqrt{|g|} g^{\alpha\beta} \Gamma_\beta - \sqrt{|g|} g^{\beta\gamma} \Gamma_{\beta\gamma}{}^\alpha) = \sqrt{|g|} L_{\text{НЕ}},$$

и, следовательно, приводит к тем же уравнениям Эйлера–Лагранжа. Он квадратичен по символам Кристоффеля и поэтому зависит только от первых производных метрики.

В разделе 21.2 действие Гильберта–Эйнштейна варьировалось по метрике и связности, которые рассматривались, как независимые переменные. В общей теории относительности этот подход называется *формализмом первого порядка*. В *формализме второго порядка* действие Гильберта–Эйнштейна $S_{\text{НЕ}}$ варьируется только по метрике. Этот подход более трудоемок для получения уравнений Эйнштейна. Чтобы его несколько упростить, полезно записать действие в виде интеграла от лагранжиана Гильберта–Эйнштейна $\sqrt{|g|} L_{\text{НЕ}}$, который не содержит вторых производных.

22.2 АДМ параметризация метрики и репера

При анализе гамильтоновой структуры уравнений общей теории относительности Арновитт, Дезер и Мизнер (АДМ) [?] использовали специальную параметризацию метрики, которая существенно упростила вычисления. В настоящем разделе будет описана АДМ параметризация метрики и соответствующая параметризация репера, которая упрощает вычисления в переменных Картана.

Рассмотрим многообразие $\mathbb{M}$, $\dim \mathbb{M} = n$, с метрикой лоренцевой сигнатуры $(+ - \ldots -)$. Пусть $\{x^\alpha\}$, $\alpha = 0, 1, \ldots, n-1$, – система локальных координат. Выделим среди координат время $t = x^0$, тогда $\{x^\alpha\} = \{x^0, x^\mu\}$, $\mu = 1, \ldots, n-1$. В дальнейшем буквы из начала греческого алфавита $\alpha, \beta, \ldots$ будут пробегать все значения индексов, а из середины $\mu, \nu, \ldots$ – только пространственные значения. Это правило легко запомнить по следующим включениям: $\{\mu, \nu, \ldots\} \subset \{\alpha, \beta, \ldots\}$ и $\{1, 2, \ldots\} \subset \{0, 1, 2, \ldots\}$. АДМ параметризация метрики имеет следующий вид

$$g_{\alpha\beta} = \begin{pmatrix} N^2 + N^\rho N_\rho & N_\nu \\ N_\mu & g_{\mu\nu} \end{pmatrix}, \tag{22.5}$$

где $g_{\mu\nu}$ – метрика на $(n-1)$-мерных сечениях многообразия $x^0 = \mathsf{const}$. В выбранной параметризации вместо n компонент метрики, содержащих хотя бы один временной индекс: g_{00} и $g_{0\mu}$, введены n функций N и N_μ. Выше $N^\rho := \hat{g}^{\rho\mu} N_\mu$, где $\hat{g}^{\rho\mu}$ – $(n-1) \times (n-1)$-матрица, обратная к $g_{\mu\nu}$:

$$\hat{g}^{\rho\mu} g_{\mu\nu} = \delta^\rho_\nu,$$

которую мы называем обратной метрикой. В настоящей главе подъем пространственных индексов будет всегда осуществляться с помощью обратной метрики $\hat{g}^{\rho\mu}$, помеченной шляпкой, которая в общем случае не совпадает с "пространственной" частью метрики $g^{\alpha\beta}$, обратной к $g_{\alpha\beta}$: $\hat{g}^{\rho\mu} \neq g^{\rho\mu}$. Функция $N = N(x)$ называется *функцией хода*, а функции $N_\mu = N_\mu(x)$ – *функциями сдвига*. Не ограничивая общности, можно считать, что функция хода положительна $N > 0$. В этом случае АДМ параметризация метрики (22.5) является взаимно однозначной.

Интервал, соответствующий параметризации (22.5), можно записать в виде

$$ds^2 = N^2 dt^2 + g_{\mu\nu}(dx^\mu + N^\mu dt)(dx^\nu + N^\nu dt).$$

Вообще, квадрат произвольного вектора $X = X^\alpha \partial_\alpha$ в АДМ параметризации метрики (22.5) имеет вид суммы двух квадратов:

$$X^2 := g_{\alpha\beta} X^\alpha X^\beta = N^2 (X^0)^2 + g_{\mu\nu}(X^\mu + N^\mu X^0)(X^\nu + N^\nu X^0),$$

где выделен квадрат временно́й компоненты.

Мы предполагаем, что координата $x^0 = t$ является временем, т.е. касательный вектор ∂_0 к координатной линии x^0 времениподобен. Формально это условие записывается в виде

$$(\partial_0, \partial_0) = g_{00} = N^2 + N^\rho N_\rho > 0. \tag{22.6}$$

В этом случае метрика $g_{\alpha\beta}$ имеет лоренцеву сигнатуру тогда и только тогда, когда матрица (теорема 5.2.1)

$$g_{\mu\nu} - \frac{N_\mu N_\nu}{N^2 + N^\rho N_\rho} \tag{22.7}$$

отрицательно определена. Отметим, что сама метрика $g_{\mu\nu}$, индуцированная на сечении $x^0 = \mathsf{const}$, может и не быть отрицательно определенной. Это значит, что сечение $x^0 = \mathsf{const}$ в общем случае не является пространственноподобным. В дальнейшем мы будем дополнительно предполагать, что координаты выбраны таким образом, чтобы все сечения $x^0 = \mathsf{const}$ были пространственноподобны, т.е. метрика $g_{\mu\nu}$ также отрицательно определена. Это удобно для постановки задачи Коши, когда начальные данные задаются на пространственноподобной поверхности, и рассматривается их эволюция во времени.

Замечание. Аналогичным образом можно параметризовать метрику и на римановом многообразии с положительно определенной метрикой, для этого вместо времени достаточно явно выделить произвольную координату. □

Метрика, обратная к (22.5), имеет вид

$$g^{\alpha\beta} = \begin{pmatrix} \dfrac{1}{N^2} & -\dfrac{N^\nu}{N^2} \\ -\dfrac{N^\mu}{N^2} & \hat{g}^{\mu\nu} + \dfrac{N^\mu N^\nu}{N^2} \end{pmatrix}. \tag{22.8}$$

Предложение 22.2.1. *Пространственная матрица в правом нижнем блоке*

$$g^{\mu\nu} = \hat{g}^{\mu\nu} + \frac{N^\mu N^\nu}{N^2}, \tag{22.9}$$

отрицательно определена.

Доказательство. Легко проверить, матрица (22.9) является обратной к метрике (22.7). Это значит, что отрицательная определенность метрики (22.7) эквивалентна отрицательной определенности матрицы $g^{\mu\nu}$. □

Заметим, что, если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру, то условие пространственноподобности всех сечений $x^0 = \mathsf{const}$ эквивалентно условию $N^2 > 0$. Действительно, из отрицательной определенности $g_{\mu\nu}$ следует отрицательная определенность обратной матрицы $\hat{g}^{\mu\nu}$. Тогда из уравнения (22.9) вытекает отрицательная определенность матрицы

$$g^{\mu\nu} - \frac{N^\mu N^\nu}{N^2}.$$

Это, в свою очередь, эквивалентно условию $g^{00} > 0$ или $N^2 > 0$.

Чтобы продемонстрировать тонкости, которые могут возникнуть при АДМ параметризации метрики, приведем простой

Пример 22.2.1. Рассмотрим двумерное пространство-время Минковского $\mathbb{R}^{1,1}$ с инерциальными координатами t, x. Введем новую систему координат ξ, η, зависящую от двух вещественных параметров a и b (см. рис.**??**)

$$\xi = t + ax, \quad \eta = t - bx, \qquad |a| \neq 1, \quad |b| \neq 1, \quad a + b \neq 0.$$

Легко получить формулы обратного преобразования

$$t = \frac{b\xi + a\eta}{a + b}, \qquad x = \frac{\xi - \eta}{a + b}.$$

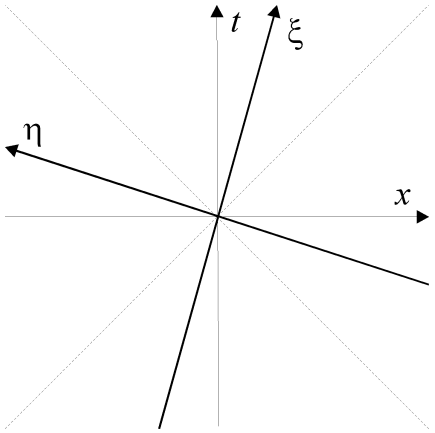


Рис. 22.1: Выбор координат на плоскости Минковского. Рисунок соответствует следующим значениям параметров: $0 < a < 1$ и $b > 1$.

В новых координатах лоренцева метрика имеет вид

$$ds^2 = dt^2 - dx^2 = \frac{1}{(a+b)^2}\left[(b^2-1)d\xi^2 + 2(ab+1)d\xi d\eta + (a^2-1)d\eta^2\right].$$

Векторы, касательные к координатным линиям ξ и η имеют следующие компоненты:

$$\partial_\xi = \frac{b}{a+b}\partial_t + \frac{1}{a+b}\partial_x, \qquad \partial_\eta = \frac{a}{a+b}\partial_t - \frac{1}{a+b}\partial_x.$$

Проанализируем АДМ параметризацию метрики в координатах $x^0 = \xi,\ x^1 = \eta$. Метрика имеет следующие компоненты:

$$g_{00} = \frac{b^2-1}{(a+b)^2}, \qquad g_{01} = \frac{ab+1}{(a+b)^2}, \qquad g_{11} = \frac{a^2-1}{(a+b)^2}.$$

Функции хода и сдвига имеют вид

$$N^2 = -\frac{1}{a^2-1}, \qquad N_1 = \frac{ab+1}{(a+b)^2}.$$

Из условий $g_{00} > 0$ и $g_{11} < 0$ следует, соответственно, что $|b| > 1$ и $|a| < 1$. Мы видим, что эти условия являются необходимыми и достаточными для того, чтобы координатная линия ξ была времениподобной, а η – пространственноподобной. Нетрудно также проверить эквивалентность условий:

$$g_{00} > 0 \qquad \Leftrightarrow \qquad g_{11} - \frac{N_1 N_1}{N^2 + N^1 N_1} = -\frac{1}{b^2-1} < 0,$$

а также условий:

$$g^{00} > 0 \qquad \Leftrightarrow \qquad \hat{g}^{11} = g^{11} - \frac{N^1 N^1}{N^2} = \frac{(a+b)^2}{a^2-1} < 0. \qquad \square$$

Обратная метрика (22.8) удобна в приложениях тем, что квадрат произвольного ковектора (1-формы) $A = dx^\alpha A_\alpha$ имеет вид суммы двух квадратов:

$$A^2 := g^{\alpha\beta}A_\alpha A_\beta = \frac{1}{N^2}(A_0 - N^\mu A_\mu)^2 + \hat{g}^{\mu\nu}A_\mu A_\nu,$$

где выделен квадрат пространственных компонент.

Используя формулу (28.7) для определителя блочных матриц, получаем выражение для определителя метрики (22.5)

$$\det g_{\alpha\beta} = N^2 \det g_{\mu\nu}. \tag{22.10}$$

Отсюда следует выражение для элемента объема

$$\sqrt{|g|} = N\sqrt{|\hat{g}|}, \qquad \sqrt{|g|} := \sqrt{|\det g_{\alpha\beta}|}, \qquad \sqrt{|\hat{g}|} := \sqrt{|\det g_{\mu\nu}|}. \tag{22.11}$$

Последняя формула является обобщением хорошо известной из школьного курса правила: объем призмы равен произведению площади основания на высоту. В рассматриваемом случае площадью основания является $\sqrt{|\hat{g}|}$, а высотой – функция хода N.

При проведении вычислений оказываются полезными следующие формулы, которые проверяются прямой проверкой,

$$\begin{aligned} g^{00}g^{\mu\nu} - g^{0\mu}g^{0\nu} &= \frac{\hat{g}^{\mu\nu}}{N^2}, \\ g^{\sigma\mu}g^{0\nu} - g^{\sigma\nu}g^{0\mu} &= \frac{N^\mu \hat{g}^{\sigma\nu} - N^\nu \hat{g}^{\sigma\mu}}{N^2}, \\ g^{\mu\nu}g_{\nu\sigma} &= \delta^\mu_\sigma + \frac{N^\mu N_\sigma}{N^2}, \\ g^{\mu\nu}g_{\mu\nu} &= n - 1 + \frac{N^\mu N_\mu}{N^2}. \end{aligned} \tag{22.12}$$

В аффинной геометрии вместо метрики $g_{\alpha\beta}$ и аффинной связности $\Gamma_{\alpha\beta}{}^\gamma$ удобно использовать переменные Картана: репер $e_\alpha{}^a$ и $\mathbb{GL}(n,\mathbb{R})$-связность $\omega_{\alpha a}{}^b$, где $a, b, \ldots = 0, 1, \ldots, n-1$ (см. главу **??**). Здесь и в дальнейшем мы примем следующее обозначение. Если латинский индекс принимает значение нуль, то мы будем писать его курсивом $a = 0$. Это правило принято потому что в дальнейшем значение нуль будут принимать как греческие, так и латинские индексы, что необходимо различать. Выделим в ортонормальном базисе касательного пространства временно́й вектор e_0, $\{e_a\} = \{e_0, e_i\}$, тогда метрика Минковского примет вид

$$\eta_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & \eta_{ij} \end{pmatrix}, \qquad i, j = 1, \ldots, n-1,$$

где $\eta_{ij} = -\delta_{ij}$ – евклидова метрика с обратным знаком для пространственных компонент тензоров. Как и раньше, мы принимаем следующее соглашение. Латинские индексы из начала алфавита $a, b, \ldots$ пробегают все значения, а индексы из середины алфавита $i, j, \ldots$ – только пространственные. Поскольку репер содержит большее число независимых компонент, чем метрика, то его параметризация, соответствующая выделенной временно́й координате, является более громоздкой:

$$e_\alpha{}^a = \begin{pmatrix} N\sqrt{1-M^2} + (MN) & N^i - NM^i + \dfrac{\sqrt{1-M^2}-1}{M^2}(MN)M^i \\ \\ M_\mu & \hat{e}_\mu{}^i + \dfrac{\sqrt{1-M^2}-1}{M^2} M_\mu M^i \end{pmatrix}, \tag{22.13}$$

где M_μ – дополнительное $(n-1)$-компонентное пространственноподобное векторное поле (относительно трехмерных диффеоморфизмов). Ниже будет показано, что это

поле пропорционально трехмерному вектору скорости, параметризующему преобразования Лоренца. Репер со шляпкой в (22.13) $\hat{e}_\mu{}^i$ на сечении $x^0 = \mathsf{const}$ определяется трехмерной метрикой

$$g_{\mu\nu} = \hat{e}_\mu{}^i \hat{e}_\nu{}^j \eta_{ij}, \tag{22.14}$$

В параметризации репера (22.13) использованы также обозначения $(MN) := M^\mu N_\mu = M^i N_i$ и $M^2 := M^\mu M_\mu = M^i M_i < 0$. Переход от пространственных латинских индексов к пространственным греческим индексам осуществляется с помощью репера $\hat{e}_\mu{}^i$ и его обратного $\hat{e}^\mu{}_i$, которые мы пометили шляпкой,

$$\hat{g}^{\mu\nu} = \hat{e}^\mu{}_i \hat{e}^\nu{}_j \eta^{ij}.$$

Например, $M^i = M_\mu \hat{e}^{\mu i}$. Как и для метрики, реперы со шляпкой и без нее необходимо различать: $\hat{e}_\mu{}^i \neq e_\mu{}^i$, $\hat{e}^\mu{}_i \neq e^\mu{}_i$. Из определения репера со шляпкой (22.14) следует равенство для определителей

$$\mathsf{det}\, \hat{e}_\mu{}^i = \sqrt{|\,\mathsf{det}\, g_{\mu\nu}|},$$

что согласуется с обозначением в (22.11).

С помощью прямых вычислений нетрудно проверить, что приведенная параметризация репера соответствует АДМ параметризации метрики (22.5)

$$N^2 + N^\rho N_\rho = e_0{}^a e_{0a}, \qquad N_\mu = e_0{}^a e_{\mu a}, \qquad g_{\mu\nu} = e_\mu{}^a e_{\nu a}.$$

Обратный репер имеет вид

$$e^\alpha{}_a = \begin{pmatrix} \dfrac{\sqrt{1-M^2}}{N} & -\dfrac{M_i}{N} \\ M^\mu - \dfrac{\sqrt{1-M^2}}{N} N^\mu & \hat{e}^\mu{}_i + \dfrac{N^\mu M_i}{N} + \dfrac{\sqrt{1-M^2}-1}{M^2} M^\mu M_i. \end{pmatrix} \tag{22.15}$$

Чтобы понять геометрический смысл параметров M_i, рассмотрим преобразования касательного пространства. Связная компонента единицы группы Лоренца $\mathbb{SO}_\uparrow(1, n-1)$ имеет подгруппу вращений $\mathbb{SO}(n-1)$, действующую на пространственные компоненты репера $\hat{e}_\mu{}^i$. Эта подгруппа параметризуется $(n-1)(n-2)/2$ параметрами. Остальные $n-1$ параметров соответствуют лоренцевым бустам, которые параметризуются функциями M_i. В этом нетрудно убедиться. Представим репер (22.13) в виде

$$e_\alpha{}^a = \mathring{e}_\alpha{}^b S_b{}^a,$$

где

$$\mathring{e}_\alpha{}^a = \begin{pmatrix} N & N^i \\ 0 & \hat{e}_\mu{}^i \end{pmatrix} \tag{22.16}$$

– фиксированный репер, который преобразуется с помощью лоренцевых бустов

$$S_b{}^a = \begin{pmatrix} \sqrt{1-M^2} & -M^j \\ M_i & S_i{}^j \end{pmatrix} \quad \in \mathbb{SO}_\uparrow(1, n-1), \tag{22.17}$$

где введено обозначение для пространственной части матрицы лоренцевых вращений

$$S_i{}^j = \delta_i^j + \frac{\sqrt{1-M^2}-1}{M^2} M_i M^j \tag{22.18}$$

Можно проверить, что матрица (22.17) действительно задает преобразования Лоренца. Репер, обратный к (22.16), имеет вид

$$\overset{\circ}{e}{}^{\alpha}{}_{a} = \begin{pmatrix} \dfrac{1}{N} & 0 \\ -\dfrac{N^{\mu}}{N} & \hat{e}^{\mu}{}_{i} \end{pmatrix}.$$

Формулы для фиксированного репера $\overset{\circ}{e}{}_{\alpha}{}^{a}$ и его обратного $\overset{\circ}{e}{}^{\alpha}{}_{a}$ получаются из (22.13) и (22.15) при $M_i = 0$.

В дальнейшем матрица пространственных вращений (22.18) часто используется в вычислениях, поэтому приведем для нее несколько полезных соотношений, которые проверяются прямой проверкой,

$$\begin{aligned} S^{-1}{}_{i}^{j} &= \delta_i^j - \frac{\sqrt{1-M^2}-1}{M^2\sqrt{1-M^2}} M_i M^j \\ S_i{}^j(M) &= S_i{}^j(-M), \qquad S_{ij} = S_{ji}, \\ S_i{}^j M_j &= \sqrt{1-M^2} M_i, \\ S_i{}^k S_{jk} &= \eta_{ij} - M_i M_j, \\ S^{-1}{}_{i}^{k} S^{-1}{}_{jk} &= \eta_{ij} + \frac{M_i M_j}{\sqrt{1-M^2}}. \end{aligned}$$

Таким образом вместо n^2 независимых компонент репера мы ввели функцию сдвига N (1 компонента), функции хода N^i ($n-1$ компонент), лоренцевы бусты M_i ($n-1$ компонент) и пространственный репер $\hat{e}_{\mu}{}^{i}$ ($(n-1)^2$ компонент). Эта параметризация при $N > 0$ является взаимно однозначной.

Отметим, что единичный нормальный вектор к сечению $x^0 = \mathsf{const}$ (21.157) относительно неголономного базиса $e_a = e^{\mu}{}_{a}\partial_{\mu}$ имеет вид

$$n = \sqrt{1-M^2} e_0 - M^i e_i.$$

То есть его компоненты относительно неголономного базиса полностью определяются лоренцевыми бустами.

Приведем более привычную параметризацию лоренцевых вращений $S_a{}^b$ с помощью вектора скорости. Вектор M^i взаимно однозначно параметризуется вектором скорости V^i:

$$M^i = \frac{V^i}{\sqrt{1-\boldsymbol{V}^2}} \quad \Leftrightarrow \quad V^i = \frac{M^i}{\sqrt{1-M^2}},$$

где $\boldsymbol{V}^2 := -V^i V_i$. По-определению, $M^2 < 0$ и, следовательно, $0 < \boldsymbol{V}^2 < 1$. Тогда

$$S_a{}^b = \begin{pmatrix} \dfrac{1}{\sqrt{1-\boldsymbol{V}^2}} & -\dfrac{V^j}{\sqrt{1-\boldsymbol{V}^2}} \\ \dfrac{V_i}{\sqrt{1-V^2}} & \delta_i^j - \dfrac{V_i V^j}{\boldsymbol{V}^2}\left(\dfrac{1}{\sqrt{1-V^2}} - 1\right) \end{pmatrix}.$$

Эта матрица для лоренцевых бустов была получена ранее (2.200).

Замечание. В римановом пространстве с положительно определенной метрикой репер можно параметризовать похожим образом, заменив матрицу лоренцевых вращений на ортогональную матрицу. □

22.3 Геометрия гиперповерхностей

При гамильтоновой формулировке моделей гравитации мы рассматриваем пространство-время как семейство пространственноподобных гиперповерхностей, которое параметризуется временем. Другими словами для каждого момента времени пространство представляет собой гиперповерхность, вложенную в пространство-время. Поскольку уравнения моделей гравитации определяют геометрию всего пространства-времени, то полезно знать, какая геометрия возникает при этом на пространственноподобных сечениях. В настоящем разделе мы подойдем к этому вопросу с общей точки зрения, предполагая, что на объемлющем многообразии задана аффинная геометрия общего вида, т.е. независимые метрика и связность, не предполагая лоренцевой сигнатуры метрики.

Рассмотрим $(n-1)$-мерную гиперповерхность $\mathbb{U}$, вложенную в n-мерное многообразие $\mathbb{M}$:

$$f: \quad \mathbb{U} \hookrightarrow \mathbb{M}. \tag{22.19}$$

Обозначим координаты на $\mathbb{M}$ и $\mathbb{U}$ соответственно через x^α, $\alpha = 1, \dots, n$, и u^i, $i = 1, \dots, n-1$. Тогда вложение $\mathbb{U}$ в $\mathbb{M}$ локально задается n функциями $x^\alpha(u)$, которые предполагаются достаточно гладкими. Произвольное векторное поле $X = \{X^i\} \in \mathcal{X}(\mathbb{U})$, на гиперповерхности, отображается в векторное поле $Y = \{Y^\alpha\} \in \mathcal{X}(\mathbb{M})$ на $\mathbb{M}$ с помощью дифференциала отображения

$$f_*: \quad \mathcal{X}(\mathbb{U}) \ni X = X^i \partial_i \;\mapsto\; Y = Y^\alpha \partial_\alpha \in \mathcal{X}(\mathbb{M}),$$

где

$$Y^\alpha = e^\alpha{}_i X^i, \qquad e^\alpha{}_i := \partial_i x^\alpha.$$

Матрица Якоби $e^\alpha{}_i$ отображения f прямоугольна, имеет размер $n \times (n-1)$, ранг $n-1$ и, естественно, необратима. Возврат отображения f каждому ковекторному полю (1-форме) на $\mathbb{M}$ ставит в соответствие ковекторное поле на $\mathbb{U}$:

$$f^*: \quad \Lambda_1(\mathbb{M}) \ni A = dx^\alpha A_\alpha \;\mapsto\; B = du^i B_i \in \Lambda_1(\mathbb{U}),$$

где

$$B_i = A_\alpha e^\alpha{}_i.$$

В дальнейшем мы будем отождествлять гиперповерхность $\mathbb{U}$ с ее образом: $\mathbb{U} = f(\mathbb{U}) \subset \mathbb{M}$.

1-форма $n = dx^\alpha n_\alpha$, определяемая системой алгебраических уравнений

$$n_\alpha e^\alpha{}_i = 0, \qquad i = 1, \dots, n-1, \tag{22.20}$$

ортогональна гиперповерхности $\mathbb{U}$ и задает на многообразии $\mathbb{M}$ распределение $(n-1)$-мерных подпространств, касательных к $\mathbb{U}$. Уравнения (22.20) имеют единственное нетривиальное решение с точностью до умножения на произвольную отличную от нуля функцию, поскольку из определения вложения следует, что ранг матрицы Якоби равен $n-1$.

Матрица Якоби $e^\alpha{}_i$ задает в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, где $x \in \mathbb{U} \subset \mathbb{M}$, набор из $n-1$ векторов $e_i = e^\alpha{}_i \partial_\alpha$, которые образуют базис касательного пространства к гиперповерхности $\mathbb{U} \subset \mathbb{M}$.

Это все, что можно сказать о гиперповерхности $\mathbb{U}$, если задано только вложение (22.19). Теория становится намного более содержательной, если на многообразии $\mathbb{M}$ заданы дополнительные структуры. Остановимся на этом вопросе подробно.

Пусть на $\mathbb{M}$ задана аффинная геометрия, т.е. метрика $g_{\alpha\beta}$ и аффинная связность $\Gamma_{\alpha\beta}{}^{\gamma}$. Рассмотрим, какая геометрия возникает на гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$. Возврат отображения f^* индуцирует на гиперповерхности единственную метрику:

$$f^* : \quad g_{\alpha\beta} \;\mapsto\; g_{ij} = g_{\alpha\beta} e^{\alpha}{}_{i} e^{\beta}{}_{j}. \tag{22.21}$$

Для многообразий с метрикой лоренцевой сигнатуры мы предполагаем, что гиперповерхность $\mathbb{U}$ пространственноподобна и, следовательно, индуцированная метрика невырождена (отрицательно определена). Наличие двух метрик $g_{\alpha\beta}$ и g_{ij} соответственно на $\mathbb{M}$ и $\mathbb{U}$ позволяет опускать и поднимать индексы у матрицы Якоби:

$$e_{\alpha}{}^{i} := g_{\alpha\beta} e^{\beta}{}_{j} g^{ij}.$$

Эта матрица проектирует произвольный вектор из $\mathbb{T}_{f(u)}(\mathbb{M})$ в касательное пространство к гиперповерхности $\mathbb{T}_u(\mathbb{U})$

$$\mathbb{T}_{f(u)}(\mathbb{M}) \ni \{X^{\alpha}\} \;\mapsto\; \{X^{i} := X^{\alpha} e_{\alpha}{}^{i}\} \in \mathbb{T}_u(\mathbb{U}). \tag{22.22}$$

Теперь определим связность на гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$, спроектировав ковариантную производную в $\mathbb{M}$ на гиперповерхность $\mathbb{U}$:

$$\hat{\nabla}_i X^k := (\nabla_{\alpha} X^{\beta}) e^{\alpha}{}_{i} e_{\beta}{}^{k}, \tag{22.23}$$

где компоненты X^i определены отображением (22.22).

Замечание. Из приведенного определения нельзя выразить ковариантную производную $\nabla_{\alpha} X^{\beta}$ в $\mathbb{M}$ через ковариантную производную $\hat{\nabla}_i X^k$ на $\mathbb{U}$, т.к. оригинал нельзя восстановить по его проекции. □

Раскрытие равенства (22.23) приводит к выражению для компонент индуцированной связности на гиперповерхности $\mathbb{U}$:

$$\hat{\Gamma}_{ij}{}^{k} = (\partial^2_{ij} x^{\gamma} + \Gamma_{\alpha\beta}{}^{\gamma} e^{\alpha}{}_{i} e^{\beta}{}_{j}) e_{\gamma}{}^{k}. \tag{22.24}$$

Эта связность единственна. Отметим, что если исходная связность $\Gamma_{\alpha\beta}{}^{\gamma}$ симметрична, то индуцированная связность $\hat{\Gamma}_{ij}{}^{k}$ также симметрична. Кроме того, связность на $\mathbb{U}$ определяется единственным образом только в том случае, если на $\mathbb{M}$ помимо связности задана также метрика.

Из уравнения (22.24) следует, что тензор кручения на $\mathbb{M}$ индуцирует кручение на гиперповерхности:

$$T_{ij}{}^{k} = T_{\alpha\beta}{}^{\gamma} e^{\alpha}{}_{i} e^{\beta}{}_{j} e_{\gamma}{}^{k}. \tag{22.25}$$

Прямые вычисления дают следующее выражение для ковариантной производной от индуцированной метрики на гиперповерхности

$$\hat{\nabla}_i g_{jk} = \partial_i g_{ij} - \hat{\Gamma}_{ij}{}^{l} g_{lk} - \hat{\Gamma}_{ik}{}^{l} g_{jl} = (\nabla_{\alpha} g_{\beta\gamma}) e^{\alpha}{}_{i} e^{\beta}{}_{j} e^{\gamma}{}_{k}.$$

Отсюда вытекает выражение для тензора неметричности на гиперповерхности

$$Q_{ijk} = Q_{\alpha\beta\gamma} e^{\alpha}{}_{i} e^{\beta}{}_{j} e^{\gamma}{}_{k}.$$

В частности, если связность $\Gamma_{\alpha\beta}{}^{\gamma}$ на $\mathbb{M}$ является метрической, то и индуцированная связность $\hat{\Gamma}_{ij}{}^{k}$ на $\mathbb{U}$ также метрическая.

Таким образом, метрика $g_{\alpha\beta}$ и связность $\Gamma_{\alpha\beta}{}^{\gamma}$ на $\mathbb{M}$ индуцируют единственные метрику g_{ij} и связность $\hat{\Gamma}_{ij}{}^{k}$ на гиперповерхности $\mathbb{U} \subset \mathbb{M}$. Обратное утверждение, конечно, неверно. Если метрика и связность заданы на гиперповерхности $\mathbb{U}$, то они не индуцируют на $\mathbb{M}$ геометрию единственным образом. Это понятно, поскольку размерность гиперповерхности меньше размерности самого многообразия.

В дальнейшем все геометрические объекты, относящиеся только к гиперповерхности и построенные по индуцированной метрике g_{ij} и связности $\hat{\Gamma}_{ij}{}^{k}$ мы будем отмечать шляпкой.

Наличие метрики $g_{\alpha\beta}$ позволяет построить единичное векторное поле $n = n^{\alpha}\partial_{\alpha}$, ортогональное к гиперповерхности. Как уже отмечалось, система уравнений $n_{\alpha}e^{\alpha}{}_{i} = 0$ определяет 1-форму $dx^{\alpha}n_{\alpha}$ с точностью до умножения на произвольную функцию. Воспользуемся этим произволом для того, чтобы в каждой точке вектор с компонентами $n^{\alpha} := g^{\alpha\beta}n_{\beta}$ имел единичную длину $(n,n) = n^{\alpha}n^{\beta}g_{\alpha\beta} = 1$. По-построению, этот вектор ортогонален всем векторам, касательным к гиперповерхности:

$$(n, e_i) = n^{\alpha}e^{\beta}{}_{i}g_{\alpha\beta} = n_{\alpha}e^{\alpha}{}_{i} = 0.$$

Если на многообразии задана гиперповерхность, то естественно рассматривать базис $\{n, e_i\}$ в касательном пространстве $\mathbb{T}_x(\mathbb{M})$, $x \in \mathbb{U}$, определяемый этой гиперповерхностью. Ему соответствует сопряженный базис $\{n := dx^{\alpha}n_{\alpha}, e^{i} := dx^{\alpha}e_{\alpha}{}^{i}\}$ в кокасательном пространстве $\mathbb{T}^{*}_{x}(\mathbb{M})$. Тогда произвольный вектор X и 1-форма A раскладываются по этим базисам:

$$\begin{aligned} X^{\alpha} &= X^{\perp}n^{\alpha} + X^{i}e^{\alpha}{}_{i}, \qquad X^{\perp} = X^{\alpha}n_{\alpha}, \quad X^{i} = X^{\alpha}e_{\alpha}{}^{i}, \\ A_{\alpha} &= A_{\perp}n_{\alpha} + A_{i}e_{\alpha}{}^{i}, \qquad A_{\perp} = A_{\alpha}n^{\alpha}, \quad A_{i} = A_{\alpha}e^{\alpha}{}_{i}. \end{aligned}$$

Аналогично можно разложить тензор произвольного ранга. В частности, разложение ковариантного тензора второго ранга имеет вид

$$A_{\alpha\beta} = A_{\perp\perp}n_{\alpha}n_{\beta} + A_{\perp i}n_{\alpha}e^{i}{}_{\beta} + A_{i\perp}e^{i}{}_{\alpha}n_{\beta} + A_{ij}e^{i}{}_{\alpha}e^{j}{}_{\beta},$$

где

$$A_{\perp\perp} = A_{\alpha\beta}n^{\alpha}n^{\beta}, \quad A_{\perp i} = A_{\alpha\beta}n^{\alpha}e^{\beta}{}_{i}, \quad A_{i\perp} = A_{\alpha\beta}e^{\alpha}{}_{i}n^{\beta}, \quad A_{ij} = A_{\alpha\beta}e^{\alpha}{}_{i}e^{\beta}{}_{j}.$$

Нетрудно проверить, что разложение для метрики гораздо проще:

$$g_{\alpha\beta} = n_{\alpha}n_{\beta} + e_{\alpha}{}^{i}e_{\beta}{}^{j}g_{ij}. \tag{22.26}$$

Разложение для обратной метрики имеет аналогичный вид:

$$g^{\alpha\beta} = n^{\alpha}n^{\beta} + e^{\alpha}{}_{i}e^{\beta}{}_{j}g^{ij}. \tag{22.27}$$

Из определения обратной метрики $g^{\alpha\beta}g_{\beta\gamma} = \delta^{\alpha}_{\gamma}$ следует правило суммирования матрицы Якоби по латинским индексам:

$$e^{\alpha}{}_{i}e_{\beta}{}^{i} = \delta^{\alpha}_{\beta} - n^{\alpha}n_{\beta}. \tag{22.28}$$

Нетрудно также проверить, что из определения обратной индуцированной метрики $g^{ij}g_{jk} = \delta^{i}_{k}$ следует равенство

$$e^{\alpha}{}_{i}e_{\alpha}{}^{j} = \delta^{j}_{i}, \tag{22.29}$$

где суммирование проводится по греческим индексам. С учетом этого правила из равенства (22.27) следует представление для обратной индуцированной метрики

$$g^{ij} = g^{\alpha\beta} e_\alpha{}^i e_\beta{}^j.$$

Индуцированные метрика (22.21) и связность (22.24) задают внутреннюю геометрию гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$. В частности, индуцированная связность задает тензор внутренней кривизны гиперповерхности

$$\hat{R}_{ijk}{}^l(\hat{\Gamma}) = \partial_i \hat{\Gamma}_{jk}{}^l - \hat{\Gamma}_{ik}{}^m \hat{\Gamma}_{jm}{}^l - (i \leftrightarrow j).$$

Вложение гиперповерхности $f: \mathbb{U} \hookrightarrow \mathbb{M}$ позволяет определить еще один важный объект, характеризующий гиперповерхность.

Определение. Тензор второго ранга с компонентами

$$K_{ij} := -(\nabla_\alpha n_\beta) e^\alpha{}_i e^\beta{}_j, \tag{22.30}$$

определен на $\mathbb{U} \subset \mathbb{M}$ и называется *внешней кривизной* гиперповерхности. Он равен ковариантной производной нормали, спроектированной на касательное пространство к гиперповерхности и взятой с обратным знаком. □

В отличии от тензора внутренней кривизны внешняя кривизна является тензором второго, а не четвертого ранга и в общем случае никакой симметрии по индексам не имеет. Этот тензор характеризует изменение нормали при ее параллельном переносе вдоль кривой на гиперповерхности.

Раскрывая определение тензора внешней кривизны, с учетом равенства (22.20) получим, что

$$K_{ij} = n_\alpha (\partial^2_{ij} x^\alpha + \Gamma_{\beta\gamma}{}^\alpha e^\beta{}_i e^\gamma{}_j).$$

Отсюда следует, что антисимметричная часть тензора внешней кривизны определяется тензором кручения

$$K_{[ij]} = \frac{1}{2} n_\alpha T_{\beta\gamma}{}^\alpha e^\beta{}_i e^\gamma{}_j = \frac{1}{2} T_{ij}{}^\perp. \tag{22.31}$$

Это доказывает следующее

Предложение 22.3.1. *Внешняя кривизна гиперповерхности $\mathbb{U} \hookrightarrow \mathbb{M}$ симметрична тогда и только тогда, когда кручение связности $\Gamma_{\beta\gamma}{}^\alpha$ на $\mathbb{U} \subset \mathbb{M}$ удовлетворяет условию $T_{ij}{}^\perp = 0$.*

Вычислим ковариантную производную от матрицы Якоби

$$\nabla_i e^\alpha{}_j = e^\beta{}_i (\partial_\beta e^\alpha{}_j + \Gamma_{\beta\gamma}{}^\alpha e^\gamma{}_j) - \hat{\Gamma}_{ij}{}^k e^\alpha{}_k. \tag{22.32}$$

Разложение этого равенства (индекс α) по базису n, e_i показывает, что эта ковариантная производная имеет только нормальную составляющую и пропорциональна внешней кривизне:

$$\nabla_i e^\alpha{}_j = n^\alpha K_{ij}. \tag{22.33}$$

Введем специальное обозначение для "половины" ковариантной производной (22.32), которая содержит связность только для греческих индексов,

$$\check{\nabla}_i e^\alpha{}_j := e^\beta{}_i (\partial_\beta e^\alpha{}_j + \Gamma_{\beta\gamma}{}^\alpha e^\gamma{}_j).$$

Она ковариантна относительно преобразований координат x^α в $\mathbb{M}$, но не на $\mathbb{U}$. Разлагая ее по базису n, e_i, получим равенство

$$\check{\nabla}_i e^\alpha{}_j = n^\alpha K_{ij} + \hat{\Gamma}_{ij}{}^k e^\alpha{}_k. \tag{22.34}$$

Это соотношение известно как *формула Гаусса–Вейнгартена.*

Из равенств (22.33) и (22.34) следует ещё одно представление для тензора внешней кривизны

$$K_{ij} = n_\alpha \nabla_i e^\alpha{}_j = n_\alpha \check{\nabla}_i e^\alpha{}_j. \tag{22.35}$$

Полный тензор кривизны $R_{\alpha\beta\gamma\delta}$, спроектированный на гиперповерхность, можно выразить через тензор внутренней кривизны $\hat{R}_{ijkl}$, построенный только по индуцированной метрике (22.21) и связности (22.24), и тензор внешней кривизны. Для этого рассмотрим коммутатор ковариантных производных ковекторного поля, который определяется тензором кривизны и кручения (7.82)

$$\begin{aligned}[\nabla_\alpha, \nabla_\beta] X_\gamma &= -R_{\alpha\beta\gamma}{}^\delta X_\delta - T_{\alpha\beta}{}^\delta \nabla_\delta X_\delta = \\ &= -R_{\alpha\beta\gamma i} X^i - R_{\alpha\beta\gamma\perp} X^\perp - T_{\alpha\beta}{}^i \nabla_i X_\gamma - T_{\alpha\beta}{}^\perp \nabla_\perp X_\gamma,\end{aligned} \tag{22.36}$$

где в правой части сначала вычисляются ковариантные производные, а затем проектируются на гиперповерхность и ортогональное направления:

$$\nabla_i X_\gamma := e^\alpha{}_i \nabla_\alpha X_\gamma, \qquad \nabla_\perp X_\gamma := n^\alpha \nabla_\alpha X_\gamma.$$

Для того, чтобы спроектировать равенство (22.36) на гиперповерхность, спроектируем сначала ковариантную производную:

$$\nabla_i X_j := e^\alpha{}_i (\nabla_\alpha X_\beta) e^\beta{}_j = e^\alpha{}_i \nabla_\alpha (X_\beta e^\beta{}_j) - e^\alpha{}_i X_\beta \nabla_\alpha e^\beta{}_j = \hat{\nabla}_i X_j - X_\perp K_{ij},$$

где мы воспользовались равенством (22.33) и

$$\hat{\nabla}_i X_j := \partial_i X_j - \hat{\Gamma}_{ij}{}^k X_k$$

– $(n-1)$-мерная ковариантная производная на гиперповерхности. Аналогично проектируется вторая ковариантная производная:

$$\begin{aligned}\nabla_i \nabla_j X_k &= \hat{\nabla}_i (\nabla_j X_k) - \nabla_\perp X_k K_{ij} - \nabla_j X_\perp K_{ik} = \\ &= \hat{\nabla}_i \hat{\nabla}_j X_k - \hat{\nabla}_i X_\perp K_{jk} - X_\perp \hat{\nabla}_i K_{jk} - \nabla_\perp X_k K_{ij} - \hat{\nabla}_j X_\perp K_{ik} - X^l K_{lj} K_{ik},\end{aligned}$$

где

$$\nabla_j X_\perp := e^\alpha{}_j (\nabla_\alpha X_\beta) n^\beta = \hat{\nabla}_j X_\perp + X^l K_{lj}.$$

Антисимметризация полученного выражения по индексам i, j дает проекцию коммутатора (22.36) на гиперповерхность

$$\begin{aligned}[\nabla_i, \nabla_j] X_k &= -R_{ijkl} X^l - R_{ijk\perp} X^\perp - T_{ij}{}^l \nabla_l X_k - T_{ij}{}^\perp \nabla_\perp X_k = \\ &= -R_{ijkl} X^l - R_{ijk\perp} X^\perp - T_{ij}{}^l \hat{\nabla}_l X_k + T_{ij}{}^l X_\perp K_{lk} - T_{ij}{}^\perp \nabla_\perp X_k.\end{aligned}$$

Учитывая независимость компонент X^l и $X^\perp$ и выражения для компонент тензора кручения (22.25), (22.31), получаем выражение для проекций полного тензора кривизны на гиперповерхность:

$$R_{ijkl} = \hat{R}_{ijkl} + K_{ik} K_{lj} - K_{jk} K_{li}, \tag{22.37}$$

$$R_{ijk\perp} = \hat{\nabla}_i K_{jk} - \hat{\nabla}_j K_{ik} + T_{ij}{}^l K_{lk}. \tag{22.38}$$

Полученные соотношения являются обобщением *уравнений Гаусса–Петерсона–Кодацци* на случай, когда на объемлющем многообразии $\mathbb{M}$ задана не риманова геометрия, а произвольная аффинная геометрия с ненулевым кручением и неметричностью.

В заключение настоящего раздела вычислим нормальные компоненты $G_{\perp\perp}$ и $G_{\perp i}$ тензора Эйнштейна

$$G_{\alpha\beta} := \tilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \tilde{R},$$

которые играют важную роль при анализе уравнений движения общей теории относительности. Знак тильды в правой части, как и ранее, означает, что кручение и неметричность связности, заданной на $\mathbb{M}$, равны нулю. В этом случае тензор кривизны обладает дополнительной симметрией относительно перестановки индексов (см. раздел 7.8) и тензор внешней кривизны симметричен: $K_{ij} = K_{ji}$.

Ниже, для простоты, мы опустим знак тильды.

Сначала вычислим скалярную кривизну

$$R = g^{\alpha\gamma} g^{\beta\delta} R_{\alpha\beta\gamma\delta} = 2R_{\perp\perp} + g^{ik} g^{jl} R_{ijkl},$$

где $R_{\perp\perp} = g^{ij} R_{i\perp j\perp}$ – нормальная составляющая тензора Риччи и учтено представление для обратной метрики (22.27). С учетом уравнений Гаусса–Петерсона–Кодацци (22.37) получаем

$$g^{ik} g^{jl} R_{ijkl} = \hat{R} + K^2 - K^{ij} K_{ij},$$

где $\hat{R}$ – скалярная внутренняя кривизна гиперповерхности, $K = g^{ij} K_{ij}$ – скалярная внешняя кривизна гиперповерхности. Отсюда следуют выражения для нормальных компонент тензора Эйнштейна:

$$\begin{aligned} G_{\perp\perp} &= -\frac{1}{2}(\hat{R} + K^2 - K^{ij} K_{ij}), \\ G_{\perp i} &= \hat{\nabla}_j K_i{}^j - \hat{\nabla}_i K. \end{aligned} \tag{22.39}$$

Важным обстоятельством является то, что эти компоненты тензора Эйнштейна вообще не содержат нормальных производных к гиперповерхности $\nabla_\perp$ от индуцированной метрики и тензора внешней кривизны. На гамильтоновом языке это означает отсутствие производных по времени, и что вакуумные уравнения Эйнштейна $G_{\perp\perp} = 0$ и $G_{\perp i} = 0$ представляют собой связи, поскольку компоненты тензора внешней кривизны K^{ij}, как будет показано в следующем разделе, пропорциональны импульсам, канонически сопряженным компонентам индуцированной метрики g_{ij}.

22.4 Кривизна в АДМ параметризации метрики

АДМ параметризация метрики (22.5) удобна для канонической формулировки общей теории относительности, в которой исходными независимыми переменными являются компоненты метрики $g_{\alpha\beta}$ и канонически сопряженные импульсы $p^{\alpha\beta}$. Однако, чтобы перейти от лагранжиана к гамильтониану, необходимо произвести довольно громоздкие вычисления, чем мы и займемся в настоящем разделе.

Для существенного упрощения вычислений следует воспользоваться результатами раздела 22.3, в котором описана геометрия гиперповерхностей. Будем считать, что пространство-время является топологическим произведением $\mathbb{M} \approx \mathbb{R} \times \mathbb{U}$, где первый сомножитель соответствует времени x^0. Тогда сечения $x^0 = \mathsf{const}$

пространства-времени $\mathbb{M}$ задают семейство гиперповерхностей $\mathbb{U} \subset \mathbb{M}$, которые, по-предположению, являются пространственноподобными. В качестве координат на гиперповерхностях выберем пространственные координаты

$$\{u^i\} \mapsto \{x^\mu\}.$$

При этом мы теряем возможность независимого преобразования координат в пространстве-времени $\mathbb{M}$ и пространственноподобной гиперповерхности $\mathbb{U}$, зато многие формулы упрощаются.

Матрица Якоби вложения гиперповерхности в рассматриваемом случае имеет вид

$$\{e^\alpha{}_i\} \mapsto \{0_\nu, \delta^\mu_\nu\}, \qquad \{e_\alpha{}^i\} \mapsto \{N^\mu, \delta^\mu_\nu\},$$

где 0_ν обозначает строку, состоящую из $n-1$ нулей. Вложение индуцирует на гиперповерхностях метрику $g_{\mu\nu}$ согласно формуле (22.21).

Построим векторное поле $n = n^\alpha \partial_\alpha$, ортогональное семейству пространственноподобных гиперповерхностей $x^0 = \mathsf{const}$. Из условия ортогональности $(n, X) = 0$, где $X = X^\mu \partial_\mu$ – произвольный вектор, касательный к сечению $x^0 = 0$, следует, что

$$n = n^0(\partial_0 - N^\mu \partial_\mu),$$

где n^0 – произвольная отличная от нуля функция. Если, кроме того, положить $n^0 = 1/N$, то длина нормального вектора будет равна единице, $n^2 = 1$. Таким образом, единичный вектор, нормальный к сечениям $x^0 = \mathsf{const}$ имеет вид

$$n = \frac{1}{N}(\partial_0 - N^\mu \partial_\mu) \tag{22.40}$$

и всегда является времениподобным. Соответствующая ортонормальная 1-форма имеет вид

$$n = dx^0 N. \tag{22.41}$$

Произвольный тензор на $\mathbb{M}$ можно разложить по базису $\{n, e_\mu\}$. В частности, для векторов и 1-форм справедливы разложения

$$X^\alpha = X^\perp n^\alpha + \tilde{X}^\mu e_\mu{}^\alpha, \qquad X_\alpha = X_\perp n_\alpha + \tilde{X}_\mu e^\mu{}_\alpha,$$

где

$$X^\perp = X^0 N, \qquad \tilde{X}^\mu = X^0 N^\mu + X^\mu,$$
$$X_\perp = \frac{1}{N}(X_0 - N^\mu X_\mu), \qquad \tilde{X}_\mu = X_\mu.$$

Представления для метрики (22.26) всего пространства-времени и ее обратной (22.27) принимает вид

$$\begin{aligned} g_{\alpha\beta} &= n_\alpha n_\beta + g_{\mu\nu} e_\alpha{}^\mu e_\beta{}^\nu, \\ g^{\alpha\beta} &= n^\alpha n^\beta + \hat{g}^{\mu\nu} e_\mu{}^\alpha e_\nu{}^\beta. \end{aligned} \tag{22.42}$$

Связность (22.24), индуцированная на гиперповерхностях – это символы Кристоффеля $\hat{\Gamma}_{\mu\nu}{}^\rho$, построенные по пространственной метрике $g_{\mu\nu}$.

Тензор внешней кривизны (22.30) гиперповерхности $x^0 = \mathsf{const}$ в АДМ параметризации метрики имеет вид

$$K_{\mu\nu} = \Gamma_{\mu\nu}{}^0 N = \frac{1}{2N}(\hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu - \dot{g}_{\mu\nu}), \tag{22.43}$$

где точка обозначает дифференцирование по времени, $\dot{g}_{\mu\nu} := \partial_0 g_{\mu\nu}$ и $\hat{\nabla}_\mu N_\nu = \partial_\mu N_\nu - \hat{\Gamma}_{\mu\nu}{}^\rho N_\rho$. Тензор внешней кривизны симметричен $K_{\mu\nu} = K_{\nu\mu}$, поскольку кручение в метрической формулировке общей теории относительности равно нулю. В дальнейшем нам понадобится также след тензора внешней кривизны

$$K := K_\mu{}^\mu = \hat{g}^{\mu\nu} K_{\mu\nu} = \frac{1}{2N}(2\hat{\nabla}_\mu N^\mu - \hat{g}^{\mu\nu}\dot{g}_{\mu\nu}).$$

При вычислении тензора кривизны $R_{\alpha\beta\gamma\delta}$ пространства-времени $\mathbb{M}$ все производные по времени от пространственной части метрики $\dot{g}_{\mu\nu}$ удобно выражать через $K_{\mu\nu}$. Кроме этого, для исключения вторых производных по времени $\ddot{g}_{\mu\nu}$ нам понадобится производная по времени от тензора внешней кривизны

$$\dot{K}_{\mu\nu} = \frac{1}{2N}\left[\hat{\nabla}_\mu \dot{N}_\nu + \hat{\nabla}_\nu \dot{N}_\mu - \ddot{g}_{\mu\nu} - N^\rho(\hat{\nabla}_\mu \dot{g}_{\nu\rho} + \hat{\nabla}_\nu \dot{g}_{\mu\rho} - \hat{\nabla}_\rho \dot{g}_{\mu\nu})\right] - \frac{\dot{N}}{N} K_{\mu\nu},$$

где

$$\hat{\nabla}_\mu \dot{N}_\nu = \partial_\mu \dot{N}_\nu - \hat{\Gamma}_{\mu\nu}{}^\rho \dot{N}_\rho$$

и куда при вычислениях мы должны подставить выражение для $\dot{g}_{\mu\nu}$ через $K_{\mu\nu}$.

Приступим к вычислению тензора кривизны $R_{\alpha\beta\gamma\delta}$ для метрики, записанной в виде (22.5). Первый шаг состоит в вычислении символов Кристоффеля (7.22). Прямые выкладки приводят к следующим выражениям для линейно независимых символов Кристоффеля

$$\begin{aligned}
\Gamma_{00}{}^0 &= \frac{1}{N}\left(\dot{N} + N^\rho \partial_\rho N + N^\rho N^\sigma K_{\rho\sigma}\right),\\
\Gamma_{00}{}^\mu &= \hat{g}^{\mu\nu}(\dot{N}_\nu - N\partial_\nu N - N^\rho \hat{\nabla}_\nu N_\rho) - \frac{N^\mu}{N}\left(\dot{N} + N^\rho \partial_\rho N + N^\rho N^\sigma K_{\rho\sigma}\right),\\
\Gamma_{0\mu}{}^0 &= \frac{1}{N}\left(\partial_\mu N + N^\nu K_{\mu\nu}\right),\\
\Gamma_{0\mu}{}^\nu &= \hat{\nabla}_\mu N^\nu - N K_\mu{}^\nu - \frac{N^\nu}{N}(\partial_\mu N + N^\rho K_{\mu\rho}),\\
\Gamma_{\mu\nu}{}^0 &= \frac{1}{N} K_{\mu\nu},\\
\Gamma_{\mu\nu}{}^\rho &= \hat{\Gamma}_{\mu\nu}{}^\rho - \frac{N^\rho}{N} K_{\mu\nu}.
\end{aligned} \tag{22.44}$$

В дальнейшем нам понадобятся два независимых следа символов Кристоффеля:

$$\Gamma_\alpha := \Gamma_{\alpha\beta}{}^\beta, \qquad \Xi^\alpha := g^{\beta\gamma}\Gamma_{\beta\gamma}{}^\alpha.$$

Несложные вычисления приводят к формулам

$$\begin{aligned}
\Gamma_0 &= \frac{\dot{N}}{N} + \hat{\nabla}_\mu N^\mu - NK,\\
\Gamma_\mu &= \hat{\Gamma}_\mu + \frac{\partial_\mu N}{N},\\
\Xi^0 &= \frac{1}{N}K + \frac{1}{N^3}(\dot{N} - N^\mu \partial_\mu N),\\
\Xi^\mu &= \left(\hat{g}^{\rho\sigma} + \frac{N^\rho N^\sigma}{N^2}\right)\hat{\Gamma}_{\rho\sigma}{}^\mu - \frac{N^\mu}{N}K - \frac{N^\mu}{N^3}(\dot{N} - N^\rho \partial_\rho N) +\\
&\quad + \frac{1}{N^2}\hat{g}^{\mu\rho}(\dot{N}_\rho - N\partial_\rho N - N^\sigma \hat{\nabla}_\rho N_\sigma - 2N^\sigma \hat{\nabla}_\sigma N_\rho + 2NN^\sigma K_{\rho\sigma}).
\end{aligned} \tag{22.45}$$

Выпишем также формулы дифференцирования по времени для символов Кристоффеля:

$$\begin{aligned}
\partial_0\hat\Gamma_{\mu\nu\rho} &= \frac{1}{2}(\hat\nabla_\mu\dot g_{\nu\rho} + \hat\nabla_\nu\dot g_{\mu\rho} - \hat\nabla_\rho\dot g_{\mu\nu}) + \hat\Gamma_{\mu\nu}{}^\sigma\dot g_{\rho\sigma},\\
\partial_0\hat\Gamma_{\mu\nu}{}^\sigma &= \frac{1}{2}\hat g^{\sigma\rho}(\hat\nabla_\mu\dot g_{\nu\rho} + \hat\nabla_\nu\dot g_{\mu\rho} - \hat\nabla_\rho\dot g_{\mu\nu}),\\
\partial_0\hat\Gamma_\mu &= \frac{1}{2}\hat g^{\nu\rho}\hat\nabla_\mu\dot g_{\nu\rho}.
\end{aligned} \tag{22.46}$$

В этих выражениях производные по времени $\dot g_{\mu\nu}$ также исключаются с помощью соотношения (22.43).

Теперь вычислим линейно независимые компоненты тензора кривизны:

$$\begin{aligned}
R_{0\mu0\nu} =& -N\dot K_{\mu\nu} + \hat R_{\mu\rho\nu\sigma}N^\rho N^\sigma + NN^\rho(\hat\nabla_\mu K_{\nu\rho} + \hat\nabla_\nu K_{\mu\rho} - \hat\nabla_\rho K_{\mu\nu}) +\\
&+ N\hat\nabla_\mu\hat\nabla_\nu N + K_{\mu\nu}N^\rho N^\sigma K_{\rho\sigma} + N(K_\mu{}^\rho\hat\nabla_\nu N_\rho + K_\nu{}^\rho\hat\nabla_\mu N_\rho) -\\
&- N^2K_\mu{}^\rho K_{\nu\rho} - N^\rho N^\sigma K_{\mu\rho}K_{\nu\sigma},\\
R_{\mu\nu\rho0} =& \hat R_{\mu\nu\rho\sigma}N^\sigma + N(\hat\nabla_\mu K_{\nu\rho} - \hat\nabla_\nu K_{\mu\rho}) + (K_{\mu\rho}K_{\nu\sigma} - K_{\nu\rho}K_{\mu\sigma})N^\sigma,\\
R_{\mu\nu\rho\sigma} =& \hat R_{\mu\nu\rho\sigma} + K_{\mu\rho}K_{\nu\sigma} - K_{\mu\sigma}K_{\nu\rho},
\end{aligned} \tag{22.47}$$

где мы воспользовались формулой для коммутатора ковариантных производных

$$(\hat\nabla_\mu\hat\nabla_\nu - \hat\nabla_\nu\hat\nabla_\mu)N_\rho = -\hat R_{\mu\nu\rho\sigma}N^\sigma.$$

Компоненты тензора кривизны, имеющие по крайней мере один временно́й индекс, относительно ортонормального базиса n, e_μ выглядят проще:

$$\begin{aligned}
R_{\perp\mu\perp\nu} &= \frac{1}{N}\left(-\dot K_{\mu\nu} + \hat\nabla_\mu\hat\nabla_\nu N + K_{\mu\rho}\hat\nabla_\nu N^\rho + K_{\nu\rho}\hat\nabla_\mu N^\rho - NK_{\mu\rho}K_\nu{}^\rho + N^\rho\hat\nabla_\rho K_{\mu\nu}\right),\\
R_{\mu\nu\rho\perp} &= \hat\nabla_\mu K_{\nu\rho} - \hat\nabla_\nu K_{\mu\rho}.
\end{aligned} \tag{22.48}$$

Фактически, компоненты тензора кривизны $R_{\mu\nu\rho\sigma}$ и $R_{\mu\nu\rho\perp}$ уже были получены в предыдущем разделе, см. формулы (22.37), (22.38), без прямых вычислений.

Тензор Риччи имеет следующие линейно независимые компоненты:

$$\begin{aligned}
R_{00} =& -N\hat g^{\mu\nu}\dot K_{\mu\nu} + \hat R_{\mu\nu}N^\mu N^\nu + NN^\mu(2\hat\nabla_\nu K^\nu{}_\mu - \partial_\mu K) + N\hat\nabla^\mu\hat\nabla_\mu N +\\
&+ N^\mu N^\nu K_{\mu\nu}K + 2NK^{\mu\nu}\hat\nabla_\mu N_\nu - N^2K^{\mu\nu}K_{\mu\nu} - 2N^\mu N^\nu K_\mu{}^\rho K_{\nu\rho} +\\
&+ \frac{N^\mu N^\nu}{N}\left(-\dot K_{\mu\nu} + N^\rho\hat\nabla_\mu K_{\nu\rho} + \hat\nabla_\mu\hat\nabla_\nu N + 2K_\mu{}^\rho\hat\nabla_\nu N_\rho\right),\\
R_{0\mu} =& \frac{N^\nu}{N}\left(-\dot K_{\mu\nu} + N^\sigma\hat\nabla_\nu K_{\mu\sigma} + \hat\nabla_\mu\hat\nabla_\nu N + K_\mu{}^\rho\hat\nabla_\nu N_\rho + K_\nu{}^\rho\hat\nabla_\mu N_\rho\right)\\
&+ \hat R_{\mu\nu}N^\nu + N\left(\hat\nabla_\nu K^\nu{}_\mu - \partial_\mu K\right) + K_{\mu\nu}N^\nu K - 2K_\mu{}^\rho K_{\nu\rho}N^\nu,\\
R_{\mu\nu} =& \hat R_{\mu\nu} + \frac{1}{N}\left(-\dot K_{\mu\nu} + \hat\nabla_\mu\hat\nabla_\nu N + K_\mu{}^\rho\hat\nabla_\nu N_\rho + K_\nu{}^\rho\hat\nabla_\mu N_\rho + N^\rho\hat\nabla_\rho K_{\mu\nu}\right) +\\
&+ K_{\mu\nu}K - 2K_\mu{}^\rho K_{\nu\rho}.
\end{aligned} \tag{22.49}$$

Опять же, компоненты тензора Риччи относительно базиса n, e_μ существенно проще:

$$\begin{aligned}
R_{\perp\perp} =& -\frac{1}{N}\hat g^{\mu\nu}\dot K_{\mu\nu} + \frac{1}{N}\hat\nabla^\mu\hat\nabla_\mu N + \frac{2}{N}K^{\mu\nu}\hat\nabla_\mu N_\nu - K^{\mu\nu}K_{\mu\nu} + \frac{N^\mu}{N}\partial_\mu K,\\
R_{\perp\mu} =& \hat\nabla_\nu K^\nu{}_\mu - \partial_\mu K.
\end{aligned} \tag{22.50}$$

Далее, вычисляем скалярную кривизну

$$R = \hat{R} + \frac{2}{N}\left(-\hat{g}^{\mu\nu}\dot{K}_{\mu\nu} + \hat{\nabla}^{\mu}\hat{\nabla}_{\mu}N + 2K^{\mu\nu}\hat{\nabla}_{\mu}N_{\nu} + N^{\mu}\partial_{\mu}K\right) - 3K^{\mu\nu}K_{\mu\nu} + K^2. \quad (22.51)$$

Вакуумные уравнения Эйнштейна

$$G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = 0$$

также можно переписать в базисе n, e^{μ}. Выражения для $G_{\perp\perp}$ и $G_{0\perp}$ уже были получены в предыдущем разделе менее трудоемким образом (22.39). Выражение для $G_{\mu\nu}$ в АДМ параметризации метрики следует из выражения для тензора Риччи (22.49) и скалярной кривизны (22.51). Оно громоздко и мы не будем его выписывать.

Как уже отмечалось, тензор Эйнштейна удовлетворяет свернутым тождествам Бианки (7.108). Эти тождества полезно переписать в АДМ параметризации метрики. Четыре тождества (7.108) перепишем в следующем виде:

$$\nabla_{\beta}G^{\beta\alpha} = 0 \quad \Leftrightarrow \quad (\nabla_{\beta}G^{\beta\alpha})n_{\alpha} = 0, \quad (\nabla_{\beta}G^{\beta\alpha})e_{\alpha}{}^{\mu} = 0.$$

После довольно утомительной алгебры получим равенства:

$$(\nabla_{\beta}G^{\beta\alpha})n_{\alpha} = \frac{1}{N}\dot{G}^{\perp\perp} - \frac{N^{\mu}}{N}\partial_{\mu}G^{\perp\perp} - G^{\perp\perp}K + \hat{\nabla}_{\mu}G^{\mu\perp} + 2\frac{\partial_{\mu}N}{N}G^{\mu\perp} + G^{\mu\nu}K_{\mu\nu}, \quad (22.52)$$

$$(\nabla_{\beta}G^{\beta\alpha})e_{\alpha}{}^{\mu} = \frac{1}{N}\dot{G}^{\perp\mu} - \frac{N^{\nu}}{N}\hat{\nabla}_{\nu}G^{\perp\mu} - G^{\perp\mu}K + \hat{\nabla}_{\nu}G^{\nu\mu} + \frac{\partial_{\nu}N}{N}G^{\nu\mu} - \\ - \frac{1}{N}G^{\perp\perp}\hat{g}^{\mu\nu}\partial_{\nu}N + \frac{1}{N}G^{\perp\nu}\hat{\nabla}_{\nu}N^{\mu} - 2G^{\perp\nu}K_{\nu}{}^{\mu}. \quad (22.53)$$

Эти тождества будут полезны при анализе вторичных связей в общей теории относительности.

22.5 Гамильтониан

Скалярная кривизна содержит вторые производные от компонент метрики как по времени, так и по пространственным координатам и поэтому неудобна для канонической формулировки общей теории относительности. Чтобы исключить из лагранжиана все вторые производные, в разделе 22.1 к действию был добавлен граничный член. Однако для канонической формулировки достаточно исключить из лагранжиана только вторые производные по времени. Наиболее простой вид лагранжиан принимает после добавления следующего граничного члена

$$L_{\text{ADM}} = N\hat{e}R + 2\partial_0\left(\hat{e}K\right) - 2\partial_{\mu}\left(\hat{e}\hat{g}^{\mu\nu}\partial_{\nu}N\right), \quad (22.54)$$

где

$$\hat{e} := \det \hat{e}_{\mu}{}^{i} = \sqrt{|\hat{g}|} = \sqrt{|\det g_{\mu\nu}|}.$$

Прямые вычисления приводят к следующему простому выражению

$$L_{\text{ADM}} = N\hat{e}\left(K^{\mu\nu}K_{\mu\nu} - K^2 + \hat{R}\right). \quad (22.55)$$

Теперь нетрудно перейти к гамильтонову формализму. Во-первых, АДМ-лагранжиан не содержит производных по времени от функции хода N и функций сдвига N_{μ}. Это значит, что теория содержит n первичных связей:

$$p^0 = \frac{\partial L_{\text{ADM}}}{\partial \dot{N}} = 0, \qquad p^{\mu} = \frac{\partial L_{\text{ADM}}}{\partial \dot{N}_{\mu}} = 0, \quad (22.56)$$

число которых совпадает с числом независимых функций, параметризующих диффеоморфизмы. Импульсы, канонически сопряженные к метрике $g_{\mu\nu}$, имеют вид

$$p^{\mu\nu} = \frac{\partial L_{\text{ADM}}}{\partial \dot{g}_{\mu\nu}} = -\frac{1}{2N}\frac{\partial L_{\text{ADM}}}{\partial K_{\mu\nu}} = -\hat{e}\,(K^{\mu\nu} - g^{\mu\nu}K)\,, \tag{22.57}$$

где мы воспользовались выражением для внешней кривизны (22.43). Отсюда следует, что обобщенные импульсы пропорциональны тензору внешней кривизны. Отметим, что импульсы являются не тензорами относительно преобразований координат x^μ, а тензорными плотностями степени -1, как и определитель репера, степень которого, по-определению, равна

$$\mathsf{deg}\,\hat{e} = -1.$$

Поэтому ковариантная производная от импульсов равна

$$\hat{\nabla}_\mu p^{\nu\rho} = \partial_\mu p^{\nu\rho} + \hat{\Gamma}_{\mu\sigma}{}^{\nu} p^{\sigma\rho} + \hat{\Gamma}_{\mu\sigma}{}^{\rho} p^{\nu\sigma} - \hat{\Gamma}_{\mu} p^{\nu\rho}.$$

Чтобы исключить из АДМ-лагранжиана скорости $\dot{g}_{\mu\nu}$, разложим импульсы на неприводимые компоненты, выделив из $p^{\alpha\beta}$ след,

$$p^{\mu\nu} = \tilde{p}^{\mu\nu} + \frac{1}{n-1} p \hat{g}^{\mu\nu}, \tag{22.58}$$

где мы ввели след импульсов

$$p := p^{\mu\nu} g_{\mu\nu} = \hat{e}(n-2)K \tag{22.59}$$

и симметричную бесследовую часть

$$\tilde{p}^{\mu\nu} := \tilde{p}^{\nu\mu} = -\hat{e}\left(K^{\mu\nu} - \frac{1}{n-1}\hat{g}^{\mu\nu}K\right), \qquad \tilde{p}^{\mu\nu} g_{\mu\nu} = 0.$$

Импульсы $p^{\mu\nu}$ и внешняя кривизна $K^{\mu\nu}$ находятся во взаимно однозначном соответствии. Из формулы (22.57) следует выражение внешней кривизны через импульсы

$$K^{\mu\nu} = -\frac{1}{\hat{e}}\left(p^{\mu\nu} - \frac{1}{n-2} g^{\mu\nu} p\right). \tag{22.60}$$

Теперь можно решить уравнение (22.57) относительно скоростей, воспользовавшись соотношением (22.43),

$$\dot{g}_{\mu\nu} = \frac{2N}{\hat{e}}\left(p_{\mu\nu} - \frac{1}{(n-2)} p g_{\mu\nu}\right) + \hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu.$$

Несложные вычисления приводят к гамильтоновой плотности

$$H = p^{\mu\nu}\dot{g}_{\mu\nu} - L_{\text{ADM}} = N H_\perp + N^\mu H_\mu + 2\partial_\mu(p^{\mu\nu} N_\nu), \tag{22.61}$$

где

$$\begin{aligned} H_\perp &= \frac{1}{\hat{e}}\left(p^{\mu\nu}p_{\mu\nu} - \frac{1}{(n-2)}p^2\right) - \hat{e}\hat{R}, \\ H_\mu &= -2\hat{\nabla}_\nu p^{\nu}{}_{\mu} = -2\partial_\nu p^{\nu}{}_{\mu} + \partial_\mu g_{\nu\rho} p^{\nu\rho}, \end{aligned} \tag{22.62}$$

и $p_{\mu\nu} = g_{\mu\rho} g_{\nu\sigma} p^{\rho\sigma}$.

Замечание. Ковариантная производная от импульсов в последнем выражении содержит только одно слагаемое с символами Кристоффеля, так как импульсы являются тензорными плотностями. □

Отбрасывая в выражении для гамильтониана (22.61) дивергенцию, приходим к окончательному выражению для гамильтониана

$$\mathcal{H}_{\text{ADM}} = \int d\boldsymbol{x} H_{\text{ADM}} = \int d\boldsymbol{x}(NH_\perp + N^\mu H_\mu). \tag{22.63}$$

Для функционалов, полученных интегрированием по пространственным сечениям $x^0 = \mathsf{const}$, будем использовать каллиграфический шрифт.

Перепишем выражение для $H_\perp$ в терминах неприводимых компонент импульсов (22.58):

$$H_\perp = \frac{1}{\hat{e}}\left[\tilde{p}^{\mu\nu}\tilde{p}_{\mu\nu} - \frac{1}{(n-1)(n-2)}p^2\right] - \hat{e}\hat{R}.$$

Отсюда следует, что квадратичная форма импульсов в $H_\perp$ при $n \geq 3$ не является положительно определенной. Отсутствие положительной определенности квадратичной формы импульсов в гамильтониане представляет серьезные трудности для физической интерпретации моделей. В рассматриваемом случае не все компоненты импульсов являются физическими, т.к. общая теория относительности инвариантна относительно общих преобразований координат. После исключения нефизических степеней свободы с помощью связей и калибровочных условий квадратичная форма импульсов для физических степеней свободы будет положительно определена.

22.6 Вторичные связи

Для завершения построения гамильтонова формализма необходимо исследовать согласованность первичных связей (22.56) с уравнениями движения. Фазовое пространство общей теории относительности бесконечномерно и описывается $n(n+1)$ сопряженными координатами и импульсами: (N, p^0), (N_μ, p^μ), $(g_{\mu\nu}, p^{\mu\nu})$, в каждой точке пространства $\boldsymbol{x} \in \mathbb{U}$. На этом фазовом пространстве задана каноническая пуассонова структура с помощью одновременны́х скобок Пуассона:

$$[N, p'^0] = \delta, \qquad [N_\mu, p'^\nu] = \delta^\nu_\mu\delta, \qquad [g_{\mu\nu}, p'^{\rho\sigma}] = \delta^{\rho\sigma}_{\mu\nu}\delta, \tag{22.64}$$

где штрих у полевой переменной означает, что она рассматривается в точке $\boldsymbol{x}' = (x'^1, \ldots, x'^{n-1})$. Все поля рассматриваются в момент времени $t = x^0$. В правых частях скобок Пуассона для краткости использованы следующие обозначения для $(n-1)$-мерной δ-функции и симметризованной комбинации символов Кронекера:

$$\begin{aligned} &\delta := \delta^{(n-1)}(\boldsymbol{x} - \boldsymbol{x}') := \delta(x^1 - x'^1)\ldots\delta(x^{n-1} - x'^{n-1}),\\ &\delta^{\rho\sigma}_{\mu\nu} := \frac{1}{2}(\delta^\rho_\mu\delta^\sigma_\nu + \delta^\sigma_\mu\delta^\rho_\nu). \end{aligned} \tag{22.65}$$

Чтобы отличать δ-функцию от вариации поля, мы будем писать все δ-функции справа.

Если заданы два функционала

$$\mathcal{F} = \int d\boldsymbol{x}\, F, \qquad \mathcal{G} = \int d\boldsymbol{x}\, G,$$

где F и G – некоторые функции от пространственной метрики $g_{\mu\nu}$, сопряженных импульсов $p^{\mu\nu}$ и их производных, то скобка Пуассона этих функционалов определяется следующим интегралом

$$[\mathcal{F},\mathcal{G}]=\int d\boldsymbol{x}\left(\frac{\delta\mathcal{F}}{\delta g_{\mu\nu}}\frac{\delta G}{\delta p^{\mu\nu}}-\frac{\delta\mathcal{F}}{\delta p^{\mu\nu}}\frac{\delta G}{\delta g_{\mu\nu}}\right). \tag{22.66}$$

Напомним, что сами компоненты метрики и импульсов также можно рассматривать в виде функционалов. Например,

$$g_{\mu\nu}(\boldsymbol{x})=\int d\boldsymbol{x}' g_{\mu\nu}(\boldsymbol{x}')\delta(\boldsymbol{x}-\boldsymbol{x}').$$

Тогда последняя скобка Пуассона в (22.64) является следствием общего определения канонической пуассоновой структуры на множестве функционалов (22.66).

Аналогично определяется скобка Пуассона для функционалов, зависящих от полного набора канонических переменных: (N,p^0), (N_μ,p^μ), $(g_{\mu\nu},p^{\mu\nu})$.

Рассмотрим теперь гамильтоновы уравнения движения для первичных связей (22.56):

$$\dot{p}^0=[p^0,\mathcal{H}_{\text{ADM}}]=-\frac{\delta\mathcal{H}_{\text{ADM}}}{\delta N}=-H_\perp,$$
$$\dot{p}^\mu=[p^\mu,\mathcal{H}_{\text{ADM}}]=-\frac{\delta\mathcal{H}_{\text{ADM}}}{\delta N^\mu}=-H^\mu.$$

Из условия согласованности первичных связей с уравнениями движения $\dot{p}^0=0$, $\dot{p}^\mu=0$ следуют вторичные связи:

$$H_\perp=0,\qquad H_\mu=0. \tag{22.67}$$

Отметим, что вторичные связи являются не тензорами, а тензорными плотностями степени -1. Кроме того, вместо связей H^μ удобнее рассматривать эквивалентную систему связей с опущенным индексом $H_\mu:=g_{\mu\nu}H^\nu$. В следующем разделе будет показано, что эти связи определяют генераторы преобразований координат на сечениях $x^0=\mathsf{const}$ и удовлетворяют более простой алгебре.

Связи H_μ линейны по импульсам и метрике. Связь $H_\perp$ квадратична по импульсам и неполиномиальна по метрике $g_{\mu\nu}$, поскольку зависит от корня из определителя метрики $\hat{e}$ и обратной метрики $g^{\mu\nu}$. Последнее обстоятельство является существенной технической трудностью при построении теории возмущений.

Вторичные связи не зависят от канонических переменных (N,p^0), (N_μ,p^μ), и их можно исключить, рассматривая $n(n-1)\times\infty^{n-1}$-мерное фазовое пространство переменных $g_{\mu\nu}$ и $p^{\mu\nu}$, на которые наложены связи (22.67). В этом случае функция хода N и функции сдвига N_μ рассматриваются, как множители Лагранжа в задаче на условный экстремум для действия

$$S=\int d^nx(p^{\mu\nu}\dot{g}_{\mu\nu}-H_{\text{ADM}}).$$

Поскольку гамильтониан общей теории относительности (22.63) представляет собой линейную комбинацию связей, то для исследования согласованности вторичных связей (22.67) с уравнениями движения необходимо вычислить скобки Пуассона связей между собой. Алгебра связей в общей теории относительности хорошо известна:

$$[H_\perp,H'_\perp]=-(H_\mu\hat{g}^{\mu\nu}+H'_\mu\hat{g}'^{\mu\nu})\delta_\nu, \tag{22.68}$$
$$[H_\perp,H'_\mu]=-H'_\perp\delta_\mu, \tag{22.69}$$
$$[H_\mu,H'_\nu]=-H_\nu\delta_\mu-H'_\mu\delta_\nu, \tag{22.70}$$

где введено сокращенное обозначение для производной δ-функции:

$$\delta_\mu := \frac{\partial}{\partial x'^\mu}\delta(x'-x).$$

Прямое вычисление алгебры связей (22.68)–(22.70) требует большого терпения и внимания. Эта алгебра впервые была выписана Дираком [?] из соображений симметрии. При этом были сделаны предположения о виде связей, которые не выполняются в общей теории относительности, и вопрос о существовании соответствующего канонического преобразования в настоящее время открыт. Поэтому получение алгебры связей Дираком нельзя считать удовлетворительным.

Вид двух скобок Пуассона (22.69) и (22.70) действительно можно найти, не прибегая к прямым вычислениям. С этой целью рассмотрим функционал

$$T_u = -\int d\boldsymbol{x}\ u^\mu H_\mu,$$

где u^μ малое векторное поле. Вычисление скобок Пуассона координат фазового пространства $g_{\mu\nu}$, $p^{\mu\nu}$ с T_u приводит к равенствам:

$$\delta_u g_{\mu\nu} = [g_{\mu\nu}, T_u] = -\partial_\mu u^\rho g_{\rho\nu} - \partial_\nu u^\rho g_{\mu\rho} - u^\rho \partial_\rho g_{\mu\nu},$$
$$\delta_u p^{\mu\nu} = [p^{\mu\nu}, T_u] = \partial_\rho u^\mu p^{\rho\nu} + \partial_\rho u^\nu p^{\mu\rho} - \partial_\rho(u^\rho p^{\mu\nu}).$$

Это значит, что функционал T_u, который определяется связями H_μ, является генератором общих преобразований координат на гиперповерхностях $x^0 = \mathsf{const}$. Напомним, что импульсы $p^{\mu\nu}$ являются тензорными плотностями степени -1. Алгебра преобразований координат хорошо известна и задается скобкой Пуассона (22.70). Скобку Пуассона (22.69) также можно не вычислять явно. Ее вид следует из того, что связь $H_\perp$ является скалярной плотностью степени -1. Таким образом, необходимо вычислить только скобку Пуассона (22.68). Эти вычисления очень громоздки, и впервые, по-видимому, были проделаны намного позже ДеВиттом [?]. В следующем разделе мы вычислим эту скобку после канонического преобразования, приводящего связи к полиномиальному виду, что существенно упрощает вычисления.

Ввиду того, что связи H_μ задают только пространственные диффеоморфизмы, мы будем называть их кинематическими. Они также не зависят от констант связи в действии, если таковые имеются. Связь $H_\perp$ называется динамической, так как она определяет развитие начальных данных во времени и существенно зависит от исходного действия, в частности, от констант связи.

Для сравнения приведем скобку Пуассона связей $H^\mu = 0$ с контравариантным индексом, которые эквивалентны связям $H_\mu = 0$,

$$[H^\mu, H'^\nu] = (\hat{g}^{\mu\nu}H^\rho + \hat{g}'^{\mu\nu}H'^\rho)\,\delta_\rho + (\hat{g}^{\mu\rho}\partial_\rho g_{\sigma\lambda}\hat{g}^{\nu\sigma} - \hat{g}^{\nu\rho}\partial_\rho g_{\sigma\lambda}\hat{g}^{\mu\sigma})\,H^\lambda\delta.$$

Как видим, данная скобка Пуассона выглядит сложнее (22.70). Это говорит о том, что выбор вида связей из эквивалентных наборов является очень важным и может привести к существенному упрощению вычислений. К сожалению, общего рецепта как поступать в таких случаях не существует, и действует метод “пристального всматривания”.

Теперь посмотрим на вторичные связи с другой точки зрения. Сравнение выражения для связей (22.62) с тензором Эйнштейна (22.39) приводит к равенствам

$$H_\perp = 2\hat{e}G_{\perp\perp} = \frac{2\hat{e}}{N^2}(G_{00} - 2N^\mu G_{0\mu} + N^\mu N^\nu G_{\mu\nu}),$$
$$H_\mu = 2\hat{e}G_{\perp\mu} = \frac{2\hat{e}}{N}(G_{0\mu} - N^\nu G_{\nu\mu}).$$

Мы видим, что вторичные связи (22.67) эквивалентны четырем уравнениям Эйнштейна

$$G_{\perp\perp} = 0, \quad G_{\perp\mu} = 0,$$

которые представляют собой определенные комбинации всех вакуумных уравнений Эйнштейна $G_{\alpha\beta} = 0$. Если выполнены все уравнения Эйнштейна, а также вторичные связи в начальный момент времени, то вторичные связи будут выполнены также во все последующие моменты времени. Это следует из того, что $\dot{G}_{\perp\perp} = 0$ и $\dot{G}_{\perp\mu} = 0$ как следствие свернутых тождеств Бианки (22.52), (22.53). Таким образом мы доказали, что вторичные связи в общей теории относительности согласованы с уравнениями движения, т.е. представляют собой связи первого рода. Для этого необязательно вычислять в явном виде алгебру связей (22.68)–(22.70). Однако знание алгебры связей полезно.

В заключение данного раздела выпишем канонические уравнения движения для пространственных компонент метрики и импульсов

$$\dot{g}_{\mu\nu} = \frac{2N}{\hat{e}} p_{\mu\nu} - \frac{2N}{\hat{e}(n-2)} g_{\mu\nu} p + \hat{\nabla}_\mu N_\nu + \hat{\nabla}_\nu N_\mu, \tag{22.71}$$

$$\begin{aligned}\dot{p}^{\mu\nu} = {} & \frac{N}{2\hat{e}} \hat{g}^{\mu\nu} \left(p^{\rho\sigma} p_{\rho\sigma} - \frac{1}{n-2} p^2 \right) - \frac{2N}{\hat{e}} \left(p^{\mu\rho} p^{\nu}{}_{\rho} - \frac{1}{n-2} p^{\mu\nu} p \right) - \\ & - \hat{e} N \left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \right) + \hat{e} (\triangle N g^{\mu\nu} - \hat{\nabla}^\mu \hat{\nabla}^\nu N) - \\ & - p^{\mu\rho} \hat{\nabla}_\rho N^\nu - p^{\nu\rho} \hat{\nabla}_\rho N^\mu + \hat{\nabla}_\rho (N^\rho p^{\mu\nu}),\end{aligned} \tag{22.72}$$

где $\triangle := \hat{\nabla}^\mu \hat{\nabla}_\mu$ – оператор Лапласа–Бельтрами.

Уравнение (22.71), конечно, совпадает с определением импульсов (22.57). Уравнения для импульсов (22.72) эквивалентны уравнениям Эйнштейна $G_{\mu\nu} = 0$. Чтобы доказать это, запишем уравнения (22.72) в виде

$$\dot{p}^{\mu\nu} - (\dots)^{\mu\nu} = 0,$$

где точки обозначают все слагаемые в правой части (22.72). Тогда можно проверить, что выполнено равенство

$$g_{\mu\rho} g_{\nu\sigma} [\dot{p}^{\rho\sigma} - (\dots)^{\rho\sigma}] = N \hat{e} G_{\mu\nu}.$$

Таким образом, канонические уравнения движения (22.71), (22.72) с учетом уравнений связей (22.67) эквивалентны полной системе вакуумных уравнений Эйнштейна $G_{\alpha\beta} = 0$.

22.7 Полиномиальная гамильтонова форма

Гамильтониан (22.63), построенный ранее, полиномиален по импульсам $p^{\mu\nu}$, но неполиномиален по компонентам метрики $g_{\mu\nu}$, т.к. содержит скалярную кривизну $\hat{R}$ и определитель репера, которые входят в динамическую связь $H_\perp$. В настоящем разделе мы опишем каноническое преобразование, которое приводит к полиномиальной форме гамильтониана и, следовательно, к полиномиальной форме уравнений движения. Это преобразование является аналогом полиномиальной лагранжевой формы действия Гильберта–Эйнштейна, рассмотренной в разделе 21.7.

Идея канонического преобразования состоит в следующем. Импульсы $p^{\mu\nu}$ приводимы и разлагаются на бесследовую часть $\tilde{p}^{\mu\nu}$ и след p (22.58). При вычислениях, как

правило, удобнее работать с неприводимыми компонентами, поскольку много слагаемых автоматически сокращаются. Поставим вопрос: “Нельзя ли совершить такое каноническое преобразование, после которого новыми импульсами будут неприводимые компоненты $\tilde{p}^{\mu\nu}$ и p ?”. Этот вопрос нетривиален, потому что разложение импульсов включает метрику, компоненты которой сами являются координатами фазового пространства. Ответ на поставленный вопрос отрицательный, потому что скобка Пуассона импульсов между собой отлична от нуля. Например, $[\tilde{p}^{\mu\nu}, p'] \neq 0$, где $p' := p(t, \boldsymbol{x}')$. Однако существует такое каноническое преобразование, что новые импульсы будут пропорциональны неприводимым компонентам $\tilde{p}^{\mu\nu}$ и p. Построением этого канонического преобразования мы и займемся в настоящем разделе.

Рассмотрим каноническое преобразование

$$(g_{\mu\nu}, p^{\rho\sigma}) \quad \mapsto \quad (k_{\mu\nu}, P^{\rho\sigma}), (\varrho, P), \tag{22.73}$$

к новым парам канонически сопряженных координат $k_{\mu\nu}$, ϱ и импульсов $P^{\mu\nu}$, P, где на координаты $k_{\mu\nu}$ и сопряженные им импульсы $P^{\mu\nu}$ наложены дополнительные условия

$$|\det k_{\mu\nu}| = 1, \qquad P^{\mu\nu} k_{\mu\nu} = 0. \tag{22.74}$$

В качестве производящего функционала канонического преобразования выберем функционал

$$\mathcal{F} = -\int d\boldsymbol{x}\, \varrho^s k_{\mu\nu} p^{\mu\nu}, \qquad s \in \mathbb{R}, \quad s \neq 0, \tag{22.75}$$

зависящий от новых координат $\varrho, k_{\mu\nu}$ и старых импульсов $p^{\mu\nu}$, а также от вещественного параметра s, который будет определен позже. Тогда старые координаты и новые импульсы определяются вариационными производными (см. раздел 20.1.11)

$$g_{\mu\nu} := -\frac{\delta \mathcal{F}}{\delta p^{\mu\nu}} = \varrho^s k_{\mu\nu}, \tag{22.76}$$

$$P^{\mu\nu} := -\frac{\delta \mathcal{F}}{\delta k_{\mu\nu}} = \varrho^s \tilde{p}^{\mu\nu}, \tag{22.77}$$

$$P := -\frac{\delta \mathcal{F}}{\delta \varrho} = \frac{s}{\varrho} p. \tag{22.78}$$

При вычислении вариационной производной по $k_{\mu\nu}$ учтено условие $|\det k_{\mu\nu}| = 1$, из которого вытекает ограничение на вариации $k^{\mu\nu}\delta k_{\mu\nu} = 0$, где $k^{\mu\nu}$ – тензорная плотность, обратная к $k_{\mu\nu}$: $k^{\mu\nu}k_{\nu\sigma} = \delta^\mu_\sigma$. Тем самым равенство нулю следа импульсов (22.74) автоматически следует из условия единичности определителя плотности $k_{\mu\nu}$ для производящего функционала (22.75). В выражении (22.78) учтено соотношение (22.76).

По сути дела, в качестве новой канонической переменной из метрики выделен ее определитель в некоторой степени, как следует из (22.76)

$$|\det g_{\mu\nu}| = \hat{e}^2 = \varrho^{s(n-1)}.$$

Поскольку метрика $g_{\mu\nu}$ невырождена, то отсюда, в частности, следует равенство $\varrho > 0$.

В дальнейшем симметричную тензорную плотность с единичным определителем $k_{\mu\nu}$ мы для краткости также будем называть метрикой.

Из формул (22.77), (22.78) следует, что новые импульсы $P^{\mu\nu}$ и P пропорциональны неприводимым компонентам старых импульсов $\tilde{p}^{\mu\nu}$ и p.

Отметим, что все новые канонические переменные являются тензорными плотностями следующих степеней:

$$\deg k_{\mu\nu} = \frac{2}{n-1}, \qquad \deg \varrho = -\frac{2}{s(n-1)},$$
$$\deg P^{\mu\nu} = -\frac{2}{n-1} - 1, \qquad \deg P = \frac{2}{s(n-1)} - 1.$$

Чтобы вычислить скалярную кривизну $\hat{R}$, входящую в связь (22.62), заметим, что каноническое преобразование (22.76) имеет вид преобразования Вейля $g_{\mu\nu} = \mathsf{e}^{2\phi} k_{\mu\nu}$, где $\phi = s \ln \varrho/2$. Поэтому можно воспользоваться полученной ранее формулой (**??**)

Прямые вычисления приводят к следующему выражению для скалярной кривизны сечения $x^0 = \mathsf{const}$ в новых координатах:

$$\hat{R} = \varrho^{-s-2}\left[\varrho^2 R^{(k)} + s(n-2)\varrho\partial_\mu(k^{\mu\nu}\partial_\nu\varrho) + s(n-2)\left(s\frac{n-3}{4} - 1\right)(\partial\varrho)^2\right], \qquad (22.79)$$

где $(\partial\varrho)^2 := k^{\mu\nu}\partial_\mu\varrho\partial_\nu\varrho$. Скалярная кривизна, построенная по метрике $k_{\mu\nu}$, принимает удивительно простой вид

$$R^{(k)} = \partial^2_{\mu\nu}k^{\mu\nu} + \frac{1}{2}k^{\mu\nu}\partial_\rho k_{\mu\sigma}\partial_\nu k^{\rho\sigma} - \frac{1}{4}k^{\mu\nu}\partial_\mu k_{\rho\sigma}\partial_\nu k^{\rho\sigma}. \qquad (22.80)$$

Из единичности определителя метрики следует, что компоненты обратной метрики $k^{\mu\nu}$ являются полиномами степени $n-2$ по компонентам $k_{\mu\nu}$:

$$k^{\mu\nu} = \frac{1}{(n-2)!}\hat{\varepsilon}^{\mu\rho_1\ldots\rho_{n-2}}\hat{\varepsilon}^{\nu\sigma_1\ldots\sigma_{n-2}}k_{\rho_1\sigma_1}\ldots k_{\rho_{n-2}\sigma_{n-2}},$$

где $\hat{\varepsilon}^{\mu_1\ldots\mu_{n-1}}$ – полностью антисимметричная тензорная плотность ранга $n-1$. Поэтому скалярная кривизна $R^{(k)}$ полиномиальна как по метрике $k_{\mu\nu}$, так и по обратной метрике $k^{\mu\nu}$.

Динамическая связь в новых переменных принимает вид

$$H_\perp = \varrho^{-\frac{s(n-1)}{2}}\left[P^{\mu\nu}P_{\mu\nu} - \frac{\varrho^2}{s^2(n-1)(n-2)}P^2\right] -$$
$$- \varrho^{\frac{s(n-1)}{2}-s-2}\left[\varrho^2 R^{(k)} + s(n-2)\varrho\partial_\mu(k^{\mu\nu}\partial_\nu\varrho) + s(n-2)\left(s\frac{n-3}{4} - 1\right)(\partial\varrho)^2\right],$$

где опускание индексов у импульсов производится с помощью новой метрики, $P_{\mu\nu} := k_{\mu\rho}k_{\nu\sigma}P^{\rho\sigma}$.

Проанализируем возможность такого выбора постоянной s, чтобы динамическая связь имела полиномиальный вид. Оба выражения в квадратных скобках полиномиальны по всем динамическим переменным. Поскольку $n \geq 3$, то для положительности степени плотности ϱ перед первой квадратной скобкой необходимо выполнение неравенства $s < 0$. В этом случае степень ϱ перед второй квадратной скобкой будет отрицательна. Таким образом, за счет выбора параметра s добиться полиномиальности самой связи $H_\perp$ нельзя. Однако, связь можно умножить целиком на произвольный множитель, отличный от нуля. При этом поверхность в фазовом пространстве, определяемая данной связью не изменится. Минимальная степень ϱ, на которую необходимо умножить $H_\perp$ будет тогда, когда степени ϱ перед квадратными скобками будут равны. Отсюда следует равенство

$$s = \frac{2}{n-2}.$$

Тогда, умножив динамическую связь на степень ρ,

$$K_\perp := \varrho^{\frac{n-1}{n-2}} H_\perp, \tag{22.81}$$

получим эквивалентную полиномиальную связь

$$K_\perp = P^{\mu\nu}P_{\mu\nu} - \frac{n-2}{4(n-1)}\varrho^2 P^2 - \varrho^2 R^{(k)} - 2\varrho\partial_\mu(k^{\mu\nu}\partial_\nu\rho) + \frac{n-1}{n-2}(\partial\varrho)^2 = 0. \tag{22.82}$$

Кинематические связи в новых динамических переменных сохраняют свою полиномиальность:

$$H_\mu = -2\partial_\nu(P^{\nu\sigma}k_{\sigma\mu}) + P^{\nu\sigma}\partial_\mu k_{\nu\sigma} - \frac{n-2}{n-1}\partial_\mu(P\varrho) + P\partial_\mu\varrho = 0. \tag{22.83}$$

Исходя из явного выражения для новых канонических переменных (22.76)–(22.78), вычислим основные скобки Пуассона. Отличными от нуля являются только три скобки:

$$[k_{\mu\nu}, P'^{\varrho\sigma}] = \left(\delta^{\rho\sigma}_{\mu\nu} - \frac{1}{n-1}k_{\mu\nu}k^{\rho\sigma}\right)\delta(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.84}$$

$$[P^{\mu\nu}, P'^{\rho\sigma}] = \frac{1}{n-1}\Big(P^{\mu\nu}k^{\rho\sigma} - P^{\rho\sigma}k^{\mu\nu}\Big)\delta(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.85}$$

$$[\varrho, P'] = \delta(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.86}$$

где

$$\delta^{\rho\sigma}_{\mu\nu} := \frac{1}{2}\Big(\delta^\rho_\mu\delta^\sigma_\nu + \delta^\sigma_\mu\delta^\rho_\nu\Big).$$

Скобка Пуассона (22.84) не имеет канонического вида для фазовых переменных. Появление дополнительного слагаемого в (22.84) связано с тем, что на поля $k_{\mu\nu}$ и $P^{\mu\nu}$ наложены дополнительные условия (22.74).

Нетрудно проверить, что пуассонова структура, определяемая скобками Пуассона (22.84)–(22.86) является вырожденной. Это значит, что все многообразие $\mathbb{N}$, задаваемое координатами $(k_{\mu\nu}, P^{\mu\nu})$ и (ϱ, P) является не симплектическим, а только пуассоновым многообразием (см. раздел 18.3). На пуассоновом многообразии $\mathbb{N}$ существуют две функции Казимира:

$$c^1 := \det k_{\mu\nu}, \qquad c^2 := P^{\mu\nu}k_{\mu\nu}.$$

Действительно, из определения пуассоновой структуры (22.84), (22.86) следует, что все скобки Пуассона равны нулю:

$$[c^{1,2}, k'_{\mu\nu}] = [c^{1,2}, P'^{\mu\nu}] = [c^{1,2}, \varrho'] = [c^{1,2}, P'] = 0.$$

Отсюда вытекает, что скобка Пуассона $[c^{1,2}, f'] = 0$, где $f \in \mathcal{C}^1(\mathbb{N})$ – произвольная дифференцируемая функция на $\mathbb{N}$. Пуассонова структура, ограниченная на сечения $\mathbb{V} \subset \mathbb{N}$, которые определяются уравнениями $c^1 = \mathsf{const}$ и $c^2 = \mathsf{const}$, невырождена. Следовательно, эти сечения являются симплектическими. Новым фазовым пространством общей теории относительности в рассматриваемом случае является подмногообразие $\mathbb{V} \subset \mathbb{N}$, определяемое двумя фиксированными значениями функций Казимира (22.74). Строго говоря, каноническое преобразование (22.73) является каноническим, т.е. сохраняющим вид скобок Пуассона, только между исходным фазовым пространством и подмногообразием $\mathbb{V}$. Полиномиальность связей достигнута за

счет расширения исходного фазового пространства до пуассонова многообразия $\mathbb{N}$. Если решить дополнительные связи (22.74) явно, то полиномиальность будет нарушена. В этом нет ничего необычного. Например, электродинамика содержит связи, явное решение которых приводит даже к нелокальному действию для физических степеней свободы (см., например, [**?**]).

Рассмотрим алгебру связей. Поскольку вместо динамической связи $H_\perp$ мы ввели новую связь $K_\perp$, то алгебра связей изменится. Несложные вычисления приводят к равенствам

$$[K_\perp, K'_\perp] = -(\varrho^2 H^\mu + \varrho'^2 H'^\mu)\delta_\mu(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.87}$$

$$[K_\perp, H'_\mu] = -(K_\perp + K'_\perp)\delta_\mu(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.88}$$

$$[H_\mu, H'_\nu] = -H_\nu\delta_\mu(\boldsymbol{x}' - \boldsymbol{x}) - H'_\mu\delta_\nu(\boldsymbol{x}' - \boldsymbol{x}), \tag{22.89}$$

где

$$H^\mu := k^{\mu\nu}H_\nu, \qquad \delta_\mu(\boldsymbol{x}' - \boldsymbol{x}) := \frac{\partial}{\partial x'^\mu}\delta(\boldsymbol{x}' - \boldsymbol{x}).$$

По сравнению с исходной алгеброй (22.68)–(22.69), изменения касаются скобок Пуассона (22.87) и (22.88). Скобка Пуассона (22.87) является результатом прямого счета. Вторая скобка Пуассона (22.88) носит кинематический характер и определяется тем, что новая связь является не функцией, а скалярной плотностью.

Получим явные выражения для геометрических объектов в новых переменных. Сначала вычислим символы Кристоффеля:

$$\hat\Gamma_{\mu\nu}{}^\rho = \frac{1}{(n-2)\varrho}(\partial_\mu\varrho\delta^\rho_\nu + \partial_\nu\varrho\delta^\rho_\mu - k_{\mu\nu}k^{\rho\sigma}\partial_\sigma\varrho) + \Gamma^{(k)\rho}_{\mu\nu},$$

где "символы Кристоффеля" $\Gamma^{(k)\rho}_{\mu\nu}$ выражаются через тензорную плотность $k_{\mu\nu}$ по тем же формулам, что и символы Кристоффеля через метрику. Конечно, "символы Кристоффеля" $\Gamma^{(k)\rho}_{\mu\nu}$ никакой связности не определяют. Отсюда следует выражение для следа символов Кристоффеля, который определяет дополнительные слагаемые в ковариантных производных тензорных плотностей,

$$\hat\Gamma_\mu = \frac{n-1}{n-2}\frac{\partial_\mu\varrho}{\varrho},$$

так как

$$\Gamma^{(k)}_\mu := \Gamma^{(k)\nu}_{\nu\mu} = \frac{1}{2}k^{\nu\rho}\partial_\mu k_{\nu\rho} = 0$$

в силу условия $|\det k_{\mu\nu}| = 1$. Нетрудно проверить ковариантное постоянство новых переменных:

$$\hat\nabla_\mu k_{\nu\rho} = 0, \qquad \hat\nabla_\mu\varrho = 0.$$

Теперь вычислим тензор Риччи. Параметризация метрики (22.76) имеет вид преобразования Вейля (**??**) от "метрики" $k_{\mu\nu}$ к метрике $g_{\mu\nu}$. Разница заключается только в трансформационных свойствах множителей ϱ^s и $\mathsf{e}^{2\phi}$. Кроме того, поле $k_{\mu\nu}(x)$ является не тензором, а тензорной плотностью. Однако, мы можем воспользоваться формулой преобразования тензора Риччи (**??**) и выписать его в новых канонических переменных

$$\hat R_{\mu\nu} = R^{(k)}_{\mu\nu} + \frac{n-3}{n-2}\frac{\partial^2_{\mu\nu}\varrho}{\varrho} - \frac{(n-1)(n-3)}{(n-2)^2}\frac{\partial_\mu\varrho\partial_\nu\varrho}{\varrho^2} - \frac{n-3}{n-2}\Gamma^{(k)\rho}_{\mu\nu}\frac{\partial_\rho\varrho}{\varrho} - \\ - \frac{1}{(n-2)^2}k_{\mu\nu}\frac{\partial\varrho^2}{\varrho^2} + \frac{1}{n-2}k_{\mu\nu}\frac{\partial_\rho(k^{\rho\sigma}\partial_\sigma\varrho)}{\varrho}, \tag{22.90}$$

где введен "тензор Риччи" $R^{(k)}_{\mu\nu}$, построенный по "метрике" $k_{\mu\nu}$. Отсюда следует выражение для тензорной плотности скалярной кривизны, которая возникает после сворачивания тензора Риччи с $k^{\mu\nu}$:

$$\check{R} := k^{\mu\nu}\hat{R}_{\mu\nu} = \varrho^{\frac{2}{n-2}}\hat{R} = R^{(k)} + 2\frac{\partial_\mu(k^{\mu\nu}\partial_\nu\varrho)}{\varrho} - \frac{n-1}{n-2}\frac{(\partial\varrho)^2}{\varrho^2}.$$

После переопределения динамической связи (22.81) гамильтониан также равен линейной комбинации связей:

$$\mathcal{H} = \int d\boldsymbol{x}(\tilde{N}K_\perp + N^\mu H_\mu),$$

только с новым множителем Лагранжа $\tilde{N} := \varrho^{-\frac{n-1}{n-2}}N$. Соответствующее действие имеет вид

$$S_{\text{не}} = \int dx\Big(P^{\mu\nu}\dot{k}_{\mu\nu} + P\dot{\varrho} - \tilde{N}K_\perp - N^\mu H_\mu - \lambda(|\det k_{\mu\nu}| - 1) - \mu P^{\mu\nu}k_{\mu\nu}\Big),$$

где мы учли связи (22.74) с помощью множителей Лагранжа λ и μ. Уравнения движения для новых канонических переменных примут вид:

$$\dot{\varrho} = -\frac{n-2}{2(n-1)}\tilde{N}\varrho^2 P + \frac{n-2}{n-1}\hat{\nabla}_\mu N^\mu\varrho, \tag{22.91}$$

$$\dot{P} = \frac{n-2}{2(n-1)}\tilde{N}\varrho P^2 + 2\tilde{N}\varrho\check{R} + 2k^{\mu\nu}\hat{\nabla}_\mu\hat{\nabla}_\nu\tilde{N}\varrho + \frac{1}{n-1}\hat{\nabla}_\mu N^\mu P + N^\mu\hat{\nabla}_\mu P, \tag{22.92}$$

$$\dot{k}_{\mu\nu} = 2\tilde{N}P_{\mu\nu} + \hat{\nabla}_\mu N^\rho k_{\nu\rho} + \hat{\nabla}_\nu N^\rho k_{\mu\rho} - \frac{2}{n-1}\hat{\nabla}_\rho N^\rho k_{\mu\nu}, \tag{22.93}$$

$$\begin{aligned}\dot{P}^{\mu\nu} = {} & -2\tilde{N}P^{\mu\rho}P^\nu{}_\rho + \hat{\nabla}_\rho(N^\rho P^{\mu\nu}) + \frac{2}{n-1}\hat{\nabla}_\rho N^\rho P^{\mu\nu} - P^{\mu\rho}\hat{\nabla}_\rho N^\nu - P^{\nu\rho}\hat{\nabla}_\rho N^\mu - \\ & - \varrho^2 k^{\mu\rho}k^{\nu\sigma}\left(\tilde{N}\hat{R}_{\rho\sigma} + \hat{\nabla}_\rho\hat{\nabla}_\sigma\tilde{N}\right) + \frac{\varrho^2 k^{\mu\nu}}{n-1}\left(\tilde{N}\check{R} + k^{\rho\sigma}\hat{\nabla}_\rho\hat{\nabla}_\sigma\tilde{N}\right).\end{aligned} \tag{22.94}$$

Отметим, что множители Лагранжа λ и μ в уравнения движения вообще не входят, т.к. связи являются функциями Казимира. Эта система уравнений движения записана для тензорных плотностей:

$$\begin{aligned}\mathsf{deg}\, k_{\mu\nu} &= \frac{2}{n-1}, \qquad \mathsf{deg}\,\varrho = -\frac{n-2}{n-1},\\ \mathsf{deg}\, P^{\mu\nu} &= -\frac{n+1}{n-1}, \qquad \mathsf{deg}\, P = -\frac{1}{n-1}.\end{aligned}$$

Новый множитель Лагранжа также является тензорной плотностью:

$$\mathsf{deg}\,\tilde{N} = 1, \qquad \mathsf{deg}\, N^\mu = 0.$$

Поскольку степени тензорных плотностей при умножении складываются, то

$$\mathsf{deg}\,(\tilde{N}\varrho) = \frac{1}{n-1}, \qquad \mathsf{deg}\,(\tilde{N}\varrho k^{\mu\nu}) = -\frac{1}{n-1}.$$

Ковариантные производные от тензорных плотностей в системе уравнений (22.91)–(22.94) имеют вид:

$$\begin{aligned}\hat{\nabla}_\mu\tilde{N} &= \partial_\mu\tilde{N} + \frac{n-1}{n-2}\frac{\partial_\mu\varrho}{\varrho}\tilde{N},\\ \hat{\nabla}_\mu P &= \partial_\mu P - \frac{1}{n-2}\frac{\partial_\mu\varrho}{\varrho}P.\end{aligned}$$

Полученная система уравнений движения (22.91)–(22.94) является полиномиальной.

22.8 Проблема энергии в теории гравитации

Определение энергии гравитационного поля является одной из главных нерешенных проблем, которая привлекает большое внимание с момента создания общей теории относительности. Взгляд на эту проблему существенно меняется с течением времени, поэтому мы опишем несколько подходов.

Сначала сформулируем в чем именно заключается проблема. В механике точечных частиц, а также в теории поля в пространстве Минковского под энергией понимают численное значение гамильтониана системы. Если гамильтониан не зависит от времени явно, то энергия сохраняется. Закон сохранения энергии можно получить также из теоремы Нетер. Если действие инвариантно относительно трансляций во времени $t \mapsto t + \mathsf{const}$, то из первой теоремы Нетер следует закон сохранения энергии (см. раздел 19.2.1), при этом компонента $T_0{}^0$ энергии-импульса в теории поля совпадает с определением гамильтоновой плотности системы полей. Этот подход к определению энергии в теории гравитации приводит к ответу, мало удовлетворительному с физической точки зрения. В теории гравитации действие инвариантно относительно общих преобразований координат и, тем более, относительно трансляций во времени. С другой стороны, мы уже показали в разделе 20.2.4, что канонический гамильтониан для любой модели, инвариантной относительно произвольного невырожденного преобразования временно́й координаты пропорционален связи и равен нулю на уравнениях движения. Это значит, что формальный подход к определению энергии в теории гравитации дает нуль для любого решения уравнений движения. Этот результат плохо согласуется с нашим интуитивным представлением об энергии, так как наличие материи ассоциируется с наличием энергии. Заметим также, что в пространстве-времени Минковского действие инвариантно относительно трансляций только в декартовой системе координат. В общековариантных моделях гравитации действие инвариантно относительно трансляций в произвольной криволинейной системе координат.

Замечание. В настоящем разделе мы рассмотрим определение энергии в общей теории относительности, то есть при нулевых тензорах кручения и неметричности. Обобщение определения на более общий случай аффинной геометрии проблемы не составляет, так как основная трудность, связанная с общей ковариантностью уравнений движения, присутствует во всех моделях. □

22.8.1 Тензор энергии-импульса полей материи

Первые попытки определить энергию, как сохраняющуюся величину для системы полей материи и гравитационного поля, были предприняты на заре исследований по общей теории относительности. В этой модели действие представляет собой сумму действия Гильберта–Эйнштейна (21.6) и действия для полей материи $S_{\text{м}}$:

$$S = S_{\text{не}} + S_{\text{м}}. \tag{22.95}$$

Предположим, что действие полей материи зависит только от полей материи и метрики. Тогда вариация действия по метрике приводит к уравнениям Эйнштейна (21.1), где

$$T_{\text{м}\alpha\beta} := \frac{2}{\sqrt{|g|}} \frac{\delta S_{\text{м}}}{\delta g^{\alpha\beta}} \tag{22.96}$$

– тензор энергии-импульса полей материи. В общем случае мы предполагаем, что действие для полей материи в моделях гравитации получено из действия, записанного в плоском пространстве-времени Минковского, путем минимальной подстановки, т.е. замены обычных производных на ковариантные $\partial_\alpha \mapsto \nabla_\alpha$ и метрики Минковского на нетривиальную метрику пространства-времени $\eta_{ab} \mapsto g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b \eta_{ab}$. Возможно также введение инвариантных слагаемых неминимального взаимодействия, которые обращаются в нуль в пространстве Минковского. В таких случаях действие зависит только от самих полей материи и репера. Соответствующее выражение для тензора энергии-импульса можно записать через вариационную производную по реперу:

$$T_{\text{м}\alpha\beta} = \frac{2}{\sqrt{|g|}} \frac{\delta S_{\text{м}}}{\delta e^\alpha{}_a} e_{\beta a}. \tag{22.97}$$

Если действие для полей материи может быть записано через метрику, то данное определение совпадает с (22.96) и приводит к симметричному тензору энергии-импульса. В общем случае это не так. Например, лагранжиан спинорного поля не может быть записан через метрику, так как лоренцева связность выражается через репер, и ее нельзя выразить через компоненты метрики. Определение тензора энергии-импульса через репер (22.97) не всегда приводит к симметричному тензору энергии-импульса.

В разделе (21.4) было показано, что инвариантное действие приводит к ковариантному закону сохранения тензора энергии-импульса:

$$\nabla_\beta T_{\text{м}\alpha}{}^\beta = 0. \tag{22.98}$$

Однако это ковариантное равенство не является законом сохранения. Покажем это. Прямые вычисления приводят к равенству

$$\nabla_\beta T_{\text{м}\alpha}{}^\beta = \frac{1}{\sqrt{|g|}} \partial_\beta \Big(\sqrt{|g|} T_{\text{м}\alpha}{}^\beta \Big) - \frac{1}{2} \partial_\alpha g_{\beta\gamma} T_{\text{м}}{}^{\beta\gamma} = 0. \tag{22.99}$$

При интегрировании этого равенства по объему первое слагаемое для каждого значения индекса α преобразуется с помощью формулы Стокса (4.80) в интеграл по граничной поверхности:

$$\int_{\mathbb{M}} dx \partial_\beta (\sqrt{|g|} T_{\text{м}\alpha}{}^\beta) = \int_{\partial\mathbb{M}} ds_\beta \sqrt{|g|} T_{\text{м}\alpha}{}^\beta,$$

где ds_α – ориентированный элемент граничной гиперповерхности $\partial\mathbb{M}$. Этот интеграл привел бы к закону сохранения в выбранной системе координат, если бы не наличие второго слагаемого в равенстве (22.99).

22.8.2 Псевдотензор энергии-импульса для гравитации

Для решения проблемы сохранения энергии к тензору энергии-импульса материи, который мы будем записывать с одним верхним и одним нижним индексом $T_{\text{м}\alpha}{}^\beta$, было предложено добавить некоторый объект с компонентами $t_\alpha{}^\beta$, зависящими только от метрики и ее первых производных, таким образом, чтобы было выполнено равенство

$$\partial_\beta \Big[\sqrt{|g|} (T_{\text{м}\alpha}{}^\beta + t_\alpha{}^\beta) \Big] = 0. \tag{22.100}$$

Объект с компонентами $t_\alpha{}^\beta$ называется *псевдотензором энергии-импульса* гравитационного поля. При этом равенство (22.100) принято рассматривать, как закон сохранения энергии-импульса полей материи и гравитационного поля. Приставка "псевдо" в данном случае означает, что для выполнения равенства (22.100) компоненты $t_\alpha{}^\beta$ не могут образовывать тензор.

Задача о нахождении явного вида псевдотензора энергии-импульса гравитационного поля $t_\alpha{}^\beta$ может быть решена следующим образом. Закон сохранения (22.100) имеет тот же вид, что и закон сохранения энергии-импульса, вытекающий из первой теоремы Нетер (19.44). Из инвариантности полного действия (22.95) относительно трансляций в пространстве-времени следует закон сохранения

$$\partial_\beta(\sqrt{|g|}T_\alpha{}^\beta) = 0, \tag{22.101}$$

где суммарный тензор энергии-импульса материи и гравитационного поля, умноженный на определитель репера, имеет вид

$$\begin{aligned}\sqrt{|g|}T_\alpha{}^\beta &= \partial_\alpha g^{\gamma\delta}\frac{\partial L_{\text{не}}}{\partial(\partial_\beta g^{\gamma\delta})} + \partial_\alpha\varphi^a\frac{\partial L_{\text{м}}}{\partial(\partial_\beta\varphi^a)} - \delta^\beta_\alpha(L_{\text{не}} + L_{\text{м}}) \\ &= \sqrt{|g|}T^{(\text{c})\beta}_{\text{м}\alpha} + \partial_\alpha g^{\gamma\delta}\frac{\partial L_{\text{не}}}{\partial(\partial_\beta g^{\gamma\delta})} - \delta^\beta_\alpha L_{\text{не}}.\end{aligned} \tag{22.102}$$

В этом выражении $T^{(\text{c})\beta}_{\text{м}\alpha}$ – канонический тензор энергии-импульса полей материи, определенный соотношением

$$\sqrt{|g|}T^{(\text{c})\beta}_{\text{м}\alpha} := \partial_\alpha\varphi^a\frac{\partial L_{\text{м}}}{\partial(\partial_\beta\varphi^a)} - \delta^\beta_\alpha L_{\text{м}}.$$

В том случае, когда определение тензора энергии-импульса (22.96) является ковариантным обобщением канонического тензора энергии-импульса материи в пространстве Минковского, имеет место формула

$$T_{\text{м}\alpha}{}^\beta = T^{(\text{c})\beta}_{\text{м}\alpha}.$$

Тогда последние два слагаемых в выражении (22.102), зависящие только от метрики и ее первых производных, можно принять за определение псевдотензора энергии-импульса гравитационного поля

$$\sqrt{|g|}t_\alpha{}^\beta := \partial_\alpha g^{\gamma\delta}\frac{\partial L_{\text{не}}}{\partial(\partial_\beta g^{\gamma\delta})} - \delta^\beta_\alpha L_{\text{не}}. \tag{22.103}$$

Напомним, что под лагранжианом $L_{\text{не}}$ понимается функция, полученная из $\kappa\sqrt{|g|}\tilde{R}(g)$ добавлением полной производной, которая приводит к сокращению вторых производных от метрики (см. раздел 22.1). Полученное выражение для псевдотензора энергии-импульса широко использовалось, в том числе классиками науки: Г. Вейлем [?], П. Дираком [?], В. Паули [?], Э. Шредингером [?] и др. После несложных вычислений можно получить явное выражение для псевдотензора энергии-импульса

$$\sqrt{|g|}t_\alpha{}^\beta = \kappa(\Gamma_{\gamma\delta}{}^\beta - \delta^\beta_\gamma\Gamma_{\delta\epsilon}{}^\epsilon)\partial_\alpha\Big(\sqrt{|g|}g^{\gamma\delta}\Big) - \delta^\beta_\alpha L_{\text{не}}. \tag{22.104}$$

Отметим недостатки такого определения псевдотензора энергии-импульса гравитационного поля. Из полученного выражения следует, что в нормальной системе

координат, где все символы Кристоффеля обращаются в нуль в некоторой точке, все компоненты псевдотензора энергии-импульса гравитационного поля равны нулю. То есть, независимо от кривизны пространства, в любой заданной точке пространства-времени можно обратить в нуль все компоненты псевдотензора энергии-импульса. В то же время в плоском пространстве-времени (отсутствие гравитационного поля) в криволинейной системе координат символы Кристоффеля и, следовательно, компоненты псевдотензора энергии-импульса в общем случае отличны от нуля. Отметим также, что компонента тензора энергии-импульса $T_0{}^0$ в законе сохранения (22.101) совпадает с гамильтонианом системы и обращается в нуль на уравнениях движения. Эти замечания ставят под сомнение возможность физической интерпретации псевдотензора энергии-импульса гравитационного поля.

Если равенство (22.100) принято за определение псевдотензора энергии-импульса гравитационного поля, то последний определен неоднозначно. Очевидно, что псевдотензор

$$t'_\alpha{}^\beta = t_\alpha{}^\beta + \partial_\gamma B_\alpha{}^{\gamma\beta}, \tag{22.105}$$

где $B_\alpha{}^{\gamma\beta} = -B_\alpha{}^{\beta\gamma}$ – произвольный "тензор" третьего ранга, антисимметричный по верхним индексам, также удовлетворяет закону сохранения. В процессе исследования общей теории относительности было предложено несколько явных выражений для псевдотензора энергии-импульса, которые обладают своими достоинствами и недостатками и отличаются между собой на некоторый "тензор" $B_\alpha{}^{\gamma\beta}$. Мы не будем останавливаться на обсуждении различных подходов, а отметим лишь общий недостаток. Псевдотензор энергии-импульса $t_\alpha{}^\beta$ не является тензором. Это значит, что энергия в различных системах координат может быть положительна, равна нулю или отрицательна, что является неудовлетворительным с физической точки зрения.

Аналогичное построение псевдотензора энергии-импульса гравитационного поля можно провести в реперном формализме. При этом все недостатки псевдотензора сохраняются.

Проведенное обсуждение показывает, что в теории гравитации невозможно дать локальное определение энергии-импульса, удовлетворив одновременно двум условиям: 1) компоненты энергии-импульса должны образовывать тензорное поле второго ранга и 2) должен быть выполнен закон сохранения (22.100).

Другой подход к законам сохранения связан с симметриями пространства-времени. Если метрика пространства-времени допускает группу движений, то появляется возможность определить законы сохранения для тензора энергии-импульса полей материи без привлечения понятия псевдотензора энергии-импульса гравитационного поля. Пусть тензор энергии-импульса полей материи – это симметричный тензор второго ранга $T_{\text{м}}^{\alpha\beta} = T_{\text{м}}^{\beta\alpha}$, удовлетворяющий условию (22.98). Предположим, что метрика $g_{\alpha\beta}$ имеет N векторов Киллинга $K_{\text{A}} = K_{\text{A}}{}^\alpha \partial_\alpha$, $\text{A} = 1, \ldots, \text{N}$. Рассмотрим N векторов

$$P_{\text{A}}^\alpha = K_{\text{A}}{}^\beta T_{\text{м}\beta}{}^\alpha.$$

Тогда

$$\nabla_\alpha P^\alpha = \nabla_\alpha K_{\text{A}}{}^\beta T_{\text{м}\beta}{}^\alpha + K_{\text{A}}{}^\beta \nabla_\alpha T_{\text{м}\beta}{}^\alpha = 0.$$

Здесь первое слагаемое равно нулю как следствие симметричности тензора энергии-импульса и уравнения Киллинга (16.4). Второе слагаемое обращается в нуль в силу уравнения (22.98). С другой стороны, справедливо тождество (7.54)

$$\nabla_\alpha P^\alpha = \frac{1}{\sqrt{|g|}} \partial_\alpha (\sqrt{|g|} P^\alpha).$$

Поэтому, если $\mathbb{M}$ – компактная ориентируемая область пространства-времени с краем $\partial\mathbb{M}$, то

$$\int_{\mathbb{M}} dx\,\sqrt{|g|}\nabla_\alpha P^\alpha = \int_{\partial\mathbb{M}} ds_\alpha P^\alpha = 0, \tag{22.106}$$

где ds_α – ориентированный элемент площади края. Таким образом каждому вектору Киллинга соответствует закон сохранения. В частности, если существует вектор Киллинга $K_0 = \partial_0$ и поля материи исчезают на пространственной бесконечности, то ему соответствует закон сохранения энергии $\partial_0 E = 0$, где

$$E = \int d\boldsymbol{x}\, P^0 = \int d\boldsymbol{x}\, T_{\text{м}0}{}^0.$$

Отметим, что в этом выражении присутствует гамильтонова плотность только для полей материи.

В четырехмерном пространстве-времени Минковского метрика имеет десять векторов Киллинга: четыре вектора соответствуют трансляциям и шесть – (псевдо-)вращениям. Этим векторам Киллинга соответствуют законы сохранения энергии-импульса и момента количества движения.

22.8.3 Полная гравитационная энергия асимптотически плоского пространства-времени

Дальнейшее развитие общей теории относительности изменило подход к определению полной энергии гравитационного поля и материи. Это определение основано на сферически симметричном решении Шварцшильда и учете граничных вкладов в действие Гильберта–Эйнштейна.

Запишем решение Шварцшильда (**??**) в асимптотически декартовой системе координат, где координаты x, y, z связаны с координатами r, θ, φ обычными формулами трехмерного евклидова пространства. С этой целью воспользуемся формулами (8.10) для перехода от сферических координат к декартовым:

$$\begin{aligned} ds^2 &= \left(1 - \frac{2M}{r}\right) dt^2 - \frac{dr^2}{1 - \frac{2M}{r}} - r^2 d(d\theta^2 + \sin^2\theta d\varphi^2) = \\ &= \left(1 - \frac{2M}{r}\right) dt^2 - (dx^2 + dy^2 + dz^2) - \frac{2M}{r^2(r-2M)}(xdx + ydy + xdz)^2. \end{aligned} \tag{22.107}$$

Метрика Шварцшильда на больших расстояниях в первом порядке по $M/r \to 0$ принимает вид

$$\begin{aligned} ds^2 \approx \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 + \frac{2M}{r}\right)(dx^2 + dy^2 + dz^2) - \\ - \frac{4M}{r}(xydxdy + xzdxdz + yzdydz). \end{aligned} \tag{22.108}$$

Очевидно, что в нулевом порядке эта метрика совпадает с метрикой Лоренца. Последнее выражение используется для определения асимптотически плоского пространства-времени.

Определение. Топологически тривиальное пространство-время $\mathbb{M} \approx \mathbb{R}^{1,3}$ называется *асимптотически плоским*, если существует такая система координат, что его

метрика при $r \to \infty$, где $r := \sqrt{x^2+y^2+z^2}$ и всех моментов времени $t \in \mathbb{R}$, достаточно быстро стремится к метрике Шварцшильда (22.108), записанной в декартовой системе координат.

Максимально продолженное пространство-время Шварцшильда описывает черные дыры и не является топологически тривиальным. Понятие асимптотически плоского пространства-времени просто обобщается на многообразия с нетривиальной топологией.

Определение. Пусть пространство-время имеет вид прямого произведения $\mathbb{M} = \mathbb{R} \times \mathbb{U}$, где явно выделено время $x^0 = t \in \mathbb{R}$. Допустим, что существует компактное подмножество $\mathbb{K} \subset \mathbb{U}$ такое, что разность $\mathbb{U} \setminus \mathbb{K}$ представляет собой несвязное объединение открытых множеств $\mathbb{S}_{\mathrm{A}}$, $\mathrm{A} = 1, \ldots, \mathrm{N}$. Если в каждом произведении $\mathbb{R} \times \mathbb{S}_{\mathrm{A}}$ существует такая система координат, что метрика при $r \to \infty$ и всех $t \in \mathbb{R}$ достаточно быстро стремится к метрике Шварцшильда (22.108) в декартовой системе координат, то пространство-время $\mathbb{M}$ называется *асимптотически плоским.* □

Рассмотрим действие Гильберта–Эйнштейна в гамильтоновой форме, которое вытекает из выражения для гамильтониана (22.63),

$$S_{\text{не}} = \int d^4x \, (p^{\mu\nu}\dot{g}_{\mu\nu} - NH_\perp - N^\mu H_\mu) \,,$$

где функции хода и сдвига рассматриваются в качестве множителей Лагранжа. Динамическая связь $H_\perp$ (22.62) содержит трехмерную скалярную кривизну и, следовательно, вторые производные от пространственных компонент метрики $g_{\mu\nu}$. При вариации действия $L_{\text{не}}$ по компонентам метрики приходится интегрировать по частям, отбрасывая граничные члены. Ниже будет показано, что в асимптотически плоском пространстве-времени (для решения Шварцшильда) граничный вклад, возникающий при интегрировании по частям слагаемого со вторыми производными, отличен от нуля. В настоящее время это граничное слагаемое принято в качестве определения полной энергии асимптотически плоского распределения масс. Рассмотрим данный вопрос подробно.

Воспользуемся формулами (22.2)–(22.4) и запишем трехмерную скалярную кривизну, умноженную на определитель репера, в виде

$$\hat{e}\hat{R} = \partial_\mu \left[\hat{e}(\hat{g}^{\mu\nu}\hat{\Gamma}_{\rho\nu}{}^{\rho} - \hat{g}^{\nu\rho}\hat{\Gamma}_{\nu\rho}{}^{\mu})\right] + \hat{e}\hat{L}_{\text{не}},$$

где последнее слагаемое $\hat{L}_{\text{не}}$ квадратично по символам Кристоффеля (его вид в настоящий момент не важен). Первое слагаемое в этом представлении имеет вид дивергенции, которую можно выразить через компоненты метрики:

$$\partial_\mu \left[\hat{e}\hat{g}^{\mu\nu}\hat{g}^{\rho\sigma}(\partial_\nu g_{\rho\sigma} - \partial_\rho g_{\nu\sigma})\right],$$

где мы воспользовались выражением символов Кристоффеля через метрику (7.22). При интегрировании этого слагаемого по частям возникает граничное слагаемое, которое, как мы увидим ниже, отлично от нуля. Поэтому для того, чтобы компенсировать его вклад, добавим к исходному действию граничный член с самого начала:

$$S_{\text{не}} \mapsto S_{\text{не}} + \kappa \int d^4x \, \partial_\mu B^\mu,$$

где

$$B^\mu := N\hat{e}\hat{g}^{\mu\nu}\hat{g}^{\rho\sigma}(\partial_\nu g_{\rho\sigma} - \partial_\rho g_{\nu\sigma}) \tag{22.109}$$

и мы восстановили гравитационную константу связи κ.

Определение. *Полной энергией* асимптотически плоского распределения масс называется интеграл

$$E := \kappa \int d^3x\, \partial_\mu B^\mu, \tag{22.110}$$

где компоненты B^μ определены в (22.109), и интегрирование проводится по пространственноподобному сечению $x^0 = \mathsf{const}$. □

Компоненты B^μ образуют вектор относительно глобальных преобразований координат в пространстве $\mathbb{R}^3$. В целом, определение полной энергии неинвариантно и зависит от выбора системы координат.

Для трехмерного евклидова пространства $\mathbb{R}^3$ в декартовой системе координат метрика имеет вид $g_{\mu\nu} = -\delta_{\mu\nu}$. Поэтому $B^\mu = 0$ и, следовательно, полная энергия плоского пространства равна нулю, $E = 0$. Уже на этом этапе видна важность выбора системы координат в определении энергии (22.110). Действительно, если система координат не является асимптотически декартовой, то полная энергия трехмерного евклидова пространства может быть отлична от нуля.

Вычислим полную энергию для решения Шварцшильда и, тем самым, для произвольного асимптотически плоского пространства-времени в первом порядке по $M/r \to \infty$. Для этого запишем решение Шварцшильда в декартовой системе координат (22.107). Пространственная часть метрики отличается от евклидовой метрики: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, где

$$h_{\mu\nu} = -\frac{2M}{r^2(r-2M)} \begin{pmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{pmatrix}$$

На больших расстояниях, $r \gg M$, эта поправка к метрике стремится к нулю. В определении вектора B^μ (22.109) выражение в скобках имеет первый порядок малости. Поэтому остальные сомножители достаточно участь в нулевом порядке:

$$N = 1, \quad \hat{e} = 1, \quad \hat{g}^{\mu\nu} = \eta^{\mu\nu}.$$

Простые вычисления дают следующие выражения для компонент вектора B^μ в первом порядке по M/r:

$$B^x \simeq \frac{4Mx}{r^3}, \quad B^y \simeq \frac{4My}{r^3}, \quad B^z \simeq \frac{4Mz}{r^3}.$$

Нормальный единичный ковектор к сфере с центром в начале координат евклидова пространства имеет вид

$$\{n_\mu\} = \left\{\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right\}, \quad n^2 = -1.$$

Поэтому

$$B^\mu n_\mu \simeq \frac{4M}{r^2}.$$

Теперь можно вычислить полную энергию:

$$E = \kappa \int d^3x\, \partial_\mu B^\mu = \int_0^\pi d\theta \int_0^{2\pi} d\varphi\, r^2 \sin\theta (B^\mu n_\mu) \simeq 16\pi\kappa M. \tag{22.111}$$

Чтобы придать полученному выражению более привлекательный вид, восстановим размерные константы. Это полезно делать хотя бы изредка. Компоненты метрики, по-определению, являются безразмерными $[g_{\alpha\beta}] = 1$. Поскольку

$$\left[\frac{M}{r}\right] = \frac{\text{г}}{\text{см}},$$

то это отношение необходимо обезразмерить, умножив на некоторую комбинацию гравитационной постоянной и скорости света, которые имеют следующие размерности:

$$[G] = \frac{\text{см}^3}{\text{г} \cdot \text{сек}^2}, \qquad [c] = \frac{\text{см}}{\text{сек}}.$$

Это можно сделать единственным образом путем замены

$$M \mapsto \frac{GM}{c^2}. \tag{22.112}$$

Если воспользоваться выражением (21.105) для константы связи κ через гравитационную постоянную G, которое было получено при рассмотрении ньютонова предела в общей теории относительности, то выражение для полной энергии примет вид

$$E = Mc^2. \tag{22.113}$$

То, что гравитационная энергия асимптотически плоского пространства-времени совпадает с энергией покоя для массы в решении Шварцшильда, привлекательно с физической точки зрения и оправдывает определение полной энергии через поверхностный интеграл (22.110). Численное значение полной гравитационной энергии зависит от выбора системы координат, т.к. данное определение не инвариантно относительно преобразований координат.

Поскольку M – постоянная интегрирования уравнений Эйнштейна, то полная энергия в асимптотически плоском пространстве-времени сохраняется.

В общем случае полная энергия не может быть положительно определена, поскольку постоянная M в решении Шварцшильда является константой интегрирования и может иметь любой знак.

Глава 23

Скалярные и калибровочные поля

Скалярные и электромагнитное поля, вместе со спинорными полями, являются важнейшими объектами квантовой теории поля. В настоящей главе мы напомним основные свойства этих полей в пространстве Минковского, рассмотрим эти поля на произвольных многообразиях с заданной аффинной геометрией и обсудим их уравнения в общей теории относительности.

23.1 Действительное скалярное поле

23.1.1 Скалярное поле в пространстве Минковского

Рассмотрим действительное скалярное поле (т.е. функцию) $\varphi(x) \in \mathcal{C}^2(\mathbb{R}^{1,n-1})$ в пространстве Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений. Обозначим декартовы координаты через x^α, $\alpha = 0, 1, \ldots, n-1$. Свободное скалярное поле описывается квадратичным лагранжианом

$$L = \frac{1}{2}\eta^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - \frac{1}{2}m^2\varphi^2, \tag{23.1}$$

где $\eta^{\alpha\beta} := \mathsf{diag}\,(+ - \ldots -)$ – обратная метрика Лоренца и $m = \mathsf{const} > 0$ – масса скалярного поля. Соответствующее действие имеет вид

$$S = \int dx L.$$

Знаки слагаемых в лагранжиане выбран таким образом, чтобы канонический гамильтониан, рассмотренный ниже, был положительно определен.

Посчитаем размерности. Действие и компоненты метрики, по-определению, безразмерны. Координаты имеют размерность длины, $[x^\alpha] = l$, $\alpha = 0, 1, \ldots, n-1$ (скорость света мы положили равной единице). Поэтому скалярное поле и масса имеют следующие размерности:

$$[\varphi] = l^{\frac{2-n}{2}}, \qquad [m] = l^{-1}. \tag{23.2}$$

Уравнение движения для свободного скалярного поля линейно:

$$S,_\varphi := \frac{\delta S}{\delta\varphi} = -(\Box + m^2)\varphi = 0, \tag{23.3}$$

где

$$\Box := \eta^{\alpha\beta}\partial_\alpha\partial_\beta = \partial_t^2 - \partial_1^2 - \ldots - \partial_{n-1}^2$$

– волновой оператор (оператор Даламбера). Это уравнение называется *уравнением Клейна–Гордона–Фока*, которое при $m = 0$ сводится к уравнению Даламбера. Уравнение (23.3) было предложено независимо в работах [**?**, **?**, **?**, **?**].

Скалярное поле называется свободным, потому что описывается линейным уравнением движения.

Поскольку свободное скалярной поле удовлетворяет волновому гиперболическому уравнению, то для постановки задачи Коши в полупространстве $x^0 \geq 0$ для однозначного определения решения необходимо задать начальные данные на гиперповерхности $x^0 = 0$ для поля φ и его производной по времени $\dot{\varphi} := \partial_0\varphi$. Говорят, что скалярное поле описывает одну динамическую степень свободы. В квантовой теории поля оно описывает нейтральные скалярные частицы.

Для того, чтобы описать скалярное поле с самодействием, к лагранжиану (23.1) добавляется потенциал взаимодействия $V(\varphi)$,

$$L \mapsto L = \frac{1}{2}\eta^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V(\varphi),$$

где $V(\varphi)$ – некоторая достаточно гладкая функция одного аргумента. Для упрощения обозначений мы включили массовый член в определение потенциала $V(\varphi)$. В общем случае функция $V(\varphi)$ зависит от некоторого набора размерных или безразмерных констант связи. Для самодействующего скалярного поля уравнение движения становится нелинейным:

$$S,_\varphi := \frac{\delta S}{\delta\varphi} = -\Box\varphi - V'(\varphi) = 0, \tag{23.4}$$

где $V' := dV/d\varphi$.

Действие для скалярного поля инвариантно относительно глобального действия группы Пуанкаре $\mathbb{IO}(1, n-1)$. Согласно первой теореме Нетер эта инвариантность приводит к законам сохранения энергии-импульса и момента количества движения (см. разделы **??**). Выражение для сохраняющегося канонического тензора энергии-импульса (21.162) скалярного поля имеет вид

$$T^{(c)}_{\alpha\beta} = \partial_\alpha\varphi\partial_\beta\varphi - \eta_{\alpha\beta}\left(\frac{1}{2}\partial\varphi^2 - V\right), \tag{23.5}$$

где использовано сокращенное обозначение

$$\partial\varphi^2 := \eta^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi.$$

Тензор энергии-импульса, очевидно, симметричен по своим индексам.

Спиновый момент скалярного поля (19.53) равен нулю, $S_{\alpha\beta}{}^\gamma = 0$, т.к. при вращениях равна нулю вариация формы скалярного поля. Отсюда вытекает, что в квантовой теории поля скалярное поле описывает частицы с нулевым спином. Момент количества движения скалярного поля полностью определяется орбитальным моментом по формуле (19.51)

$$J_{\alpha\beta}{}^\gamma = M_{\alpha\beta}{}^\gamma := x_\beta T_\alpha{}^\gamma - x_\alpha T_\beta{}^\gamma. \tag{23.6}$$

Функция $V(\varphi)$ может быть положительно определена и при "неправильном" знаке квадрата массы, $m^2 < 0$. Например, положим (см. рис.23.1,*a*)

$$V(\varphi) = \frac{1}{4}\lambda(\varphi^2 - a^2)^2 = \frac{1}{4}\lambda(\varphi^4 - 2\varphi^2a^2 + a^4), \qquad a > 0. \tag{23.7}$$

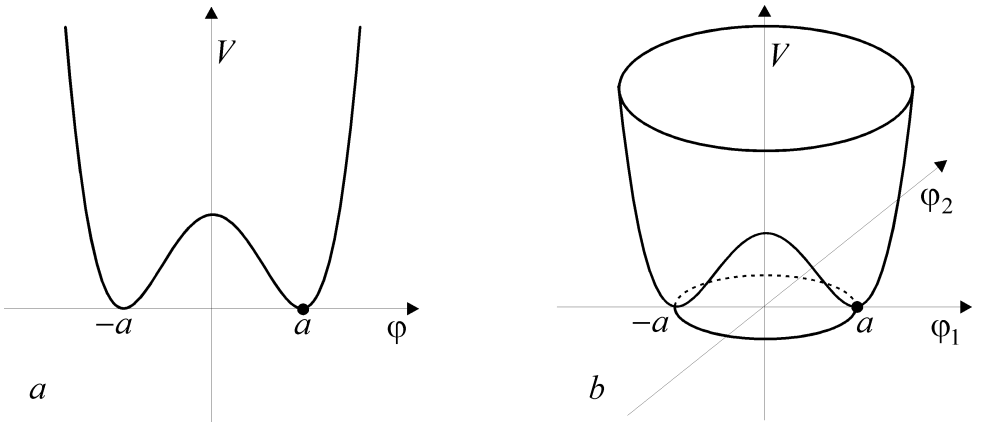


Рис. 23.1: Минимум потенциала действительного (a) и комплексного (b) скалярного поля.

При $\lambda > 0$ потенциал положительно определен, однако " квадрат массы", $m^2 = -\lambda a^2$, отрицателен. Здесь возникает вопрос: "Что называется массой частицы, соответствующей полю φ?" В квантовой теории поля масса частиц определяется квадратичным приближением лагранжиана, вокруг которого строится теория возмущений. Для этого необходимо выбрать вакуумное решение уравнений движения, вблизи которого будет происходить разложение полей. Простейшее решение уравнения движения (23.4) – это постоянное поле $\varphi_0 =$ const, которое удовлетворяет условию

$$V'(\varphi_0) = 0.$$

То есть мы предполагаем, что вакуум однороден и статичен. Для потенциала (23.7) уравнение равновесия принимает вид

$$\varphi(\varphi^2 - a^2) = 0.$$

Оно имеет три решения: $\varphi_0 = 0$, $\varphi_0 = \pm a$. Решение $\varphi_0 = 0$ соответствует неустойчивому положению равновесия. Именно поэтому вблизи данного решения квадрат массы имеет неправильный знак. Оба решения $\varphi_0 = \pm a$ соответствуют устойчивым положениям равновесия и имеют одинаковую нулевую плотность энергии. Любое из этих решений можно выбрать в качестве вакуума, вокруг которого строится теория возмущений. Положим

$$\varphi = a + \tilde{\varphi}.$$

Тогда потенциал примет вид

$$V = \lambda \left(a^2\tilde{\varphi}^2 + a\tilde{\varphi}^3 + \frac{1}{4}\tilde{\varphi}^4 \right).$$

Из квадратичного приближения следует, что масса скалярного поля положительна, $m^2 = 2\lambda a^2$. То есть масса частиц положительна, что соответствует устойчивому положению равновесия.

Перейдем к канонической формулировке. Импульс, канонически сопряженный скалярному полю, равен производной по времени от скалярного поля

$$p := \frac{\partial L}{\partial(\dot{\varphi})} = \dot{\varphi}. \tag{23.8}$$

Одновременны́е скобки Пуассона для канонически сопряженных переменных имеют вид

$$[\varphi, p'] := [\varphi(t, \boldsymbol{x}), p(t, \boldsymbol{x}')] = \delta(\boldsymbol{x} - \boldsymbol{x}'),$$

где введена пространственная δ-функция

$$\delta(\boldsymbol{x} - \boldsymbol{x}') := \delta(x^1 - x'^1) \dots \delta(x^{n-1} - x'^{n-1}).$$

Соответствующая плотность гамильтониана равна

$$H = T^{(c)}_{00} = \frac{1}{2}p^2 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + V(\varphi), \tag{23.9}$$

где греческие буквы из середины алфавита пробегают, как и раньше, только пространственные значения: $\mu, \nu = 1, \dots, n-1$. Если функция $V(\varphi)$ положительно определена, то гамильтонова плотность H также положительно определена, т.к. $\eta^{\mu\nu} := -\delta^{\mu\nu}$.

Гамильтониан скалярного поля получается интегрированием гамильтоновой плотности по всему пространству:

$$\mathcal{H} = \int d\boldsymbol{x} H. \tag{23.10}$$

Если этот интеграл сходится, то полная энергия скалярного поля сохраняется.

Уравнение движения второго порядка (23.3) в гамильтоновой форме эквивалентно системе двух уравнений движения первого порядка относительно производных по времени:

$$\begin{aligned} \dot\varphi &= [\varphi, \mathcal{H}] = p, \\ \dot p &= [p, \mathcal{H}] = \triangle\varphi - m^2\varphi - V', \end{aligned} \tag{23.11}$$

где $\triangle$ – оператор Лапласа.

Компонента $T^{(c)}_{00}$ тензора энергии-импульса скалярного поля (23.5) совпадает с плотностью гамильтониана (23.9). Остальные компоненты канонического тензора энергии-импульса также могут быть выражены через канонические переменные:

$$\begin{aligned} H_\mu &:= T^{(c)}_{0\mu} = p\partial_\mu\varphi, \\ T_{\mu\nu} &= \partial_\mu\varphi\partial_\nu\varphi - \left[\frac{1}{2}(p^2 + \partial^\rho\varphi\partial_\rho\varphi - V(\varphi)\right]\eta_{\mu\nu}. \end{aligned}$$

Ковектор энергии-импульса $\{P_\alpha\} = \{P_0 := \mathcal{H}, P_\mu\}$ для скалярного поля получается из компонент канонического тензора энергии-импульса интегрированием по пространству:

$$P_0 := \int d\boldsymbol{x}\, H, \qquad P_\mu := \int d\boldsymbol{x}\, H_\mu, \tag{23.12}$$

где интегрирование проводится по сечениям $x^0 = \mathsf{const}$. Численное значение P_0 и P_μ дают полную энергию и полный импульс скалярного поля. Если поле достаточно быстро убывает на бесконечности, что соответствует, в частности, отсутствию излучения, то полные энергия и импульс сохраняются:

$$\dot P_0 = 0, \qquad \dot P_\mu = 0.$$

Отсюда следует, что компоненты канонического тензора энергии-импульса $T^{(c)}_{0\mu}$ имеют физический смысл плотности импульса скалярного поля.

Полный момент импульса скалярного поля получается интегрированием нулевых компонент тензора момента количества движения (23.6):

$$\mathcal{J}_{\alpha\beta} := \int d\boldsymbol{x} J_{\alpha\beta}{}^{0}.$$

Если скалярное поле достаточно быстро убывает на пространственной бесконечности, то он также сохраняется,

$$\dot{\mathcal{J}}_{\alpha\beta} = 0.$$

Строго говоря, полным моментом импульса являются только пространственные компоненты $\mathcal{J}_{\mu\nu} = -\mathcal{J}_{\nu\mu}$.

23.1.2 Скалярное поле в аффинной геометрии

Пусть на многообразии (пространстве-времени) $\mathbb{M}$ произвольной размерности n задана аффинная геометрия, т.е. задана метрика g лоренцевой сигнатуры и связность Г. Лагранжиан скалярного поля, минимальным образом взаимодействующего с гравитацией, выбирается в виде

$$L = \sqrt{|g|}\left(\frac{1}{2}\partial\varphi^2 - V(\varphi)\right), \tag{23.13}$$

где введено сокращенное обозначение,

$$\partial\varphi^2 := g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi,$$

для кинетической части лагранжиана. Лагранжиан (23.13) зависит только от метрики, а аффинная связность в него не входит. Это означает, что лагранжиан скалярного поля при минимальной подстановке имеет один и тот же вид как в римановой, так и в аффинной геометрии.

Вычислим вариационные производные действия:

$$S,_{\varphi} := \frac{\delta S}{\delta\varphi} = -\sqrt{|g|}\left(\widetilde{\Box}\varphi + V'(\varphi)\right) = 0, \tag{23.14}$$

$$S,_{\alpha\beta} := \frac{\delta S}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\alpha\beta}, \tag{23.15}$$

где

$$\widetilde{\Box} := g^{\alpha\beta}\widetilde{\nabla}_\alpha\widetilde{\nabla}_\beta.$$

– инвариантный волновой оператор, построенный по псевдоримановой метрике $g_{\alpha\beta}$ и

$$T_{\alpha\beta} = \partial_\alpha\varphi\partial_\beta\varphi - g_{\alpha\beta}\left(\frac{1}{2}\partial\varphi^2 - V\right) \tag{23.16}$$

– ковариантное обобщение тензора энергии-импульса (23.5) для пространства Минковского. Уравнение (23.14) является инвариантным уравнением движения для скалярного поля в аффинной геометрии. Однако в него входит не аффинная связность, которая может быть задана на многообразии, а только символы Кристоффеля. Это происходит потому что скалярное поле не чувствует кручения и неметричности при

минимальной подстановке. Взаимодействие скалярного поля с кручением и неметричностью можно ввести неминимальным образом, добавив к лагранжиану соответствующие слагаемые. При этом существует много возможностей, поэтому мы их обсуждать не будем. Вариационная производная (23.16) определяет тензор энергии-импульса скалярного поля, который служит источником для гравитационного поля (компонент метрики) в уравнениях Эйнштейна.

Замечание. То, что вариационная производная действия по метрике (23.16) пропорциональна ковариантному обобщению канонического тензора энергии-импульса не является общим свойством. Дальнейшие примеры покажут, что вариация действия по метрике в общем случае не всегда пропорциональна ковариантному обобщению тензора энергии-импульса. □

Предложение 23.1.1. *Если метрика на многообразии* $\mathbb{M}$ *имеет лоренцеву сигнатуру и координата* x^0 *является временем, то временна́я компонента тензора энергии-импульса действительного скалярного поля,*

$$T_{00} = \partial_0\varphi\partial_0\varphi - g_{00}\left(\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi - V\right), \tag{23.17}$$

при $V \geq 0$ *положительно определена и, следовательно удовлетворяет слабому энергетическому условию (21.129).*

Доказательство. Обозначим $u^\alpha := g^{\alpha\beta}\partial_\beta\varphi$. Тогда

$$\partial_0\varphi = g_{00}u^0 + g_{0\mu}u^\mu.$$

Легко проверить равенство

$$(\partial_0\varphi)^2 - \frac{1}{2}g_{00}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi = \frac{1}{2}(g_{00}u^0 - g_{0\mu}u^\mu)^2 - \frac{1}{2}g_{00}\left(g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}\right)u^\mu u^\nu.$$

Если метрика имеет лоренцеву сигнатуру и координаты выбраны так, что x^0 – это время, то из теоремы 5.2.1 следует, что $g_{00} > 0$ и матрица

$$g_{\mu\nu} - \frac{g_{0\mu}g_{0\nu}}{g_{00}}$$

отрицательно определена. Это означает, что временна́я компонента (23.17) является положительно определенной как сумма положительно определенных слагаемых. □

Если действие инвариантно относительно общих преобразований координат, которые параметризуются n функциями, то согласно второй теореме Нетер уравнения движения удовлетворяют n тождествам. Допустим, что действие зависит только от метрики и скалярного поля. Тогда инвариантность действия означает равенство нулю вариации

$$\delta S = \int dx\sqrt{|g|}(S,^{\alpha\beta}\delta g_{\alpha\beta} + S,_\varphi\delta\varphi) = 0.$$

Отсюда с учетом явного вида вариации компонент метрики (3.102) и скалярного поля (3.97) получаем, что уравнения движения (23.14) и (23.15) удовлетворяют n тождествам:

$$2\widetilde{\nabla}_\alpha S,^\alpha{}_\beta - S,_\varphi\partial_\beta\varphi = 0, \tag{23.18}$$

где $\widetilde{\nabla}_\alpha$ – ковариантная производная с символами Кристоффеля.

С формальной точки зрения сдвиги на постоянный вектор, $x^a l \mapsto x^\alpha + a^a l$, $a^\alpha = \mathsf{const}$, образуют подгруппу группы общих преобразований координат. Поэтому, так же, как и в пространстве Минковского, можно построить полную "энергию" и "импульс" скалярного поля. Эти величины будут сохраняться на уравнениях движения, однако им не всегда можно придать физический смысл, т.к. понятие декартовой системы координат в общем случае отсутствует. Это построение имеет смысл в асимптотически плоском пространстве-времени, когда на больших расстояниях пространство-время становится приближается к пространству Минковского. Это же верно и для момента количества движения.

23.1.3 Скалярное поле в общей теории относительности

Рассмотрим действительное скалярное поле, которое минимальным образом взаимодействует с гравитацией в общей теории относительности. В этом случае действие имеет вид

$$S = \kappa S_{\text{не}} + S_\varphi, \tag{23.19}$$

где $S_{\text{не}}$ – действие Гильберта–Эйнштейна (21.6) и S_φ – действие для скалярного поля с лагранжианом (23.13). Добавление к действию скалярного поля действия Гильберта–Эйнштейна дает кинетический член для метрики. Поэтому вариация этого действия по обратной метрике приводит к уравнениям движения Эйнштейна

$$\frac{1}{\sqrt{|g|}}\frac{\delta S}{\delta g^{\alpha\beta}}: \qquad \kappa\left(\widetilde{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\widetilde{R}\right) + g_{\alpha\beta}\Lambda + \frac{1}{2}T_{\alpha\beta} = 0, \tag{23.20}$$

где $T_{\alpha\beta}$ – тензор энергии-импульса скалярного поля (23.16). Поскольку действие Гильберта–Эйнштейна не зависит от скалярного поля, то уравнение движения для скалярного поля остается прежним (23.14). Таким образом, полная система уравнений движения для скалярного поля в общей теории относительности состоит из уравнений (23.14) и (23.20).

Действие (23.19) инвариантно относительно общих преобразований координат. Поэтому, согласно второй теореме Нетер, между уравнениями движения существует линейная зависимость (23.18).

Перепишем уравнения Эйнштейна в другом виде. След уравнений Эйнштейна (23.20),

$$\kappa\left(1 - \frac{n}{2}\right)\widetilde{R} + n\Lambda + \frac{1}{2}\left(1 - \frac{n}{2}\right)\partial\varphi^2 + \frac{n}{2}V = 0,$$

позволяет исключить скалярную кривизну. В результате уравнения движения (23.20) можно записать в эквивалентной форме

$$\kappa\widetilde{R}_{\alpha\beta} = -\frac{1}{2}\partial_\alpha\varphi\partial_\beta\varphi + \frac{1}{n-2}g_{\alpha\beta}(2\Lambda + V). \tag{23.21}$$

При нулевой космологической постоянной $\Lambda = 0$ и потенциале $V = 0$ уравнения (23.21) существенно упрощаются,

$$\kappa\widetilde{R}_{\alpha\beta} = -\frac{1}{2}\partial_\alpha\varphi\partial_\beta\varphi. \tag{23.22}$$

Эта модель безмассового скалярного поля в общей теории относительности привлекает в последнее время большое внимание из-за своей относительной простоты.

Покажем, что уравнение движения для скалярного поля (23.14) являются следствием уравнений Эйнштейна (23.20).

Предложение 23.1.2. *Если в некоторой области пространства-времени $\mathbb{U} \subset \mathbb{M}$ градиент скалярного поля отличен от нуля, $\partial_\alpha \varphi \neq 0$, то в этой области уравнение для скалярного поля (23.14) является следствием уравнений Эйнштейна (23.20).*

Доказательство. Подействуем оператором ковариантного дифференцирования $\widetilde{\nabla}^\beta$ на уравнение (23.20). В силу свернутых тождеств Бианки (7.108)

$$\widetilde{\nabla}^\beta \left(\widetilde{R}_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} \widetilde{R} \right) = 0$$

получим равенство

$$\widetilde{\nabla}^\beta T_{\alpha\beta} = \partial_\alpha \varphi (\widetilde{\Box} \varphi + V') = 0.$$

Отсюда вытекает сделанное утверждение. □

Доказанное предложение позволяет вместо решения полной системы уравнений для скалярного поля и метрики ограничится решением только уравнений Эйнштейна.

В рассматриваемом случае ситуация аналогична точечным частицам в общей теории относительности, предложение 21.8.2.

До сих пор мы рассматривали минимальное взаимодействие скалярного с метрикой. Представляет также интерес неминимальное взаимодействие, поскольку в этом случае возможно появление дополнительной важной локальной инвариантности.

Теорема 23.1.1. *Действие*

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\frac{1}{2} \partial\varphi^2 - \frac{n-2}{8(n-1)} \varphi^2 \widetilde{R} - \lambda \varphi^{\frac{2n}{n-2}} \right), \qquad \lambda = \mathsf{const}, \tag{23.23}$$

с точностью до граничных слагаемых инвариантно относительно преобразований Вейля:

$$\begin{aligned} g_{\alpha\beta} &\mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \\ \varphi &\mapsto \bar{\varphi} = \mathsf{e}^{-\frac{n-2}{2}\phi} \varphi, \end{aligned} \tag{23.24}$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ – произвольная функция.

Доказательство. Прямая проверка. При этом удобно использовать формулу преобразования скалярной кривизны (**??**) и правило интегрирования по частям (7.63). □

Учитывая размерность скалярного поля (23.2), получаем, что константа связи λ самодействия скалярного поля в действии (23.23) безразмерна.

Коэффициенты перед первыми двумя слагаемыми в действии (23.23) фиксированы требованием инвариантности относительно преобразований Вейля. Их можно умножить только на общую отличную от нуля постоянную. Фиксирован также показатель степени в преобразовании Вейля для скалярного поля (23.24). Нетрудно также проверить, что добавление к действию (23.23) действия Гильберта–Эйнштейна и космологической постоянной нарушает вейлевскую инвариантность.

Уравнения движения для действия (23.23) имеют следующий вид

$$\begin{aligned} \frac{1}{\sqrt{|g|}} S_{,}^{\alpha\beta} : \quad & \frac{n-2}{8(n-1)} \left[\varphi^2 \left(\widetilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \widetilde{R} \right) - g^{\alpha\beta} \Box \varphi^2 + \widetilde{\nabla}^\alpha \widetilde{\nabla}^\beta \varphi^2 \right] + \\ & + \frac{1}{4} g^{\alpha\beta} \partial\varphi^2 - \frac{1}{2} \partial^\alpha \varphi \partial^\beta \varphi - \frac{1}{2} \lambda g^{\alpha\beta} \varphi^{\frac{2n}{n-2}} = 0, \end{aligned} \tag{23.25}$$

$$\frac{1}{\sqrt{|g|}} S_{,\varphi} : \quad - \widetilde{\Box} \varphi - \frac{n-2}{4(n-1)} \widetilde{R} \varphi - \lambda \frac{2n}{n-2} \varphi^{\frac{n+2}{n-2}} = 0. \tag{23.26}$$

Эти уравнения движения не инвариантны, а ковариантны относительно преобразований Вейля, т.к. при преобразовании умножаются на некоторую функцию.

Поскольку действие (23.23) инвариантно относительно преобразований Вейля, то в силу второй теоремы Нетер между уравнениями движения существует линейная зависимость. Чтобы ее найти, выпишем вариации формы для метрики и скалярного поля при вейлевских преобразованиях (23.24):

$$\delta g_{\alpha\beta} = 2\epsilon g_{\alpha\beta}, \qquad \delta\varphi = -\frac{n-2}{2}\epsilon\varphi,$$

где $\epsilon(x)$ – бесконечно малый параметр преобразования. Поскольку действие инвариантно, то выполнено условие

$$\delta S = \int dx \left(\frac{\delta S}{\delta g_{\alpha\beta}} \delta g_{\alpha\beta} + \frac{\delta S}{\delta\varphi}\delta\varphi \right) = 0.$$

Подставив сюда выражения для вариации полей, получим зависимость уравнений движения:

$$2S_{,}{}^{\alpha\beta}\, g_{\alpha\beta} - \frac{n-2}{2}\varphi S_{,\varphi} = 0. \tag{23.27}$$

Нетрудно проверить, что уравнения движения (23.25), (23.26) действительно удовлетворяют этому тождеству.

В важном случае четырехмерного пространства-времени, действие принимает хорошо известную форму

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\frac{1}{2}\partial\varphi^2 - \frac{1}{12}\varphi^2 \widetilde{R} - \lambda\varphi^4 \right). \tag{23.28}$$

Оно инвариантно относительно преобразований Вейля

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \qquad \varphi \mapsto \bar{\varphi} = \mathsf{e}^{-\phi}\varphi. \tag{23.29}$$

Соответствующие уравнения движения,

$$\frac{1}{\sqrt{|g|}} S_{,}{}^{\alpha\beta}: \quad \frac{1}{12}\left[\varphi^2 \left(\widetilde{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta}\widetilde{R} \right) - g^{\alpha\beta}\Box\varphi^2 + \widetilde{\nabla}^\alpha \widetilde{\nabla}^\beta \varphi^2 \right] - \\ - \frac{1}{2}\partial^\alpha\varphi\partial^\beta\varphi + \frac{1}{4} g^{\alpha\beta}\partial\varphi^2 - \frac{1}{2}\lambda g^{\alpha\beta}\varphi^4 = 0, \tag{23.30}$$

$$\frac{1}{\sqrt{|g|}} S_{,\varphi}: \quad -\widetilde{\Box}\varphi - \frac{1}{6}\widetilde{R}\varphi - 4\lambda\varphi^3 = 0, \tag{23.31}$$

ковариантны.

Действие (23.28) является выделенным случаем скалярно-тензорных моделей гравитации, рассмотренных в разделе 21.6. Чтобы привести лагранжиан (21.32) к рассматриваемому виду, необходимо положить

$$\phi = -\frac{\varphi^2}{12}, \qquad \omega(\phi) = -\frac{3}{2}, \qquad V(\phi) = \lambda\varphi^4, \qquad \kappa = 1.$$

Построение действия для других полей, которое было бы инвариантно относительно преобразований Вейля, представляет собой отдельную задачу, на которой мы останавливаться не будем.

23.2 Комплексное скалярное поле

23.2.1 Комплексное скалярное поле в пространстве Минковского

Пусть $\varphi(x) = \varphi_1(x) + i\varphi_2(x)$ – комплексное скалярное поле (функция) в пространстве Минковского $\mathbb{R}^{1,n-1}$, где $\varphi_{1,2}(x) \in \mathcal{C}^2(\mathbb{R}^{1,n-1})$ – два действительных скалярных поля (действительная и мнимая части). Комплексно сопряженное поле будем обозначать крестом $\varphi^\dagger := \varphi_1 - i\varphi_2$, как это принято в квантовой теории поля. Лагранжиан свободного поля

$$\begin{aligned} L &= \eta^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - m^2\varphi^\dagger\varphi = \\ &= \eta^{\alpha\beta}(\partial_\alpha\varphi_1\partial_\beta\varphi_1 + \partial_\alpha\varphi_2\partial_\beta\varphi_2) - m^2(\varphi_1^2 + \varphi_2^2) \end{aligned} \tag{23.32}$$

с точностью до множителя $1/2$ является суммой лагранжианов (23.1) для действительной и мнимой части. Он является действительной функцией.

Размерность комплексного скалярного поля такая же как и действительного (23.2).

Комплексное скалярное поле задается двумя действительными скалярными полями φ_1 и φ_2. Для получения уравнений движения действие можно варьировать либо по действительной и мнимой частям φ_1 и φ_2, либо по полям φ и $\varphi^\dagger$, рассматривая их в качестве независимых переменных. Мы остановимся на второй возможности как более удобной. Тогда вариация действия по $\varphi^\dagger$ приводит к волновому уравнению для φ (уравнению Клейна–Гордона–Фока)

$$\frac{\delta S}{\delta\varphi^\dagger} = -(\Box + m^2)\varphi = 0, \tag{23.33}$$

которое эквивалентно двум волновым уравнениям для действительной и мнимой части. Вариация действия по φ дает уравнение, сопряженное к (23.33).

При рассмотрении задачи Коши для волнового уравнения (23.33) необходимо задать начальные условия независимо для полей φ_1 и φ_2. Поэтому комплексное скалярное поле описывает две степени свободы и в квантовой теории поля описывает заряженные скалярные частицы.

Для описания самодействия комплексного скалярного поля вводится потенциал $V(\varphi^\dagger, \varphi)$, который в общем случае является достаточно гладкой вещественнозначной функцией двух аргументов $\varphi^\dagger$ и φ. Мы предположим, что потенциал зависит только от одного аргумента – произведения $\varphi^\dagger\varphi$ – и обозначим $V' := dV/d(\varphi^\dagger\varphi)$. Тогда лагранжиан примет вид

$$L = \eta^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - V(\varphi^\dagger\varphi), \tag{23.34}$$

где, как и в случае действительного скалярного поля, мы включили массовый член в определение потенциала V. Уравнение движения для скалярного поля с самодействием становится нелинейным:

$$\frac{\delta S}{\delta\varphi^\dagger} = -(\Box + V')\varphi = 0, \tag{23.35}$$

Как и в случае действительного скалярного поля, действие для комплексного скалярного поля инвариантно относительно группы Пуанкаре $\mathbb{IO}(1, n-1)$. При этом трансляциям соответствует сохраняющийся канонический тензор энергии-импульса

$$T^{(c)}_{\alpha\beta} = \partial_\alpha\varphi^\dagger\partial_\beta\varphi + \partial_\beta\varphi^\dagger\partial_\alpha\varphi - \eta_{\alpha\beta}L. \tag{23.36}$$

Его компоненты, очевидно, вещественны. Спиновый момент комплексного скалярного поля равен нулю и момент количества движения равен орбитальному моменту (23.6).

Действие для комплексного скалярного поля инвариантно также относительно глобальных $\mathbb{U}(1)$ преобразований, меняющих фазу скалярного поля сразу во всем пространстве

$$\varphi' = e^{i\epsilon}\varphi, \quad \varphi'^* = e^{-i\epsilon}\varphi^\dagger, \qquad \epsilon = \mathsf{const}. \tag{23.37}$$

Согласно первой теореме Нетер, этой инвариантности соответствует сохраняющийся ток (19.37)

$$\partial_\alpha J^\alpha = 0.$$

где

$$J_\alpha = i(\varphi^\dagger \partial_\alpha \varphi - \partial_\alpha \varphi^\dagger \varphi). \tag{23.38}$$

Позже мы увидим, что $\mathbb{U}(1)$ инвариантность модели комплексного скалярного поля соответствует закону сохранения заряда.

Рассмотрим гамильтонову формулировку теории. Из лагранжиана (23.34) следуют выражения для канонических импульсов, сопряженных полям φ и $\varphi^\dagger$,

$$p := \frac{\partial L}{\partial(\partial_0 \varphi)} = \partial_0 \varphi^\dagger, \tag{23.39}$$

$$p^\dagger := \frac{\partial L}{\partial(\partial_0 \varphi^\dagger)} = \partial_0 \varphi. \tag{23.40}$$

Ненулевые одновременны́е скобки Пуассона для координат и импульсов имеют вид

$$[\varphi(t,\boldsymbol{x}), p(t,\boldsymbol{x}')] = [\varphi^\dagger(t,\boldsymbol{x}), p^\dagger(t,\boldsymbol{x}')] = \delta(\boldsymbol{x}-\boldsymbol{x}').$$

Для положительно определенного потенциала V плотность гамильтониана комплексного скалярного поля,

$$H = p^\dagger p - \eta^{\mu\nu}\partial_\mu \varphi^\dagger \partial_\nu \varphi + V(\varphi^\dagger\varphi), \tag{23.41}$$

положительно определена. Канонические уравнения движения принимают вид

$$\begin{aligned} \dot\varphi &= p^\dagger, \\ \dot\varphi^\dagger &= p, \\ \dot p &= \triangle\varphi^\dagger - V'\varphi^\dagger, \\ \dot p^\dagger &= \triangle\varphi - V'\varphi. \end{aligned} \tag{23.42}$$

Теперь рассмотрим обобщение модели (23.7) для комплексного скалярного поля. Потенциал самодействия выберем в виде

$$V(\varphi^\dagger\varphi) = \frac{1}{2}\lambda\Big[(\varphi^\dagger\varphi)^2 - a^2\Big]^2 = \frac{1}{2}\lambda\Big[(\varphi^\dagger\varphi)^2 - 2\varphi^\dagger\varphi a^2 + a^4\Big], \quad a>0. \tag{23.43}$$

При $\lambda > 0$ этот потенциал положительно определен. Однородное и статичное решение φ_0 для вакуума должно удовлетворять уравнению

$$\varphi_0(\varphi_0^\dagger\varphi_0 - a^2) = 0.$$

Это уравнение имеет бесконечно много решений:

$$\varphi_0 = 0, \quad \varphi_0 = a\,\mathsf{e}^{i\alpha}, \quad \alpha = \mathsf{const},$$

которые зависят от параметра α. Это связано с наличием унитарной $\mathbb{U}(1)$ глобальной симметрии (23.37). Решение $\varphi_0 = 0$ соответствует неустойчивому положению равновесия. Поэтому выберем вакуумное решение в виде $\varphi_0 = a$ и разложим поля вблизи данного решения:

$$\varphi_1 = a + \tilde{\varphi}_1, \qquad \varphi_2 = \varphi_2. \tag{23.44}$$

В новых переменных потенциал примет вид

$$V = \frac{1}{2}\lambda[4a^2\tilde{\varphi}_1^2 + 4a\tilde{\varphi}_1^3 + 4a\tilde{\varphi}_1\varphi_2^2 + \tilde{\varphi}_1^4 + \varphi_2^4 + 2\tilde{\varphi}_1^2\varphi_2^2].$$

Его квадратичная часть равна

$$V \simeq 2\lambda a^2\tilde{\varphi}_1^2.$$

Таким образом, вещественная часть скалярного поля имеет положительную массу, в то время как мнимая часть вблизи вакуумного решения является безмассовой. Это ясно из рис.23.1,*b*. Точка $\varphi = a$ соответствует минимуму потенциальной энергии, однако в направлении φ_2 этот минимум является безразличным.

Описанная выше ситуация возникновения безмассовых полей в моделях с симметриями является типичной. Допустим, что модель зависит от некоторого набора скалярных полей $\varphi = \{\varphi^a\}$, $a = 1, \ldots, \mathrm{N}$, которые можно рассматривать, как компоненты некоторого вектора. Предположим также, что потенциал взаимодействия инвариантен относительно некоторой группы Ли преобразований $\mathbb{G}$. Если группа Ли K-мерна и вблизи единицы параметризуется набором параметров ϵ^{A}, $\mathrm{A} = 1, \ldots, \mathrm{K}$, то инфинитезимальные преобразования полей имеют вид

$$\varphi(x) \mapsto \varphi(x) + \delta\varphi(x) = \varphi(x) + \epsilon^{\mathrm{A}} R_{\mathrm{A}}{}^a(\varphi),$$

где $R_{\mathrm{A}}{}^a(\varphi)$ – некоторые функции, возможно, нелинейные, полей φ^a. Из инвариантности потенциала вытекает равенство нулю его вариации:

$$\delta V(\varphi) = \epsilon^{\mathrm{A}} R_{\mathrm{A}}{}^a \frac{\partial V}{\partial \varphi^a} = 0. \tag{23.45}$$

Это условие должно выполняться при всех значениях φ. Допустим, что существует вакуумное решение $\varphi_0 = \mathsf{const}$, которое соответствует минимуму потенциала $V(\varphi)$. Тогда выполнено уравнение

$$\left.\frac{\partial V}{\partial \varphi^a}\right|_{\varphi_0} = 0.$$

Само вакуумное решение может нарушать симметрию полностью или частично. Если все векторы $R_{\mathrm{A}}(\varphi_0) \neq 0$ линейно независимы, то симметрия полностью нарушена. Если существуют некоторые наборы параметров ϵ^{A}, для которых $\epsilon^{\mathrm{A}} R_{\mathrm{A}}(\varphi_0) = 0$, то при соответствующих преобразованиях вакуум не меняется, и соответствующая симметрия является симметрией вакуумного решения. Очевидно, что множество преобразований, сохраняющих вакуум, образуют некоторую подгруппу $\mathbb{H} \subset \mathbb{G}$. Описанное явление называется *спонтанным нарушением симметрии*, и было описано в [**?**]. Слово "спонтанный" в данном случае означает, что симметрия нарушена не руками, а выбором вакуумного решения уравнений движения.

Теорема 23.2.1 (Голдстоун). *Если потенциал самодействия скалярных полей инвариантен относительно некоторой группы Ли $\mathbb{G}$ преобразований (23.45) и вакуумное решение нарушает симметрию до некоторой собственной подгруппы $\mathbb{H} \subset \mathbb{G}$, то соответствующая модель содержит безмассовые скалярные поля, соответствующие нарушенным симметриям. Число безмассовых полей больше или равно разности* $\mathsf{dim}\,\mathbb{G} - \mathsf{dim}\,\mathbb{H}$.

Доказательство. Продифференцируем равенство (23.45) по φ^b:

$$\epsilon^{\mathrm{A}}\left(\frac{\partial R_{\mathrm{A}}{}^a}{\partial\varphi^b}\frac{\partial V}{\partial\varphi^a}+R_{\mathrm{A}}{}^a\frac{\partial^2 V}{\partial\varphi^a\partial\varphi^b}\right)=0.$$

Если φ_0 – вакуумное решение уравнений движения, то первое слагаемое обращается в нуль и, следовательно, должно быть выполнено равенство

$$\epsilon^{\mathrm{A}}R_{\mathrm{A}}{}^a\frac{\partial^2 V}{\partial\varphi^a\partial\varphi^b}\bigg|_{\varphi_0}=0. \tag{23.46}$$

Массы полей определяются квадратичным приближением потенциала вблизи вакуумного решения. Поэтому массовая матрица имеет вид

$$m^2_{ab}:=\frac{\partial^2 V}{\partial\varphi^a\partial\varphi^b}\bigg|_{\varphi_0}.$$

Собственные значения этой матрицы не могут быть отрицательны, т.к. φ_0 соответствует минимуму потенциала. Она симметрична и может быть приведена к диагональному виду с помощью ортогональной матрицы (теорема 28.1.9). Система линейных уравнений (23.46) для $\epsilon^{\mathrm{A}}R_{\mathrm{A}}$ имеет нетривиальные решения тогда и только тогда, когда массовая матрица вырождена. Каждой нарушенной симметрии соответствует некоторый нетривиальный вектор $\epsilon^{\mathrm{A}}R_{\mathrm{A}}$. Таких линейно независимых векторов ровно $\dim\mathbb{G}-\dim\mathbb{H}$. Это означает, что массовая матрица должна иметь по крайней мере $\dim\mathbb{G}-\dim\mathbb{H}$ нулевых собственных значений. В общем случае возможно появление дополнительных нулевых собственных значений, вызванных не нарушением симметрии, а спецификой модели. Нулевые собственные числа массовой матрицы означают наличие безмассовых скалярных полей. □

Замечание. В доказанной теореме возможная зависимость параметров преобразования ϵ^{A} от точки пространства-времени никак не использовалась. Поэтому теорема Голдстоуна справедлива как для глобальных, так и для локальных (калибровочных) преобразований. □

Пример 23.2.1. Рассмотрим комплексное скалярное поле с положительно определенным потенциалом

$$V=\frac{1}{4}\nu\Big(\varphi^\dagger\varphi-a^2\Big)^4,\qquad \nu,a>0.$$

Вакуумное решение должно удовлетворять уравнению

$$\varphi_0(\varphi_0^\dagger\varphi_0-a^2)^3=0.$$

Как и для потенциала (23.43) минимуму энергии соответствует бесконечно много решений. Выберем вакуумное решение в виде $\varphi_0=a$ и произведем вблизи него разложение (23.44). Тогда потенциал примет вид

$$V=\frac{1}{4}\nu\Big(2a\tilde\varphi_1+\tilde\varphi_1^2+\varphi_2^2\Big)^4.$$

Этот потенциал вообще не имеет квадратичных слагаемых, и, значит, оба поля $\tilde\varphi_1$ и φ_2 являются безмассовыми. В то же время, вакуумное решение $\varphi_0=a$ нарушает $\mathbb{U}(1)$-симметрию. Таким образом, число безмассовых полей превышает разность $\dim\mathbb{G}-\dim\mathbb{H}=1$. □

В рассмотренном примере появление дополнительного безмассового поля связано не с нарушением симметрии, а со спецификой модели. В квантовой теории поля безмассовые частицы, соответствующие безмассовым скалярным полям в спонтанно нарушенных моделях теории поля, которые связаны именно с нарушением симметрии, называются *голдстоуновскими бозонами*. Их число равно разности $\dim \mathbb{G} - \dim \mathbb{H}$.

23.2.2 Комплексное скалярное поле в аффинной геометрии

В аффинной геометрии минимальная подстановка для комплексного скалярного поля приводит к следующему лагранжиану

$$L = \sqrt{|g|}\Big[g^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - V(\varphi^\dagger\varphi)\Big], \tag{23.47}$$

и аффинная связность в него не входит. Вариация соответствующего действия по полям приводит к следующим уравнениям движения и тензору энергии-импульса

$$S,_{\varphi^\dagger} := \frac{\delta S}{\delta\varphi^\dagger} = -\sqrt{|g|}(\widetilde{\Box} + V')\varphi = 0, \tag{23.48}$$

$$S,_{\varphi} := \frac{\delta S}{\delta\varphi} = -\sqrt{|g|}(\widetilde{\Box} + V')\varphi^\dagger = 0, \tag{23.49}$$

$$\frac{\delta S}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\alpha\beta}, \tag{23.50}$$

где

$$T_{\alpha\beta} = \partial_\alpha\varphi^\dagger\partial_\beta\varphi + \partial_\beta\varphi^\dagger\partial_\alpha\varphi - g_{\alpha\beta}(g^{\gamma\delta}\partial_\gamma\varphi^\dagger\partial_\delta\varphi - V) \tag{23.51}$$

– ковариантное обобщение канонического тензора энергии-импульса (23.36).

Перепишем тензор энергии-импульса через действительную и мнимую части,

$$T_{\alpha\beta} = 2\partial_\alpha\varphi_1\partial_\beta\varphi_1 + 2\partial_\alpha\varphi_2\partial_\beta\varphi_2 - g_{\alpha\beta}\Big(\partial\varphi_1^2 + \partial\varphi_2^2 - V\Big).$$

Слагаемые с производными имеют тот же вид, что и в случае двух действительных скалярных полей. Поэтому из доказательства предложения 23.1.1 вытекает положительная определенность их временны́х компонент. Следовательно, справедливо

Предложение 23.2.1. *Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временна́я компонента тензора энергии-импульса комплексного скалярного поля,*

$$T_{00} = \partial_0\varphi^\dagger\partial_0\varphi + \partial_0\varphi^\dagger\partial_0\varphi - g_{00}(g^{\gamma\delta}\partial_\gamma\varphi^\dagger\partial_\delta\varphi - V), \tag{23.52}$$

при $V \geq 0$ положительно определена и, следовательно, удовлетворяет слабому энергетическому условию.

Если инвариантное действие зависит только от метрики и скалярного поля, то из инвариантности действия относительно общих преобразований координат следует зависимость уравнений движения:

$$2\widetilde{\nabla}_\alpha S,^\alpha{}_\beta - S,_\varphi\partial_\beta\varphi - S,_{\varphi^\dagger}\partial_\beta\varphi^\dagger = 0. \tag{23.53}$$

Эта зависимость уравнений движения будет иметь место, например, в общей теории относительности, когда к действию Гильберта–Эйнштейна добавляется действие для комплексного скалярного поля.

Действие для комплексного скалярного поля инвариантно также относительно $\mathbb{U}(1)$ преобразований в пространстве-мишени (23.37). Эта инвариантность сохраняется и в аффинной геометрии, если компоненты метрики и связности не преобразуются. Таким образом, закон сохранения заряда для комплексного скалярного поля

$$\partial_\alpha J^\alpha = 0.$$

где

$$J^\alpha = i\sqrt{|g|}g^{\alpha\beta}(\varphi^\dagger\partial_\beta\varphi - \varphi\partial_\beta\varphi^\dagger), \tag{23.54}$$

имеет место также на многообразиях с нетривиальной геометрией.

23.2.3 Комплексное скалярное поле в общей теории относительности

В общей теории относительности действие для комплексного скалярного поля при минимальной подстановке имеет вид

$$S = \kappa S_{\text{не}} + S_\varphi,$$

где

$$S_\varphi = \int dx\sqrt{|g|}\big[g^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - V(\varphi^\dagger\varphi)\big].$$

Уравнения Эйнштейна имеют прежний вид (23.20), где тензор энергии-импульса определен равенством (23.51). Их необходимо дополнить двумя уравнениями для скалярного поля (23.48) и (23.49).

Как и в случае действительного скалярного поля действие оператора ковариантного дифференцирования на уравнения Эйнштейна приводит к ковариантному закону сохранения тензора энергии-импульса

$$\widetilde{\nabla}^\beta T_{\alpha\beta} = \partial_\alpha\varphi^\dagger(\widetilde{\Box}\varphi + V'\varphi) + \partial_\alpha\varphi(\widetilde{\Box}\varphi^\dagger + V'\varphi^\dagger) = 0.$$

Ясно, что если выполнены уравнения движения (23.48), (23.49), то тензор энергии-импульса сохраняется. Однако обратное утверждение неверно, и аналога предложения 23.1.2 в комплексном случае нет.

Как и в случае действительного скалярного поля (23.21) из уравнений Эйнштейна можно исключить скалярную кривизну и переписать их в эквивалентном виде

$$\kappa\widetilde{R}_{\alpha\beta} = -\partial_\alpha\varphi^\dagger\partial_\beta\varphi + \frac{2}{n-2}g_{\alpha\beta}(\Lambda + V). \tag{23.55}$$

При неминимальной подстановке также возможно появление инвариантности относительно преобразований Вейля.

Предложение 23.2.2. *Действие*

$$S = \int_{\mathbb{M}} dx\sqrt{|g|}\left(g^{\alpha\beta}\partial_\alpha\varphi^\dagger\partial_\beta\varphi - \frac{n-2}{4(n-1)}\varphi^\dagger\varphi\widetilde{R} - \lambda(\varphi^\dagger\varphi)^{\frac{n}{n-2}}\right), \qquad \lambda = \mathsf{const}, \tag{23.56}$$

с точностью до граничных слагаемых инвариантно относительно преобразований Вейля

$$\begin{aligned} g_{\alpha\beta} &\mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi}g_{\alpha\beta},\\ \varphi &\mapsto \bar{\varphi} = \mathsf{e}^{-\frac{n-2}{2}\phi}\varphi,\\ \varphi^\dagger &\mapsto \bar{\varphi}^\dagger = \mathsf{e}^{-\frac{n-2}{2}\phi}\varphi^\dagger, \end{aligned} \tag{23.57}$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ – произвольная вещественнозначная функция.

Доказательство. Перепишем действие (23.56) через вещественную и мнимую части:

$$S = \int_{\mathbb{M}} dx \sqrt{|g|} \left(\partial\varphi_1^2 + \partial\varphi_2^2 - \frac{n-2}{4(n-1)} (\varphi_1^2 + \varphi_2^2) \widetilde{R} - \lambda (\varphi_1^2 + \varphi_2^2)^{\frac{n}{n-2}} \right).$$

Первые три слагаемых инвариантны как сумма двух инвариантных слагаемых для φ_1 и φ_2. Инвариантность последнего слагаемого просто проверяется. □

23.3 Электромагнитное поле

Электромагнитное поле является важнейшим геометрическим объектом в математической физике. Оно наблюдается в экспериментах и широко используется на протяжении многих лет.

С геометрической точки зрения в природе существует следующая конструкция. Пространство-время, в котором мы живем, представляет собой четырехмерное многообразие $\mathbb{M}$. Строится главное расслоение $\mathbb{P}\big(\mathbb{M}, \pi, \mathbb{U}(1)\big)$ (см. главу 13), базой которого является пространство-время, а структурной группой – мультипликативная группа Ли комплексных чисел $\mathbb{U}(1)$, равных по модулю единице, которая изоморфна группе двумерных вращений $\mathbb{U}(1) \simeq \mathbb{SO}(2)$, и, как многообразие, диффеоморфна окружности, $\mathbb{U}(1) \approx \mathbb{S}^1$. На этом главном расслоении задается связность (см. главу 14). Связность, в свою очередь, определяет форму связности и, после проектирования на базу с какого либо сечения, локальную форму связности:

$$\omega = dx^\alpha A_\alpha.$$

Группа $\mathbb{U}(1)$ одномерна, и поэтому форма ω принимает значения в $\mathbb{R}$. Компоненты локальной формы связности $A_\alpha(x)$, после добавления уравнений Максвелла, называются в физике *потенциалом электромагнитного поля*, а соответствующие компоненты

$$F_{\alpha\beta} := \partial_\alpha A_\beta - \partial_\beta A_\alpha, \tag{23.58}$$

локальной формы кривизны,

$$F = \frac{1}{2} dx^\alpha \wedge dx^\beta F_{\alpha\beta}, \tag{23.59}$$

– называются *напряженностью электромагнитного поля* или просто *электромагнитным полем.*

Как и любая другая форма кривизны, напряженность электромагнитного поля удовлетворяет тождествам Бианки (14.44), которые в рассматриваемом случае имеют вид

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0. \tag{23.60}$$

Обозначим два локальных сечения через

$$\mathsf{e}^{ia(x)},\ \mathsf{e}^{ia'(x)} \in \mathbb{U}(1), \qquad a, a' \in [0, 2\pi].$$

Тогда существует некоторая функция $\phi(x)$, связывающая эти сечения:

$$\mathsf{e}^{ia'(x)} = \mathsf{e}^{i\phi(x)}\, \mathsf{e}^{ia(x)}.$$

При переходе от одного локального сечения к другому компоненты локальной формы связности меняются по-правилу (14.15), которое в данном случае принимает простой вид

$$A'_\alpha := A_\alpha + \partial_\alpha \phi. \tag{23.61}$$

Нетрудно видеть, что при калибровочном преобразовании (23.61) компоненты напряженности электромагнитного поля (23.58) не меняются.

Подчеркнем, что преобразование компонент локальной формы связности возникает в дифференциальной геометрии до того, как вводится понятие ковариантной производной (см. раздел 14.1).

Для построения полевых моделей математической физики одного электромагнитного поля недостаточно и поэтому выбирается некоторый набор дополнительных полей $\psi := \{\psi^{\mathrm{A}}(x)\}$, $\mathrm{A} = 1, \ldots, \mathrm{N}$. Если поля вещественнозначные, то группа $\mathbb{U}(1)$ действует на них тривиально, и они описывают в квантовой теории поля нейтральные частицы. Поэтому рассмотрим комплекснозначные поля. Мы предполагаем, что эти поля являются локальными сечениями некоторого векторного расслоения $\mathbb{E}\Big(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{U}(1), \mathbb{P}\Big)$, которое ассоциировано с главным расслоением $\mathbb{P}\Big(\mathbb{M}, \pi, \mathbb{U}(1)\Big)$. Оно имеет ту же базу $\mathbb{M}$ и структурную группу $\mathbb{U}(1)$. Типичным слоем ассоциированного расслоения является комплексное векторное пространство $\mathbb{V}$, комплексной размерности $\mathsf{dim}_{\,\mathbb{C}}\mathbb{V} = \mathrm{N}$, в котором принимает значения поле $\psi \in \mathbb{V}$. Структурная группа $\mathbb{U}(1)$ действует в $\mathbb{V}$ в соответствии с некоторым представлением. Все неприводимые комплексные представления группы $\mathbb{U}(1)$ хорошо известны (см., например, [**?**]). Они одномерны и параметризуются произвольным целым числом k:

$$\mathbb{U}(1) \ni \mathsf{e}^{i\phi} \mapsto \mathsf{e}^{ik\phi} \in \mathbb{C}, \qquad k \in \mathbb{Z}. \tag{23.62}$$

Число k обязано быть целым, поскольку для группы $\mathbb{U}(1)$ мы отождествляем точки $\phi = 0$ и $\phi = 2\pi$.

Далее, мы предполагаем, что каждая компонента векторного поля ψ^{A} при действии структурной группы преобразуется по некоторому неприводимому представлению:

$$\mathbb{U}(1) \ni \mathsf{e}^{i\phi} : \qquad \psi^{\mathrm{A}} \mapsto \mathsf{e}^{ik_{\mathrm{A}}\phi}\psi^{\mathrm{A}}, \qquad k_{\mathrm{A}} \in \mathbb{Z}. \tag{23.63}$$

При переходе от одного локального сечения к другому, функция $\phi = \phi(x)$ является некоторой достаточно гладкой функцией на пространстве-времени, $x \in \mathbb{M}$, которая принимает значения в интервале $[0, 2\pi]$, концы которого отождествлены.

Преобразование (23.61), (23.63) является калибровочным, т.к. параметр преобразования $\phi(x)$ зависит от точки пространства-времени. Это преобразование называется также *градиентным*.

Связность в главном расслоении $\mathbb{P}$ определяет связности во всех ассоциированных расслоениях $\mathbb{E}$ и, следовательно, понятие ковариантной производной (14.24) для локальных сечений. Для определения ковариантной производных, необходимо знать представление генератора группы $\mathbb{U}(1)$ в $\mathbb{C}$. Разлагая "матрицу" представления (23.63) вблизи единицы, получим

$$\mathsf{e}^{ik_{\mathrm{A}}\phi} = 1 + ik_{\mathrm{A}}\phi + \mathsf{o}(\phi).$$

Отсюда вытекает, что ik_{A} является представлением генератора группы $\mathbb{U}(1)$, и, следовательно, ковариантная производная (14.24) каждой компоненты поля ψ^{A} имеет вид

$$\nabla_\alpha\psi^{\mathrm{A}} := (\partial_\alpha - ik_{\mathrm{A}}A_\alpha)\psi^{\mathrm{A}}. \tag{23.64}$$

Замечательным свойством ковариантной производной является, как легко проверить, ее тензорный характер при калибровочных преобразованиях:

$$\nabla_\alpha\psi^{\mathrm{A}} \mapsto \nabla'_\alpha\psi'^{\mathrm{A}} = \mathsf{e}^{ik_{\mathrm{A}}\phi(x)}\nabla_\alpha\psi^{\mathrm{A}},$$

где

$$\nabla'_\alpha \psi'^{\mathrm{A}} := (\partial_\alpha - ik_{\mathrm{A}} A'_\alpha)\psi'^{\mathrm{A}}. \tag{23.65}$$

Все приведенные формулы уже встречались нам в главе 14 в общем виде. Калибровочное преобразование (23.63), (23.65) можно интерпретировать либо как переход между двумя сечениями главного и ассоциированного расслоений, которые определяют локальные формы связности и кривизны, либо как вертикальный автоморфизм (пример 13.3.1). В первом случае калибровочное преобразование интерпретируется как пассивное, когда точки расслоения остаются на месте, а меняется только локальное сечение. Во втором случае калибровочное преобразование рассматривается как активное, когда точки расслоения сдвигаются под действием структурной группы.

Замечание. Описанная выше конструкция не зависит от того, задана ли или нет на многообразии $\mathbb{M}$ аффинная геометрия, т.е. метрика и аффинная связность. Ни одна из приведенных выше формул, включая тождества Бианки (23.60), не содержит метрики и аффинной связности. □

При построении моделей математической физики для полей ψ необходимо написать некоторую систему уравнений. Мы предполагаем, что уравнения движения для полей, не взаимодействующих с электромагнитным полем, следуют из принципа наименьшего действия для некоторого инвариантного интеграла

$$S = \int_{\mathbb{M}} dx\, L(\psi, \partial\psi),$$

где лагранжиан L зависит от полей $\psi = \{\psi^{\mathrm{A}}\}$ и их производных $\partial\psi = \{\partial_\alpha \psi^{\mathrm{A}}\}$. Как правило, лагранжиан также зависит от метрики и аффинной связности, которые используется для построения инвариантов. Тогда взаимодействие с электромагнитным полем вводится путем *минимальной подстановки*, которая заключается в замене всех частных производных на ковариантные:

$$\partial_\alpha \psi^{\mathrm{A}} \mapsto \nabla_\alpha \psi^{\mathrm{A}} := (\partial_\alpha - ik_{\mathrm{A}} A_\alpha)\psi^{\mathrm{A}}, \tag{23.66}$$

и добавлению к исходному лагранжиану лагранжиана для свободного электромагнитного поля:

$$L(\psi, \partial\psi) \mapsto -\frac{1}{4e^2} F^{\alpha\beta} F_{\alpha\beta} + L(\psi, \nabla\psi),$$

где подъем индексов напряженности электромагнитного поля осуществляется с помощью пространственно-временно́й метрики и e – универсальная постоянная электромагнитного взаимодействия.

Теперь сделаем переобозначение $A_\alpha \mapsto eA_\alpha$ для того, чтобы уравнения движения свободного электромагнитного поля не содержали константу связи e и можно было бы переходить к пределу $e \to 0$. Тогда лагранжиан примет вид

$$-\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} + L(\psi, \nabla\psi),$$

где изменится выражение для ковариантной производной

$$\nabla_\alpha \psi^{\mathrm{A}} := \partial_\alpha \psi^{\mathrm{A}} - iek_{\mathrm{A}} A_\alpha \psi^{\mathrm{A}}. \tag{23.67}$$

Коэффициент $q_{\mathrm{A}} := ek_{\mathrm{A}}$ описывает "силу" взаимодействия компоненты поля ψ^{A} с электромагнитным полем и называется *электрическим зарядом* данного поля. Для

слабых электромагнитных полей именно величина q_{A} входит в закон Кулона (см., например, [?]).

Мы видим, что при минимальной подстановке заряды всех полей квантуются: они должны быть пропорциональны некоторому универсальному заряду e. Причина этого в том, что все неприводимые представления группы $\mathbb{U}(1)$ параметризуются целыми числами (23.62). И это соответствует действительности, т.к. заряды всех известных в настоящее время элементарных частиц кратны заряду электрона. Поэтому естественно считать, что e – это заряд электрона. Таким образом, дифференциальная геометрия диктует нам *квантование электрического заряда*.

Замечание. В моделях, объединяющих электрослабые и сильные взаимодействия, существенную роль играют кварки, электрический заряд которых пропорционален одной трети заряда электрона. В этом нет ничего страшного. Если кварки будут обнаружены экспериментально, то под e надо понимать $1/3$ заряда электрона. □

Замечание. Электроны в квантовой электродинамике описываются спинорным полем, которое имеет четыре комплексные компоненты. В этом случае для каждой из этих компонент следует положить $k_{\text{A}} = -1$. □

В релятивистских теориях поля пространство-время отождествляется с пространством Минковского $\mathbb{R}^{1,n-1}$ некоторого числа измерений. Поскольку пространство Минковского $\mathbb{R}^{1,n-1}$ как многообразие диффеоморфно евклидову пространству $\mathbb{R}^n$, то согласно теореме 13.1.1 все расслоения тривиальны:

$$\mathbb{P} \simeq \mathbb{R}^{1,n-1} \times \mathbb{U}(1), \qquad \mathbb{E} \simeq \mathbb{R}^{1,n-1} \times \mathbb{V}.$$

Это означает, что топология пространства-времени в релятивистских моделях при наличии электромагнитного поля всегда тривиальна, и все поля определены глобально в соответствующей системе координат, покрывающей все пространство Минковского $\mathbb{R}^{1,n-1}$. Нетривиальной может быть только геометрия, если кривизна $\mathbb{U}(1)$-связности отлична от нуля.

Однако, в моделях математической физики, как правило, компоненты связности ищутся путем решения некоторых уравнений. Если окажется так, что соответствующие решения определены не во всем пространстве Минковского, а лишь на некотором подмногообразии $\mathbb{U} \subset \mathbb{R}^{1,n-1}$, то возникает главное расслоение $\mathbb{P}\big(\mathbb{U}, \pi, \mathbb{U}(1)\big)$ с другой базой $\mathbb{U}$. Если топология $\mathbb{U}$ нетривиальна, то можно говорить о топологических эффектах. Такова, например, ситуация с монополем Дирака.

Точечные частицы в (псевдо-)римановом пространстве движутся вдоль мировых линий $q_{\text{I}}(t) = \{q^\alpha_{\text{I}}(t)\} \in \mathbb{M}$, $\text{I} = 1, 2, \ldots$. Если каждой частице приписать заряд e_{I}, то минимальная подстановка состоит в добавлении к лагранжиану свободной частицы (см. 20.2.5) лагранжиана взаимодействия:

$$-m\sqrt{g_{\alpha\beta}\dot q^\alpha_{\text{I}}\dot q^\beta_{\text{I}}} \mapsto -m\sqrt{g_{\alpha\beta}\dot q^\alpha_{\text{I}}\dot q^\beta_{\text{I}}} + e_{\text{I}}A_\alpha \dot q^\alpha_{\text{I}},$$

где $\dot q_{\text{I}} := dq_{\text{I}}/dt$ – скорость частицы.

Мы считаем, что при калибровочном $\mathbb{U}(1)$ преобразовании (23.61) координаты частиц не меняются. Тогда лагранжиан минимального взаимодействия с частицей сам по себе не инвариантен:

$$e_{\text{I}}A_\alpha \dot q^\alpha_{\text{I}} \mapsto e_{\text{I}}A_\alpha \dot q^\alpha_{\text{I}} + e_{\text{I}}\partial_\alpha\phi \dot q^\alpha_{\text{I}} = e_{\text{I}}A_\alpha \dot q^\alpha_{\text{I}} + e_{\text{I}}\frac{d\phi}{dt}.$$

Однако при интегрировании вдоль мировой линии частицы,

$$e_{\text{I}} \int_{q_1}^{q_2} dt\, \partial_\alpha \phi \dot{q}^\alpha_{\text{I}} = e_{\text{I}} \big[\phi(q_2) - \phi(q_1) \big],$$

результат зависит только от положения начальной и конечной точек q_1, q_2, но не от траектории частицы. Отсюда вытекает инвариантность действия при фиксированных граничных условиях и, следовательно, ковариантность уравнений движения для точечной частицы, минимальным образом взаимодействующей с электромагнитным полем.

Замечание. Если рассматривать только замкнутые траектории с началом и концом в точке q_1, то интеграл

$$\oint dt\, \dot{q}^\alpha A_\alpha, \tag{23.68}$$

определяет элемент группы голономии $\mathbb{U}(1)$-связности в главном расслоении $\mathbb{P}\big(\mathbb{M}, \pi, \mathbb{U}(1)\big)$ в точке $q_1 \in \mathbb{M}$ (см. разделы 14.4 и 14.5). □

В моделях математической физики считается, что в экспериментах можно наблюдать только калибровочно инвариантные объекты. Таковыми являются компоненты напряженности электромагнитного поля, которые, как мы увидим ниже, отождествляются с напряженностью электрического и магнитного полей. При этом сам потенциал электромагнитного поля A_α считается ненаблюдаемым, т.к. зависит от выбора калибровки. Однако, наблюдаемым также является элемент группы голономии (23.68), который инвариантен относительно калибровочных преобразований. Это свойство называется эффектом Аароновa–Бома (см. раздел 15.3).

Для точеных частиц заряд e_{I} может быть произволен, и его квантование не вытекает из описанного выше подхода.

В настоящей главе мы не будем рассматривать взаимодействие точечных частиц с электромагнитным полем.

Теперь перейдем к описанию свободного электромагнитного поля в пространстве Минковского.

23.3.1 Лагранжева формулировка

Пусть задано пространство Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений и декартова система координат x^α, $\alpha = 0, 1, \ldots, n-1$. *Электромагнитное поле* описывается компонентами $A_\alpha(x)$ локальной формы $U(1)$-связности на пространстве Минковского. Оно является ковекторным полем по отношению к преобразованиям координат – в данном случае – по отношению к преобразованиям Лоренца. Временну́ю и пространственные компоненты электромагнитного поля будем обозначать следующим образом:

$$A_\alpha = \{A_0, A_\mu\}, \qquad \mu = 1, \ldots, n-1. \tag{23.69}$$

То есть греческие буквы из середины алфавита будут пробегать только пространственные значения.

Действие для свободного электромагнитного поля, по определению, имеет вид

$$S_{\text{ЕМ}} = \int dx L_{\text{ЕМ}}, \quad L_{\text{ЕМ}} = -\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta}, \tag{23.70}$$

где подъем и опускание индексов производится с помощью метрики Минковского $\eta_{\alpha\beta} = \mathsf{diag}\,(+ - \ldots -)$ и $F_{\alpha\beta}$ – напряженность электромагнитного поля (23.58). Поле

$F_{\alpha\beta} = -F_{\beta\alpha}$ является антисимметричным ковариантным тензором второго ранга типа $(0,2)$ и представляет собой компоненты локальной 2-формы кривизны (23.149) $U(1)$-связности.

Уравнения движения для действия (23.70) линейны:

$$S_{\text{ЕМ}},{}^{\alpha} := \frac{\delta S_{\text{ЕМ}}}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} = 0. \tag{23.71}$$

Поэтому мы говорим, что действие (23.70) описывает свободное электромагнитное поле.

Лагранжиан (23.70), напряженность (23.58) и уравнения движения (23.71) инвариантны относительно калибровочных преобразований (23.61) действующих только на поля и не затрагивающие координаты. Из калибровочной инвариантности, согласно второй теореме Нетер, следует, что между уравнениями движения имеется линейная зависимость:

$$\partial_\alpha S_{\text{ЕМ}},{}^{\alpha} = 0. \tag{23.72}$$

Замечание. Легко видеть, что добавление массового члена $-\frac{1}{2}m^2 A^\alpha A_\alpha$ к лагранжиану электромагнитного поля (23.70) нарушает калибровочную инвариантность электродинамики. □

Пример 23.3.1. Напомним, что в *четырехмерном* пространстве Минковского $\mathbb{R}^{1,3}$ напряженности электрического и магнитного полей связаны с потенциалом $A_\alpha = \{A_0, \boldsymbol{A}\}$ известными соотношениями (см., например, [**?**]):

$$\boldsymbol{E} = \mathsf{grad}\, A_0 - \frac{\partial \boldsymbol{A}}{dt}, \qquad \boldsymbol{H} = \mathsf{rot}\, \boldsymbol{A}. \tag{23.73}$$

Это значит, что временна́я компонента электромагнитного поля A_0 является электростатическим потенциалом, а трехмерный ковектор $\boldsymbol{A}$ – векторным потенциалом магнитного поля. Отсюда вытекает, что все компоненты напряженности электромагнитного поля определяются напряженностями электрического и магнитного полей:

$$F_{\alpha\beta} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & H_3 & -H_2 \\ E_2 & -H_3 & 0 & H_1 \\ E_3 & H_2 & -H_1 & 0 \end{pmatrix} \tag{23.74}$$

Лагранжиан электромагнитного поля (23.70) также можно выразить через напряженности электрического и магнитного полей:

$$L_{\text{ЕМ}} = \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{H}^2). \qquad \square$$

Проанализируем уравнения Эйлера–Лагранжа (23.71) в пространстве Минковского произвольного числа измерений подробнее. Для временно́й и пространственных компонент электромагнитного поля уравнения движения имеют вид

$$-\triangle A_0 - \partial_0 \partial_\mu A^\mu = 0, \tag{23.75}$$

$$\Box A_\mu - \partial_\mu \partial_\nu A^\nu - \partial_\mu \partial_0 A^0 = 0, \tag{23.76}$$

где введены операторы Лапласа и Даламбера:

$$\triangle := -\partial^\mu \partial_\mu = \partial_1^2 + \partial_2^2 + \ldots + \partial_{n-1}^2, \qquad \Box := \partial^\alpha \partial_\alpha = \partial_0^2 - \triangle.$$

Уравнение для временнóй составляющей (23.75) для каждого момента времени представляет собой уравнение Пуассона в евклидовом пространстве $\mathbb{R}^{n-1}$ с правой частью, определяемой пространственными компонентами. Сформулируем две теоремы о существовании и единственности решения этого уравнения во всем пространстве и в ограниченной области, доказательство которых приведено, например, в [**?**].

Теорема 23.3.1. *Если правая часть уравнения Пуассона дифференцируема и убывает на бесконечности, то решение уравнения Пуассона во всем пространстве $\mathbb{R}^{n-1}$, $n \geq 2$, существует и единственно в классе обобщенных функций, обращающихся в нуль на бесконечности. При этом решением уравнения Пуассона является дважды непрерывно дифференцируемая функция во всем пространстве и убывающая на бесконечности.*

Теорема 23.3.2. *Если правая часть уравнения Пуассона дифференцируема в ограниченной области $\mathbb{U} \subset \mathbb{R}^{n-1}$, $n \geq 2$, с кусочно гладкой границей $\partial\mathbb{U}$ и непрерывна в замыкании $\overline{\mathbb{U}}$, то решение внутренней задачи Дирихле для уравнения Пуассона существует, единственно и непрерывно зависит от граничных условий на $\partial\mathbb{U}$.*

Если для внешней задачи Дирихле предположить, что правая часть уравнения Пуассона и решение стремятся к нулю на бесконечности, то справедлива аналогичная теорема для внешней задачи Дирихле. Аналогичные теоремы справедливы также для задачи Неймана, когда на границе задается значение нормальной производной искомой функции. В последнем случае решение определяется с точностью до аддитивной постоянной.

Приведенные теоремы показывают, что при достаточно общих предположениях оператор Лапласа в евклидовом пространстве, а также в ограниченной области с заданными граничными условиями обратим, и решение уравнения Пуассона (23.75) можно записать в виде

$$A_0 = -\triangle^{-1}\partial_0\partial_\mu A^\mu. \tag{23.77}$$

Это означает, что в каждый момент времени временнáя компонента электромагнитного поля однозначно определяется пространственными компонентами.

Теперь рассмотрим уравнение для пространственных компонент (23.76). Пространственные компоненты электромагнитного поля A_μ образуют ковектор в евклидовом пространстве $\mathbb{R}^{n-1}$. Если поле A_μ непрерывно дифференцируемо, то при достаточно общих граничных условиях его можно взаимно однозначно разложить на *поперечную (соленоидальную)* и *продольную (потенциальную)* составляющие:

$$A_\mu = A^{\mathrm{T}}_\mu + A^{\mathrm{L}}_\mu, \tag{23.78}$$

которые удовлетворяют следующим условиям:

$$\partial_\mu A^{\mathrm{T}\mu} = 0, \tag{23.79}$$

$$\partial_\mu A^{\mathrm{L}}_\nu - \partial_\nu A^{\mathrm{L}}_\mu = 0. \tag{23.80}$$

Формула (23.78) однозначно определяет ковекторное поле A_μ при заданных полях A^{T}_μ и A^{L}_μ. Покажем, что пространственная часть электромагнитного поля A_μ также однозначно определяет свою продольную и поперечную составляющую. В силу леммы Пуанкаре (4.3.3) из условия (23.80) следует, что продольная составляющая, по крайней мере локально, является градиентом от некоторой функции $\varphi(x)$, определенной с точностью до несущественной постоянной. Для односвязных многообразий, как в

нашем случае, это утверждение является глобальным. Запишем данное свойство в виде равенства

$$A_\mu = A^{\mathrm{T}}_\mu + \partial_\mu \varphi. \tag{23.81}$$

Дифференцирование данного соотношения с учетом условия (23.79) приводит к тому, что скалярное поле, соответствующее продольной составляющей, является решением уравнения Пуассона:

$$\triangle \varphi = -\partial_\mu A^\mu, \qquad \triangle := -\partial^\mu \partial_\mu, \tag{23.82}$$

которое существует и единственно при достаточно общих граничных условиях (см. теоремы 23.3.1, 23.3.2). После этого поперечная составляющая однозначно определена:

$$A^{\mathrm{T}}_\mu = A_\mu - \partial_\mu \varphi. \tag{23.83}$$

Нетрудно проверить, что ее дивергенция действительно равна нулю. Таким образом, вместо $n-1$ компоненты ковекторного поля A_μ можно рассматривать скалярное поле φ, определяемое уравнением (23.82), и соленоидальное ковекторное поле A^{T}_μ (23.83). При таком разложении число независимых компонент сохранено, т.к. поперечное ковекторное поле имеет $n-2$ независимые компоненты в силу дополнительного условия (23.79).

Замечание. При $n = 2$ поперечная составляющая у электромагнитного поля отсутствует. □

Перепишем уравнения движения (23.75), (23.76) для продольной и поперечной части:

$$-\triangle A_0 + \triangle \dot{\varphi} = 0, \tag{23.84}$$

$$\partial_\mu \Box \varphi + \partial_\mu \triangle \varphi - \partial_\mu \ddot{\varphi} = 0, \tag{23.85}$$

$$\Box A^{\mathrm{T}}_\mu = 0, \tag{23.86}$$

где уравнение (23.76) также расщеплено на продольную и поперечную части. Для полей, убывающих на бесконечности, уравнение (23.84) для временно́й компоненты электромагнитного поля легко решается:

$$A_0 = \dot{\varphi}, \tag{23.87}$$

поскольку оператор Лапласа обратим. Если подставить это решение в уравнение (23.85), то получим тождество $0 = 0$. Поскольку все проделанные операции обратимы, то мы доказали следующее

Предложение 23.3.1. *Для полей, убывающих на бесконечности, лагранжевы уравнения движения для электромагнитного поля (23.71) эквивалентны следующей системе уравнений:*

$$A_0 = \dot{\varphi}, \tag{23.88}$$

$$\Box A^{\mathrm{T}}_\mu = 0, \tag{23.89}$$

дополненной определением поля φ (23.81), (23.82).

Тот факт, что n лагранжевых уравнений оказались эквивалентными $n-1$ уравнению (23.88), (23.89) не является удивительным. Как уже отмечалось, из-за калибровочной симметрии, в силу второй теоремы Нетер, между исходными уравнениями движения имеется одно линейное соотношение (23.72).

Если рассматривать задачу Коши для электромагнитного поля, то мы видим, что она может быть поставлена только для поперечных компонент электромагнитного потенциала. Следовательно, электромагнитное поле описывает только $n-2$ физических (распространяющихся) степени свободы. При этом скалярное поле φ является произвольной функцией (непрерывно дифференцируемой и удовлетворяющей граничным условиям для рассматриваемой физической задачи), т.к. для нее нет ни одного уравнения. Временна́я компонента электромагнитного потенциала выражается через φ. Конечно, можно считать наоборот, что компонента A_0 произвольна, а функция φ находится из уравнения (23.88).

Поперечные степени свободы являются распространяющимися и соответствуют электромагнитным волнам. В двух измерениях электромагнитное поле не описывает ни одной степени свободы и электромагнитное излучение отсутствует. В пространстве Минковского трех и четырех измерений электромагнитное поле имеет соответственно одну и две степени свободы.

То обстоятельство, что мы не получили ни одного уравнения на продольную составляющую является следствием калибровочной инвариантности электродинамики относительно преобразований (23.61). Нетрудно проверить, что напряженность электромагнитного поля при выполнении уравнения движения (23.88) зависит только от поперечных компонент:

$$F_{0\mu} = \partial_0 A^{\mathrm{T}}_\mu, \qquad F_{\mu\nu} = \partial_\mu A^{\mathrm{T}}_\nu - \partial_\nu A^{\mathrm{T}}_\mu. \tag{23.90}$$

Калибровочные преобразования (23.61) для компонент электромагнитного поля выглядят следующим образом:

$$A'_0 = A_0 + \partial_0\phi, \tag{23.91}$$

$$\varphi' = \varphi + \phi, \tag{23.92}$$

$$A'^{\mathrm{T}}_\mu = A^{\mathrm{T}}_\mu. \tag{23.93}$$

Используя произвольную функцию $\phi(x)$ можно зафиксировать одну из компонент электромагнитного поля A_0 или φ, то есть придать им определенные наперед заданные значения, которые, конечно, должны быть согласованы с уравнением (23.87). Это называется *фиксированием калибровки* и упрощает решение уравнений Эйлера–Лагранжа, т.к. устраняет функциональный произвол в общем решении и уменьшает число неизвестных функций. Наложение калибровочных условий необходимо также для построения квантовой теории. Например, можно положить $A_0 = 0$ и $\varphi = 0$. Для полей, убывающих на бесконечности, последнее условие эквивалентно *кулоновской* калибровке $\partial_\mu A^\mu = 0$. Среди калибровок, наиболее часто использующихся в электродинамике, отметим следующие[1]:

$$\begin{aligned} \partial_\mu A^\mu = 0 &\quad \text{– кулоновская калибровка,} \\ \partial_\alpha A^\alpha = 0 &\quad \text{– лоренцева калибровка,} \\ A_0 = 0 &\quad \text{– временна́я калибровка,} \\ A_3 = 0 &\quad \text{– аксиальная калибровка.} \end{aligned} \tag{23.94}$$

Из этих калибровок только кулоновская и аксиальная являются каноническими (см. раздел 20.2.3). Лоренцева калибровка содержит производную по времени:

[1]Лоренцеву калибровку ввел датский физик Ludwig Lorenz в 1867 году, а не голландский физик H. A. Lorentz (на русский язык обе фамилии в настоящее время принято переводить одинаково, хотя ранее фамилия Lorentz переводилась, как Лорентц), в честь которого названы преобразования Лоренца. По этому поводу см. [**?**, **?**]

$\partial_0 A^0 + \partial_\mu A^\mu = 0$. Временна́я калибровка с гамильтоновой точки зрения фиксирует множитель Лагранжа и оставляет произвол в определении пространственных компонент электромагнитного потенциала A_μ.

Мы называем поля A_0 и A^{L}_μ нефизическими только потому, что для них нет задачи Коши. Однако это не означает, что они всегда равны нулю и не наблюдаются. Напротив, электростатический потенциал A_0 очень даже наблюдается, например, при взаимодействии точечных зарядов в законе Кулона.

23.3.2 Законы сохранения

Действие для свободного электромагнитного поля в пространстве Минковского инвариантно относительно глобального действия группы Пуанкаре и локальных калибровочных преобразований (23.61). Из теоремы Нетер отсюда следуют законы сохранения тензора энергии-импульса, момента количества движения и вектора электромагнитного тока. Поскольку калибровочные преобразования локальны, то вторая теорема Нетер приводит также к зависимости уравнений движения (23.72), которая была отмечена раньше.

Явное выражение для тензора энергии-импульса (21.162) в случае электромагнитного поля принимает вид

$$\widehat{T}_{\alpha\beta} := \widehat{T}_\alpha{}^\gamma \eta_{\gamma\beta} = -\partial_\alpha A_\gamma F_\beta{}^\gamma + \frac{1}{4}\eta_{\alpha\beta} F^2. \tag{23.95}$$

Первое слагаемое в этой формуле явно несимметрично относительно перестановки индексов. Воспользуемся произволом (19.39) в определении тензора энергии-импульса и определим новый симметричный тензор энергии-импульса

$$T^{\text{s}}_{\alpha\beta} = -F_{\alpha\gamma} F_\beta{}^\gamma + \frac{1}{4}\eta_{\alpha\beta} F^2 = \widehat{T}_{\alpha\beta} + \partial_\gamma A_\alpha F_\beta{}^\gamma. \tag{23.96}$$

С учетом уравнений движения для электромагнитного поля (23.71) разность между новым и старым выражением принимает вид

$$T^{\text{s}}{}_\alpha{}^\beta - \widehat{T}_\alpha{}^\beta = \partial_\gamma (A_\alpha F^{\beta\gamma}), \tag{23.97}$$

что согласуется с видом добавочного члена (19.39). Важным обстоятельством является то, что симметричный тензор энергии-импульса (23.96) калибровочно инвариантен. Закон сохранения энергии-импульса имеет вид (19.44),

$$\partial_\beta T^{\text{s}}{}_\alpha{}^\beta = \partial_\beta \widehat{T}_\alpha{}^\beta = 0,$$

для любого решения уравнений движения.

Лоренцевы вращения для электромагнитного поля записываются в виде

$$\delta x^\alpha = -x^\beta \omega_\beta{}^\alpha, \qquad \delta A_\alpha = \omega_\alpha{}^\beta A_\beta. \tag{23.98}$$

Поэтому для спинового момента (19.53) получаем следующее выражение

$$S_{\alpha\beta}{}^\gamma = A_\alpha F_\beta{}^\gamma - A_\beta F_\alpha{}^\gamma. \tag{23.99}$$

Полный тензор момента количества движения (19.51) состоит из орбитального и спинового моментов:

$$\widehat{J}_{\alpha\beta}{}^\gamma = x_\beta \widehat{T}_\alpha{}^\gamma - x_\alpha \widehat{T}_\beta{}^\gamma + S_{\alpha\beta}{}^\gamma. \tag{23.100}$$

Используя выражения для симметричного тензора энергии-импульса (23.96) и спинового момента (23.99) его можно переписать в виде

$$\widehat{J}_{\alpha\beta}{}^{\gamma} = M_{\alpha\beta}{}^{\gamma} + \partial_{\delta}(x_{\alpha}A_{\beta}F^{\gamma\delta} - x_{\beta}A_{\alpha}F^{\gamma\delta}), \tag{23.101}$$

где

$$M_{\alpha\beta}{}^{\gamma} := x_{\beta}T^{\mathrm{s}}{}_{\alpha}{}^{\gamma} - x_{\alpha}T^{\mathrm{s}}{}_{\beta}{}^{\gamma}$$

– орбитальный момент, построенный по симметричному тензору энергии-импульса (23.97). Поскольку второе слагаемое в выражении (23.101) имеет вид (19.39), то тензор момента количества движения можно переопределить, положив

$$J_{\alpha\beta}{}^{\gamma} = M_{\alpha\beta}{}^{\gamma} = x_{\beta}T^{\mathrm{s}}{}_{\alpha}{}^{\gamma} - x_{\alpha}T^{\mathrm{s}}{}_{\beta}{}^{\gamma}, \tag{23.102}$$

который также сохраняется

$$\partial_{\gamma}J_{\alpha\beta}{}^{\gamma} = 0.$$

Отметим полное отсутствие спинового момента в этом выражении.

С калибровочной инвариантностью электродинамики связан закон сохранения заряда

$$\partial_{\alpha}J^{\alpha} = 0, \tag{23.103}$$

где

$$J^{\alpha} = \partial_{\beta}F^{\beta\alpha} \tag{23.104}$$

– ток, соответствующий первой теореме Нетер, то есть, когда параметр калибровочных преобразований не зависит от точки пространства-времени. В данном случае закон сохранения заряда (23.103) совпадает с зависимостью уравнений движения, вытекающей из второй теоремы Нетер (23.72). При отсутствии полей материи этот закон тривиален, т.к. ток равен нулю в силу уравнений движения (23.71).

23.3.3 Гамильтонова формулировка

Свободное электромагнитное поле является относительно простым и важным примером калибровочной модели. Ниже мы используем общий гамильтонов анализ для систем со связями, описанный в разделе 20.2.3.

Электромагнитному потенциалу A_{α} соответствуют канонически сопряженные импульсы

$$P^{\alpha} := \frac{\partial L_{\text{ЕМ}}}{\partial(\partial_0 A_{\alpha})} = F^{\alpha 0}. \tag{23.105}$$

По определению, скобки Пуассона между координатами и импульсами в каждый момент времени имеют вид

$$\left[A_{\alpha}, P'^{\beta}\right] := \left[A_{\alpha}(t,\boldsymbol{x}), P^{\beta}(t,\boldsymbol{x}')\right] = \delta_{\alpha}^{\beta}\delta(\boldsymbol{x}-\boldsymbol{x}'), \tag{23.106}$$

где $(n-1)$-мерная δ-функция зависит только от пространственных координат. Ввиду антисимметрии напряженности электромагнитного поля в теории возникает первичная связь

$$G_1 := P^0 \approx 0, \tag{23.107}$$

где волнистый знак равенства $\approx$ обозначает, что равенство должно быть выполнено на поверхности всех связей. Других первичных связей электродинамика без полей материи не содержит. Гамильтониан получается простым вычислением

$$\mathcal{H} = \int d\boldsymbol{x}\,\left(-\frac{1}{2}P^{\mu}P_{\mu} + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - A_0\partial_{\mu}P^{\mu} + \lambda P^0\right), \tag{23.108}$$

где мы добавили первичную связь (23.107) с множителем Лагранжа λ. Гамильтоновы уравнения движения для построенного гамильтониана имеют вид

$$\dot{A}_0 = \lambda, \tag{23.109}$$

$$\dot{A}_\mu = -P_\mu + \partial_\mu A_0, \tag{23.110}$$

$$\dot{P}^0 = \partial_\mu P^\mu, \tag{23.111}$$

$$\dot{P}^\mu = \partial_\nu F^{\nu\mu}, \tag{23.112}$$

которые необходимо дополнить уравнением первичной связи (23.107), возникающей при варьировании соответствующего действия по множителю Лагранжа.

В уравнении (23.109) множитель Лагранжа действительно является произвольной функцией. Чтобы убедиться в этом, необходимо проверить самосогласованность уравнений движения при произвольных λ. Поскольку в теории есть связь, то она должна сохраняться во времени. То есть правая часть уравнения (23.111) должна быть равна нулю для любого решения. Вычислим производную по времени:

$$\partial_0(\partial_\mu P^\mu) = [\partial_\mu P^\mu, \mathcal{H}] = \partial^2_{\mu\nu} F^{\nu\mu} = 0,$$

где мы использовали антисимметрию компонент напряженности. Таким образом, в гамильтоновых уравнениях функция λ действительно произвольна, и уравнения движения не накладывают на нее никаких ограничений.

Предложение 23.3.2. *Для полей, убывающих на бесконечности, гамильтоновы уравнения движения для электромагнитного поля (23.107), (23.109)–(23.112) эквивалентны лагранжевым уравнениям (23.71).*

Доказательство. Пусть выполнены гамильтоновы уравнения. Из уравнения (23.110) находим импульсы

$$P_\mu = -\dot{A}_\mu + \partial_\mu A_0 = F_{\mu 0}.$$

Тогда уравнение (23.112) примет вид лагранжева уравнения

$$\partial_\alpha F^{\alpha\mu} = 0.$$

Поскольку $P^0 = 0$, то уравнение (23.111) примет вид $\partial_\mu P^\mu$ или, с учетом выражения для импульсов

$$\partial_\mu F^{\mu 0} = 0.$$

Таким образом, лагранжевы уравнения являются следствием гамильтоновых. При этом мы никак не использовали уравнение для временно́й компоненты (23.109).

Обратно. Пусть выполнены лагранжевы уравнения. Обозначим $P^\alpha := F^{\alpha 0}$. Тогда выполнены уравнения (23.110) и связь $P^0 = 0$. Лагранжевы уравнения в новых переменных примут вид

$$\partial_\mu P^\mu = 0, \qquad -\dot{P}^\mu + \partial_\nu F^{\nu\mu} = 0,$$

что эквивалентно гамильтоновым уравнениям (23.111), (23.112) с учетом условия $P^0 = 0$. Осталось показать, что уравнение (23.109) является следствием лагранжевых уравнений. Ранее мы доказали, что лагранжевы уравнения эквивалентны системе уравнений (23.88), (23.89). В уравнении (23.88) функция φ, а в уравнении (23.109) функция λ являются совершенно произвольными. Поэтому можно положить $\lambda = \ddot{\varphi}$, и тогда эти уравнения совпадут. □

Замечание. В гамильтониан (23.108) необходимо добавить первичную связь, т.к. в противном случае эквивалентности между гамильтоновыми и лагранжевыми уравнениями не будет. □

Продолжим построение гамильтонова формализма согласно общей схеме анализа систем со связями. Производная по времени от первичной связи (23.107) равна

$$\dot{G}_1 = \Big[P^0, \mathcal{H}\Big] = \partial_\mu P^\mu. \tag{23.113}$$

Это приводит к вторичной связи

$$G_2 := \partial_\mu P^\mu \approx 0. \tag{23.114}$$

Нетрудно проверить, что других связей в теории нет. Связи G_1 и G_2 являются функционально независимыми связями первого рода, т.к. их скобка Пуассона равна нулю:

$$\Big[G_1, G_2'\Big] = 0.$$

К полному гамильтониану теории мы добавляем все связи:

$$\mathcal{H}_\text{т} = \int d\boldsymbol{x}\, \left(-\frac{1}{2}P^\mu P_\mu + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \lambda P^0 + \mu\partial_\mu P^\nu\right), \tag{23.115}$$

где $\lambda(x)$ и $\mu(x)$ – множители Лагранжа. На поверхности связей $G_1 = 0$ и $G_2 = 0$ гамильтонова плотность, очевидно, положительно определена.

В лагранжевом формализме действие инвариантно относительно калибровочных преобразований (**??**), которые параметризуются одной произвольной функцией $\phi(x)$. Согласно второй теореме Нетер это приводит к зависимости уравнений движения (23.72). Покажем, что в гамильтоновом формализме действие инвариантно относительно калибровочных преобразований, параметризующихся не одной, а двумя произвольными функциями, по числу связей первого рода, и уравнения движения удовлетворяют двум тождествам. Удвоение числа локальных симметрий произошло потому что в гамильтоновом формализме калибровочный параметр $\phi(x)$ и его временна́я производная $\dot{\phi}(x)$ рассматриваются, как независимые функции и параметризуют, соответственно, различные калибровочные преобразования.

Выразим полное действие для свободного электромагнитного поля через координаты и импульсы:

$$S_\text{т} = \int dx\, (P^\alpha \dot{A}_\alpha + \frac{1}{2}P^\mu P_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \lambda P^0 - \mu\partial_\mu P^\nu). \tag{23.116}$$

Полное действие совпадает с исходным действием (23.70) в лагранжевой формулировке только при $\lambda = 0$ и $\mu = A_0$ и поэтому является более общим. Уравнения Эйлера–Лагранжа для действия (23.116) принимают вид

$$\frac{\delta S_\text{т}}{\delta P^0} = \dot{A}_0 - \lambda = 0, \tag{23.117}$$

$$\frac{\delta S_\text{т}}{\delta P^\mu} = \dot{A}_\mu + P_\mu + \partial_\mu\mu = 0, \tag{23.118}$$

$$\frac{\delta S_\text{т}}{\delta A_0} = -\dot{P}^0 = 0, \tag{23.119}$$

$$\frac{\delta S_\text{т}}{\delta A_\mu} = -\dot{P}^\mu - \partial_\nu F^{\mu\nu} = 0, \tag{23.120}$$

$$\frac{\delta S_\text{т}}{\delta \lambda} = -P^0 = 0, \tag{23.121}$$

$$\frac{\delta S_\text{т}}{\delta \mu} = -\partial_\mu P^\mu = 0. \tag{23.122}$$

Согласно теореме 20.2.3 каждой связи первого рода соответствует калибровочное преобразование, относительно которого полное действие инвариантно. Первичной связи (23.107) соответствует генератор калибровочных преобразований

$$T_1 = \int d\boldsymbol{x}\, \epsilon_1 P^0$$

с некоторым малым параметром $\epsilon_1(x)$. При этом преобразуется только временна́я компонента электромагнитного поля A_0 и множитель Лагранжа λ:

$$\delta_1 A_0 = [A_0, T_1] = \epsilon_1, \qquad \delta_1 \lambda = \dot{\epsilon}_1. \tag{23.123}$$

Все остальные поля остаются без изменения. Согласно второй теореме Нетер из инвариантности действия относительно калибровочного преобразования (23.251) следует зависимость уравнений движения

$$\frac{\delta S_{\text{т}}}{\delta A_0} - \partial_t \frac{\delta S_{\text{т}}}{\delta \lambda} = 0.$$

В справедливости этого тождества нетрудно убедиться прямой проверкой.

Вторичной связи первого рода (23.114) также соответствует генератор преобразований

$$T_2 = \int d\boldsymbol{x}\, \epsilon_2 \partial_\mu P^\mu,$$

где $\epsilon_2(x)$ – малый параметр второго калибровочного преобразования. В этом случае преобразуются только пространственные компоненты электромагнитного потенциала A_μ и множитель Лагранжа μ:

$$\delta_2 A_\mu = -\partial_\mu \epsilon_2, \qquad \delta_2 \mu = \dot{\epsilon}_2.$$

Из инвариантности действия следует зависимость уравнений Эйлера–Лагранжа

$$\partial_\mu \frac{\delta S_{\text{т}}}{\delta A_\mu} - \partial_t \frac{\delta S_{\text{т}}}{\delta \mu} = 0,$$

что также легко проверить. Таким образом, число параметров калибровочных преобразований полного действия $S_{\text{т}}$ в гамильтоновом формализме по сравнению с лоренцевой формулировкой удвоилось. Эта ситуация типична для калибровочных моделей, когда в лагранжевом формализме локальные преобразования зависят как от самого параметра, так и от его производной по времени.

Генератором калибровочных преобразований в лагранжевом формализме (**??**) является линейная комбинация связей первого рода:

$$T = \int dx\, (\dot{\epsilon} P^0 - \epsilon \partial_\mu P^\mu).$$

Действительно,

$$\delta A_0 = [A_0, T] = \dot{\epsilon},$$
$$\delta A_\mu = [A_\mu, T] = \partial_\mu \epsilon.$$

Теперь проанализируем гамильтоновы уравнения движения и связи (**??**) без фиксирования какой-либо калибровки. Для определенности и простоты будем рассматривать задачу в евклидовом пространстве $\mathbb{R}^{n-1}$ для гладких функций достаточно

быстро убывающих на бесконечности. По аналогии с разложением электромагнитного поля на поперечную и продольную части (23.81) разложим также и импульсы:

$$P_\mu = -P^\mu = P^{\text{т}}_\mu + P^{\text{л}}_\mu,$$

где

$$\partial_\mu P^{\text{т}\mu} = 0, \qquad P^{\text{л}}_\mu = \partial_\mu \chi.$$

В этом разложении скалярное поле $\chi(x)$ для каждого момента времени является решением уравнения Пуассона

$$\triangle\chi = -\partial_\mu P^\mu = -\partial_\mu P^{\text{л}\mu}.$$

Для широкого класса граничных условий в односвязных областях разложение импульсов на продольную и поперечную части является взаимно однозначным. Тогда связь (23.114) примет вид

$$\partial_\mu P^\mu = \partial_\mu P^{\text{л}\mu} = 0 \quad \Leftrightarrow \quad \triangle\chi = 0.$$

В рассматриваемом случае мы имеем единственное решение $\chi = 0$, что влечет за собой отсутствие продольной составляющей у импульса, $P^{\text{л}\mu} = 0$.

Запишем уравнение движения для электромагнитного потенциала (23.118) отдельно для продольной и поперечной компоненты:

$$\dot{A}^{\text{т}}_\mu = -P^{\text{т}}_\mu, \qquad \dot{A}^{\text{л}}_\mu = -P^{\text{л}}_\mu - \partial_\mu \mu. \tag{23.124}$$

Последнее уравнение можно переписать в терминах потенциала $A^{\text{л}}_\mu = \partial_\mu \varphi$ и один раз проинтегрировать:

$$\dot{\varphi} = -\chi - \mu = \triangle^{-1}\partial_\mu P^\mu - \mu.$$

Отсюда следует, что эволюция продольной $A^{\text{л}}_\mu$ составляющей электромагнитного поля произвольна, так как определяется множителем Лагранжа.

С учетом уравнения движения для временно́й компоненты A_0, мы видим, что эволюция компонент A_0 и $A^{\text{л}}_\mu$ совершенно произвольна, т.к. определяется множителями Лагранжа. При этом уравнения связей позволяют найти сопряженные импульсы $P^0 = 0$ и $P^{\text{л}\mu} = 0$. В то же время множители Лагранжа не входят в уравнения движения для поперечных компонент $A^{\text{т}}_\mu$ и $P^{\text{т}\mu}$. Мы видим, что временна́я и продольные компоненты соответствуют калибровочным степеням свободы и являются нефизическими.

В теории свободного электромагнитного поля связи можно решить и выписать в явном виде эффективное действие для физических степеней свободы (см. раздел 20.2.3). Ввиду важности этой процедуры, остановимся на ней подробно.

Электродинамика без полей материи содержит две связи первого рода

$$G_{\text{а}} = \{G_1, G_2\} = 0, \qquad \text{а} = 1, 2.$$

В соответствии с общим правилом (см. раздел 20.2.3) наложим два калибровочных условия:

$$F^{\text{а}} = \{F^1, F^2\} = \{A_0, \partial_\mu A^\mu\} = 0.$$

Второе калибровочное условие $\partial_\mu A^\mu = 0$ является уравнением только на продольную составляющую электромагнитного поля:

$$\partial_\mu A^\mu = \partial_\mu A^{\text{л}\mu} = 0 \quad \Leftrightarrow \quad \triangle\varphi = 0.$$

При нулевых граничных условиях на бесконечности это уравнение в евклидовом пространстве $\mathbb{R}^{n-1}$ имеет единственное решение $\varphi = 0$, т.е. у пространственных компонент ковектора A_μ отсутствует продольная составляющая. Эта калибровка называется кулоновской.

Обобщенный гамильтониан электродинамики, который получается после добавления всех связей и калибровочных условий, примет вид

$$\mathcal{H}_{\text{E}} = \int d\boldsymbol{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \lambda P^0 + \mu \partial_\mu P^\mu + \nu A_0 + \xi \partial_\mu A^\nu \right),$$

где λ, μ, ν, ξ – множители Лагранжа. Из условия сохранения калибровочных условий и связей во времени получаем уравнения на множители Лагранжа:

$$\lambda = 0, \quad \triangle \mu = 0, \quad \nu = 0, \quad \triangle \xi = 0.$$

При сделанных предположениях уравнения на μ и ξ имеют единственное решение $\mu = 0$, $\xi = 0$. Таким образом, кулоновская калибровка является канонической, поскольку однозначно фиксирует множители Лагранжа.

Для точечных частиц отличие от нуля определителя (20.107) является необходимым и достаточным условием однозначного определения множителей Лагранжа в расширенном гамильтониане S_{E}. Вопрос о критерии однозначного определения множителей Лагранжа в теории поля сложен. Действительно, введем обозначение для полной системы связей и калибровочных условий:

$$\Phi^{\text{M}} = \{G_1, F^1, G_2, F^2\}, \qquad \text{M} = 1, 2, 3, 4.$$

Тогда матрица, составленная из скобок Пуассона, примет вид

$$[\Phi^{\text{M}}, \Phi'^{\text{N}}] = \begin{pmatrix} 0 & -\delta(\boldsymbol{x}' - \boldsymbol{x}) & 0 & 0 \\ \delta(\boldsymbol{x}' - \boldsymbol{x}) & 0 & 0 & 0 \\ 0 & 0 & 0 & \triangle\delta(\boldsymbol{x}' - \boldsymbol{x}) \\ 0 & 0 & -\triangle\delta(\boldsymbol{x}' - \boldsymbol{x}) & 0 \end{pmatrix}.$$

Вопрос о том, что такое определитель этой матрицы сложен, и мы на нем останавливаться не будем.

Фазовое пространство электродинамики $\mathbb{N}$ бесконечномерно, $\mathsf{dim}\,\mathbb{N} = 2n \times \infty$, и задается координатами $A_\alpha(\boldsymbol{x})$ и сопряженными импульсами $P^\alpha(\boldsymbol{x})$. Физическое подпространство (поверхность связей и калибровочных условий $\Phi = 0$) $\mathbb{M} \subset \mathbb{N}$, $\mathsf{dim}\,\mathbb{M} = 2(n-2) \times \infty$ определяется уравнениями $\Phi^{\text{M}} = 0$. Построим на $\mathbb{N}$ такую систему координат, которая фигурирует в теореме 20.2.2. Сначала заметим, что на поверхности связей и калибровочных условий $\Phi^{\text{M}} = 0$ продольные компоненты координат и импульсов отсутствуют:

$$A_\mu|_{\Phi=0} = A^{\text{T}}_\mu, \qquad P^\mu|_{\Phi=0} = P^{\text{T}\mu}.$$

Исходной системой координат на $\mathbb{N}$ являются координаты A_α и сопряженные импульсы P^α с канонической скобкой Пуассона (23.106). Введем вместо пространственных компонент A_μ и P^μ их разложение на поперечную и продольную составляющую:

$$\{A_\mu, P^\mu\} \quad \mapsto \quad \{A^{\text{T}}_\mu, A^{\text{L}}_\mu, P^{\text{T}\mu}, P^{\text{L}\mu}\}.$$

Явные формулы для поперечных и продольных компонент имеют вид:

$$A^{\mathrm{T}}_\mu = A_\mu + \frac{\partial_\mu\partial_\nu}{\triangle}A^\nu, \qquad A^{\mathrm{L}}_\mu = -\frac{\partial_\mu\partial_\nu}{\triangle}A^\nu,$$
$$P^{\mathrm{T}\mu} = P^\mu + \frac{\partial^\mu\partial_\nu}{\triangle}P^\nu, \qquad P^{\mathrm{L}\mu} = -\frac{\partial^\mu\partial_\nu}{\triangle}P^\nu,$$

где $\partial^\mu = -\partial_\mu$. Пуассонова структура в новых координатах уже не будет иметь канонического вида, т.к. на них имеются дополнительные соотношения (23.79), (23.80) и такие же соотношения на импульсы. Простые вычисления показывают, что для пространственных компонент отличны от нуля только две скобки Пуассона:

$$[A^{\mathrm{T}}_\mu, P'^{\mathrm{T}\nu}] = \left(\delta^\nu_\mu + \frac{\partial_\mu\partial^\nu}{\triangle}\right)\delta(\boldsymbol{x}-\boldsymbol{x}'), \qquad [A^{\mathrm{L}}_\mu, P'^{\mathrm{L}\nu}] = -\frac{\partial_\mu\partial^\nu}{\triangle}\delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{23.125}$$

В качестве независимых координат на поверхности связей и калибровочных условий $\mathbb{M}$ выберем первые $n-2$ поперечные компоненты электромагнитного поля

$$q^* = \{A^{\mathrm{T}}_a\}, \qquad a = 1,\ldots,n-2.$$

Тогда последняя компонента A_{n-1} находится из уравнения поперечности (23.79)

$$A^{\mathrm{T}}_{n-1} = -\int_{-\infty}^{x^{n-1}} dx'\partial_a A^{\mathrm{T}a}(x^0,x^1,\ldots,x^{n-2},x') := \frac{1}{\partial_{n-1}}\partial_a A^{\mathrm{T}a},$$

где мы ввели обозначение $1/\partial_{n-1}$ для оператора, обратного к частной производной ∂_{n-1}. При этом мы предполагаем, что все компоненты полей достаточно быстро убывают при $x^{n-1}\to-\infty$ и, следовательно, оператор частной производной ∂_{n-1} обратим. Импульсы, сопряженные к координатам q^*, должны иметь с ними каноническую скобку Пуассона. Учитывая (23.125), нетрудно проверить, что импульсы, сопряженные к q^*, имеют вид

$$p^* = \left\{P^{\mathrm{T}a} - \frac{1}{\partial_{n-1}}\partial^a P^{\mathrm{T},n-1}\right\}, \qquad a = 1,\ldots,n-2.$$

Координаты q^* и сопряженные им импульсы p^* образуют систему координат на физическом подмногообразии $\mathbb{M}\subset\mathbb{N}$ фазового пространства. Заметим, что при $x^{n-1}\to\infty$ компоненты импульсов p^* в общем случае не обязательно убывают, поскольку содержат обратный оператор $1/\partial_{n-1}$. Эта ситуация типична в теории поля, где связи представляют собой дифференциальные уравнения.

Таким образом, мы построили канонически сопряженные координаты на поверхности связей. Это – не единственный выбор. В качестве независимых координат можно было бы выбрать непосредственно компоненты $\{q^*,p^*\} = \{A_a, P^a\}$, $a=1,\ldots,n-2$. В этом случае преобразования координат фазового пространства были бы проще, однако выражение для эффективного гамильтониана для физических степеней свободы – сложнее.

Поверхность связей и калибровочных условий $\Phi^{\mathrm{M}}=0$ можно задать канонически сопряженными координатами и импульсами:

$$Q_{\mathrm{A}} = \{A_0, \partial_\mu A^\mu\}, \qquad P^{\mathrm{A}} = \{P^0, \triangle^{-1}\partial_\mu P^\mu\}, \qquad \mathrm{A}=1,2.$$

Действительно, нетрудно проверить, что скобки Пуассона имеют канонический вид:

$$[Q_{\mathrm{A}}, P'^{\mathrm{B}}] = \delta^{\mathrm{B}}_{\mathrm{A}}\delta, \qquad [Q_{\mathrm{A}}, Q'_{\mathrm{B}}] = 0, \quad [P^{\mathrm{A}}, P'^{\mathrm{B}}] = 0.$$

Эти координаты описывают нефизические (калибровочные) степени свободы.

Таким образом, в соответствии с теоремой 20.2.2 мы построили систему координат на фазовом пространстве $\mathbb{N}$

$$\{q^*, p^*, Q_{\text{A}}, P^{\text{A}}\}$$

такую, что канонически сопряженные координаты и импульсы q^*, p^* являются координатами на физическом подпространстве $\mathbb{M} \subset \mathbb{N}$, а поверхность связей и калибровочных условий задается особенно просто:

$$\Phi^{\text{M}} = 0 \quad \Leftrightarrow \quad Q_{\text{A}} = 0,\ P^{\text{A}} = 0.$$

Построенная система координат нелокальна в том смысле, что содержит обратные дифференциальные операторы. Эта ситуация типична для калибровочных теорий поля, где связи являются дифференциальными уравнениями по пространственным координатам, поскольку их решения задаются интегралами по пространству.

Решения дифференциальных уравнений в частных производных существенно зависят от области, в которой ищется решение, и граничных условий. В частности, оператор Лапласа может быть необратим, и это может приводить к физическим следствиям.

Таким образом, электродинамика в пространстве Минковского $\mathbb{R}^{1,n-1}$ описывает $n-2$ физические (распространяющиеся) степени свободы. В четырехмерном пространстве-времени мы говорим, что фотон имеет две поляризации. При этом временна́я и продольная компоненты электромагнитного поля являются нефизическими (нераспространяющимися). Это не значит, что они равны нулю. Например, в электростатике временна́я компонента A_0 нетривиальна и наблюдаема, т.к. представляет собой электрический кулонов потенциал взаимодействия зарядов. Нефизичность в рассматриваемом контексте означает только то, что для соответствующих полей нельзя поставить задачу Коши.

В двумерном пространстве-времени, $n = 2$, пространственная часть электромагнитного поля A_μ имеет только одну продольную составляющую. Все координаты и импульсы находятся из решения связей и калибровочных условий. Поэтому электромагнитное поле в двух измерениях не имеет физических (распространяющихся) степеней свободы. В этом смысле модель тривиальна. Тем не менее она представляет интерес, особенно при включении взаимодействия с другими полями.

Пример 23.3.2. В четырехмерном пространстве-времени на поверхности связей и калибровочных условий физический (эффективный) гамильтониан принимает простой вид

$$\mathcal{H}_{\text{ph}} = \int d\boldsymbol{x} \left(-\frac{1}{2} P^{\text{T}\mu} P^{\text{T}}_{\mu} + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \frac{1}{2} \int d\boldsymbol{x} (\boldsymbol{E}^{\text{T}2} + \boldsymbol{H}^2),$$

где мы учли выражения для напряженностей электрического и магнитного полей (23.73). Напомним, что из определения напряженности магнитного поля следует, что $\boldsymbol{H} = \boldsymbol{H}^{\text{T}}$. Таким образом, мы получили стандартное выражение для плотности энергии свободного электромагнитного поля, которое, очевидно, положительно определено. □

Гамильтонова формулировка свободной электродинамики на произвольном многообразии $\mathbb{M}$ с метрикой g проводится, по существу, так же как и в пространстве Минковского. Необходимо произвести замену $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}$ и рассматривать компоненты $g_{\alpha\beta}$ как внешнее поле. Почти все формулы останутся при этом без изменения, только в ковариантных производных появятся символы Кристоффеля, возникающие при интегрировании по частям.

23.3.4 Скалярная электродинамика

Рассмотрим простейший пример нетривиального взаимодействия полей материи с электромагнитным полем в пространстве Минковского $\mathbb{R}^{1,n-1}$ произвольного числа измерений. Эта модель – *скалярная электродинамика* – описывает комплексное скалярное поле φ, минимально взаимодействующее с электромагнитным полем A_α. Соответствующий лагранжиан имеет вид

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}\nabla_\alpha\varphi^\dagger\nabla_\beta\varphi - V(\varphi^\dagger\varphi), \tag{23.126}$$

где ковариантные производные определены следующими соотношениями:

$$\begin{aligned}\nabla_\alpha\varphi &= \partial_\alpha\varphi - iqA_\alpha\varphi,\\ \nabla_\alpha\varphi^\dagger &= \partial_\alpha\varphi^\dagger + iqA_\alpha\varphi^\dagger,\end{aligned} \tag{23.127}$$

и $q = \mathsf{const}$ – заряд поля (как мы увидим позже).

Уравнения движения для скалярной электродинамики имеют вид

$$\frac{\delta S}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} + iq\varphi^\dagger\nabla^\alpha\varphi - iq\varphi\nabla^\alpha\varphi^\dagger = 0, \tag{23.128}$$

$$\frac{\delta S}{\delta \varphi^\dagger} = -\nabla^\alpha\nabla_\alpha\varphi - V'\varphi = 0, \tag{23.129}$$

$$\frac{\delta S}{\delta \varphi} = -\nabla^\alpha\nabla_\alpha\varphi^\dagger - V'\varphi^\dagger = 0. \tag{23.130}$$

Сравнение этих уравнений с уравнениями Максвелла в той форме, как они обычно записываются в физической литературе (см., например, [**?**]), показывает, что константа связи q имеет физическую интерпретацию заряда поля.

Лагранжиан (23.126) инвариантен относительно калибровочных преобразований

$$\begin{aligned}\varphi &\mapsto \varphi' = e^{iq\phi}\varphi,\\ \varphi^\dagger &\mapsto \varphi'^\dagger = e^{-iq\phi}\varphi^\dagger,\\ A_\alpha &\mapsto A'_\alpha = A_\alpha + \partial_\alpha\phi,\end{aligned} \tag{23.131}$$

где $\phi(x)$ – локальный параметр преобразования. Согласно второй теореме Нетер отсюда вытекает линейная зависимость уравнений движения:

$$-\partial_\alpha\frac{\delta S}{\delta A_\alpha} + iq\varphi\frac{\delta S}{\delta\varphi} - iq\varphi^\dagger\frac{\delta S}{\delta\varphi^\dagger} = 0. \tag{23.132}$$

В случае свободного электромагнитного поля зависимость уравнений движения (23.72) выглядит тривиально. Для скалярной электродинамики доказать зависимость уравнений (23.128)–(23.130) без обращения к калибровочной инвариантности намного сложнее.

В пространстве Минковского лагранжиан инвариантен относительно действия группы Пуанкаре $\mathbb{ISO}(1,n-1)$. Этой инвариантности, согласно первой теореме Нетер, соответствует закон сохранения энергии импульса и момента количества движения. Канонический тензора энергии-импульса равен сумме тензора энергии-импульса электромагнитного (23.95) и скалярного полей (23.36):

$$\widehat{T}^{(c)}_{\alpha\beta} = -\partial_\alpha A_\gamma F_\beta{}^\gamma + \frac{1}{4}\eta_{\alpha\beta}F^2 + \partial_\alpha\varphi^\dagger\nabla_\beta\varphi + \nabla_\beta\varphi^\dagger\partial_\alpha\varphi - \eta_{\alpha\beta}\Big[\nabla^\gamma\varphi^\dagger\nabla_\gamma\varphi - V(\varphi^\dagger\varphi)\Big], \tag{23.133}$$

который, очевидно, несимметричен. Как и в случае свободного электромагнитного поля, добавление дивергенции (23.97) с учетом уравнений движения позволяет определить симметричный и калибровочно инвариантный тензор энергии-импульса

$$T^{(s)}_{\alpha\beta} = -F_{\alpha\gamma}F_{\beta}{}^{\gamma} + \frac{1}{4}\eta_{\alpha\beta}F^2 + \nabla_\alpha\varphi^\dagger\nabla_\beta\varphi + \nabla_\beta\varphi^\dagger\nabla_\alpha\varphi - \eta_{\alpha\beta}\Big[\nabla^\gamma\varphi^\dagger\nabla_\gamma\varphi - V(\varphi^\dagger\varphi)\Big]. \quad (23.134)$$

Выражение для тензора спина в скалярной электродинамике такое же, как и для свободного электромагнитного поля (23.99). Связь полного момента количества движения с орбитальным моментом, определяемым симметричным тензором энергии-импульса также остается прежней (23.102).

Сохраняющийся вектор электромагнитного тока, соответствующий глобальным преобразованиям (23.131), имеет вид

$$J^\alpha = -\partial_\beta F^{\alpha\beta} + iq(\varphi^\dagger\nabla^\alpha\varphi - \varphi\nabla^\alpha\varphi^\dagger). \quad (23.135)$$

Заметим, что зависимость уравнений движения (23.132), вытекающая из второй теоремы Нетер, и сохранение тока $\partial_\alpha J^\alpha$ – не одно и то же.

Спектр модели, т.е. массы, спины и заряды частиц после вторичного квантования, определяется квадратичным приближением лагранжиана, которое, в свою очередь, зависит от вида потенциала самодействия скалярного поля. Если скалярное поле свободно, т.е.

$$V(\varphi^\dagger\varphi) = m^2\varphi^\dagger\varphi, \qquad m > 0,$$

то модель (23.126) в низкоэнергетическом пределе описывает заряженные скалярные частицы массы m, взаимодействующие с безмассовым электромагнитным полем.

Если потенциал скалярного поля описывает самодействие, то ситуация может сильно измениться. Продемонстрируем это на примере потенциала "$\lambda\varphi^4$" (23.43). Сначала перепишем лагранжиан (23.126) через действительную и мнимую части:

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}\Big(\nabla_\alpha\varphi_1\nabla_\beta\varphi_1 + \nabla_\alpha\varphi_2\nabla_\beta\varphi_2\Big) - V(\varphi_1^2 + \varphi_2^2),$$

где

$$\nabla_\alpha\varphi_1 = \partial_\alpha\varphi_1 + qA_\alpha\varphi_2,$$
$$\nabla_\alpha\varphi_2 = \partial_\alpha\varphi_2 - qA_\alpha\varphi_1.$$

В таком виде скалярные поля φ_1 и φ_2 преобразуются по неприводимому двумерному вещественному представлению унитарной группы $\mathbb{U}(1)$ (векторному представлению группы вращений $\mathbb{SO}(2)$) Вакуумное решение выберем однородным и изотропным:

$$\varphi_0 = a, \qquad A_\alpha = 0$$

и произведем разложение скалярного поля (23.44) вблизи вакуума. В новых переменных лагранжиан (23.126) примет вид

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}(\partial_\alpha\tilde\varphi_1\partial_\beta\tilde\varphi_1 + \partial_\alpha\varphi_2\partial_\beta\varphi_2) +$$
$$+ 2qA^\alpha(\varphi_2\partial_\alpha\tilde\varphi_1 - \tilde\varphi_1\partial_\alpha\varphi_2 - a\partial_\alpha\varphi_2) + q^2A^\alpha A_\alpha(a^2 + 2a\tilde\varphi_1 + \tilde\varphi_1^2 + \varphi_2^2) +$$
$$- \frac{1}{2}\lambda[4a^2\tilde\varphi_1^2 + 4a\tilde\varphi_1^3 + 4a\tilde\varphi_1\varphi_2^2 + \tilde\varphi_1^4 + \varphi_2^4 + 2\tilde\varphi_1^2\varphi_2^2].$$

Проанализируем квадратичное приближение. Мы видим, что вблизи вакуума векторное поле приобрело массовый член $q^2a^2A^\alpha A_\alpha$, и появилось слагаемое с перемешиванием $-2qaA^\alpha\partial_\alpha\varphi_2$. Заметим, что квадратичные слагаемые по A_α и $\partial_\alpha\varphi_2$ можно записать в виде полного квадрата:

$$\eta^{\alpha\beta}\partial_\alpha\varphi_2\partial_\beta\varphi_2 - 2qaA^\alpha\partial_\alpha\varphi_2 + q^2a^2A^\alpha A_\alpha = q^2a^2\eta^{\alpha\beta}\left(A_\alpha - \frac{1}{qa}\partial_\alpha\varphi_2\right)\left(A_\beta - \frac{1}{qa}\partial_\beta\varphi_2\right).$$

Поскольку замена компонент векторного поля

$$A_\alpha \mapsto B_\alpha := A_\alpha - \frac{1}{qa}\partial_\alpha\varphi_2$$

не изменяет кинетического члена, то это преобразование диагонализирует квадратичную часть лагранжиана. В результате возникает лагранжиан с квадратичной частью, которая вообще не содержит голдстоуновское поле φ_2:

$$-\frac{1}{4}F_{\text{в}}^{\alpha\beta}F_{\text{в}\alpha\beta} + q^2a^2B^\alpha B_\alpha + \partial\tilde{\varphi}_1^2 - 2\lambda a^2\tilde{\varphi}_1^2, \tag{23.136}$$

где

$$F_{\text{в}\alpha\beta} := \partial_\alpha B_\beta - \partial_\beta B_\alpha.$$

Теперь можно описать спектр теории. Модель описывает массивное векторное поле, взаимодействующее с нейтральным массивным скалярным полем. Таким образом, безмассовое голдстоуновское поле исчезло, а векторное поле приобрело массу. При этом число степеней свободы не изменилось, т.к. массивное векторное поле имеет на одну степень свободы больше, чем безмассовое. То есть одна степень свободы комплексного скалярного поля перешла в продольную компоненту массивного векторного поля.

Обратим внимание, что массы полей имеют правильные знаки.

Поле φ_2 при описанном преобразовании остается в высших порядках лагранжиана, однако это не влияет на спектр теории.

Появление массивного векторного поля в физическом секторе модели можно пояснить по-другому. Параметризуем комплексное векторное поле его амплитудой и фазой:

$$\varphi := \rho\,\mathsf{e}^{iq\theta}, \qquad \rho = \rho(x), \quad \theta = \theta(x).$$

Поскольку исходная теория калибровочно инвариантна, то физические степени свободы остаются после фиксирования калибровки. Совершим калибровочное преобразование

$$\begin{aligned}\varphi &\mapsto \varphi' = \mathsf{e}^{-iq\theta},\\ A_\alpha &\mapsto A'_\alpha = A_\alpha - \partial_\alpha\theta.\end{aligned}$$

Другими словами, зафиксируем калибровку $\mathsf{im}\,\varphi = 0$. Тогда лагранжиан примет вид (опустив штрихи)

$$\begin{aligned}L &= -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \eta^{\alpha\beta}(\partial_\alpha\rho + iqA_\alpha\rho)(\partial_\beta\rho - iqA_\beta\rho) - V(\rho^2) =\\ &= -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \partial\rho^2 + q^2A^\alpha A_\alpha\rho^2 - V(\rho^2).\end{aligned}$$

Теперь разложим поля вблизи вакуумного решения

$$\rho = a + \tilde{\rho}, \qquad A_\alpha = A_\alpha.$$

Тогда квадратичная часть лагранжиана примет вид

$$-\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \partial\tilde{\rho}^2 + q^2a^2A^\alpha A_\alpha - 2\lambda a^2\tilde{\rho}^2.$$

Это выражение с точностью до обозначений совпадает с полученным ранее выражением (23.136). То есть мы получили прежний спектр модели. При таком подходе голдстоуновское нефизическое поле, каковым является фаза скалярного поля, вообще исключается из лагранжиана.

Описанная выше конструкция обобщается на теорию полей Янга–Миллса, взаимодействующих со скалярными полями. При этом безмассовые голдстоуновские бозоны исчезают, а калибровочные поля приобретают массу. Это явление приобретения масс первоначально безмассовыми векторными полями за счет спонтанного нарушения симметрии называется *механизмом Хиггса*. Оно было независимо открыто Энглертом, Броутом и Хиггсом [**?**, **?**].

Как видим, спектр теории, который определяется низкоэнергетическим приближением, существенно меняется при спонтанном нарушении симметрии. Современные модели, объединяющие электромагнитные, слабые и сильные взаимодействия, основываются на калибровочной симметрии и соответствующих полях Янга–Миллса. Поэтому теория изначально содержит большое число безмассовых векторных полей. С другой стороны, нам экспериментально известно только одно безмассовое векторное поле – электромагнитное. Чтобы устранить возникающее противоречие теории и эксперимента вводится взаимодействие со скалярными полями (хиггсовскими бозонами) таким образом, чтобы вакуумное решение уравнений движения нарушало калибровочную симметрию. В результате калибровочные поля приобретают массу за счет механизма Хиггса, и можно сказать, что мы эти поля не наблюдаем, потому что они слишком массивны.

Важно отметить, что перенормируемость квантовополевой модели определяется не низко-, а высокоэнергетическим приближением, которое у спонтанно нарушенной теории такое же, как и до нарушения симметрии. Поэтому свойство перенормируемости сохраняется при спонтанном нарушении симметрии.

23.3.5 Электромагнитное поле в аффинной геометрии

Лагранжиан электромагнитного поля на многообразии $\mathbb{M}$ с заданной аффинной геометрией при минимальной подстановке принимает вид

$$L_{\text{ем}} = -\frac{1}{4}\sqrt{|g|}g^{\alpha\beta}g^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} := -\frac{1}{4}\sqrt{|g|}F^2, \tag{23.137}$$

где $F_{\alpha\beta}$ – напряженность электромагнитного поля (**??**). Ему соответствует действие

$$S_{\text{ем}} = \int_{\mathbb{M}} dx L_{\text{ем}}. \tag{23.138}$$

Замечание. Как и в случае скалярных полей, лагранжиан (23.137) не зависит от аффинной связности. Это связано с тем, что $F_{\alpha\beta}$ – это компоненты локальной формы кривизны для $U(1)$-связности и являются компонентами тензора независимо от

того задана или нет на многообразии $\mathbb{M}$ аффинная геометрия. С геометрической точки зрения в выражении для $F_{\alpha\beta}$ при минимальной подстановке нет никакой необходимости заменять частные производные на ковариантные, т.к. локальная форма кривизны для $U(1)$-связности уже является тензором относительно общих преобразований координат.

Если все же в определении компонент напряженности электромагнитного поля заменить частные производные на ковариантные, определяемые аффинной связностью Γ, то получим добавочное слагаемое с тензором кручения:

$$\nabla_\alpha A_\beta - \nabla_\beta A_\alpha = \partial_\alpha A_\beta - \partial_\beta A_\alpha - T_{\alpha\beta}{}^\gamma A_\gamma.$$

Полученное выражение ковариантно относительно общих преобразований координат, однако добавочный член явно нарушает калибровочную инвариантность. Такая замена соответствует неминимальной подстановке. □

Замечание. В настоящее время считается, что спектр модели: массы, спины и заряды частиц, описываемых данным полем, определяется линейным приближением в уравнениях движения. При этом метрика, как правило, раскладывается вблизи метрики Минковского. Это означает, что электромагнитное поле в аффинной геометрии описывает безмассовые нейтральные частицы со спином единица. Такой подход вызывает много вопросов. Например, если пространство-время имеет нетривиальную топологию, то говорить про метрику Лоренца не имеет никакого смысла. Мы не будем обсуждать эти важные вопросы, потому что лучших предложений пока не поступало. □

Уравнения движения для электромагнитного поля и тензор энергии-импульса получаются варьированием соответствующего действия

$$\sqrt{|g|}S_{\text{ем}},{}^\alpha := \frac{\delta S_{\text{ем}}}{\delta A_\alpha} = \partial_\beta(\sqrt{|g|}F^{\beta\alpha}) = \sqrt{|g|}\widetilde{\nabla}_\beta F^{\beta\alpha} = 0, \tag{23.139}$$

$$\sqrt{|g|}S_{\text{ем}},{}_{\alpha\beta} := \frac{\delta S_{\text{ем}}}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\text{ем}\alpha\beta} = -\frac{1}{2}\sqrt{|g|}F_{\alpha\gamma}F_\beta{}^\gamma + \frac{1}{8}\sqrt{|g|}g_{\alpha\beta}F^2. \tag{23.140}$$

Отсюда следует выражение для тензора энергии-импульса электромагнитного поля

$$T_{\text{ем}\alpha\beta} = -F_{\alpha\gamma}F_\beta{}^\gamma + \frac{1}{4}g_{\alpha\beta}F^2. \tag{23.141}$$

Замечание. След тензора энергии-импульса электромагнитного поля в четырехмерном пространстве-времени равен нулю, $T_{\text{ем}}{}^\alpha{}_\alpha = 0$. Здесь важна четырехмерность пространства-времени, т.к. если размерность не равна четырем, то след тензора энергии-импульса отличен от нуля. □

Замечание. Вариация действия по метрике (23.140) приводит к тензору энергии-импульса электромагнитного поля (23.141), стоящему в правой части уравнений Эйнштейна (21.1). Он является ковариантным обобщением симметричного канонического тензора энергии-импульса (23.96) в пространстве Минковского. В этом отношении ситуация с электромагнитным полем такая же, как и для скалярного поля. В общем случае вариация действия полей материи по метрике совсем необязательно приводит к ковариантному обобщению канонического тензора энергии-импульса, следующему из теоремы Нетер. □

Если присутствуют источники электромагнитного поля, то в правой части уравнения (23.139) появляется электрический ток:

$$\widetilde{\nabla}_\beta F^{\beta\alpha} = J^\alpha, \tag{23.142}$$

где $J = \{J^\alpha\}$ – вектор тока. Происхождение электрического тока в настоящий момент не имеет значения.

Предложение 23.3.3. *Если пространство-время топологически тривиально (диффеоморфно $\mathbb{R}^n$), то система уравнений второго порядка для потенциала A_α (23.142) эквивалентна системе уравнений первого порядка для компонент напряженности $F_{\alpha\beta} = -F_{\beta\alpha}$ электромагнитного поля*

$$\begin{aligned} \widetilde{\nabla}_\beta F^{\beta\alpha} &= J^\alpha, \\ \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} &= 0 \end{aligned} \tag{23.143}$$

Доказательство. Пусть $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$. Тогда из уравнения (23.142) следует система уравнений (23.143). Обратно. В силу леммы Пуанкаре 4.3.3 из второго уравнения (23.143) для односвязных многообразий следует существование потенциала A_α. Подстановка соответствующего выражения (**??**) в первое уравнение (23.143) приводит к (23.142). □

Определение. Система уравнений первого порядка (23.143) для напряженности электромагнитного поля $F_{\alpha\beta} = -F_{\beta\alpha}$ называется *уравнениями Максвелла.* □

Предложение 23.3.4. *Уравнения движения для действия (23.138) эквивалентны уравнениям Эйлера–Лагранжа для действия*

$$S = -\int_{\mathbb{M}} dx \sqrt{|g|} \left[\frac{1}{2} F^{\alpha\beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) - \frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} \right], \tag{23.144}$$

в котором компоненты $F^{\alpha\beta}$ и A_α рассматриваются в качестве независимых переменных.

Доказательство. Уравнения движения для компонент напряженности являются алгебраическими:

$$\frac{\delta S}{\delta F^{\alpha\beta}} = \frac{1}{2} \big(\partial_\alpha A_\beta - \partial_\beta A_\alpha - F_{\alpha\beta} \big) = 0.$$

Их решение приводит к выражению компонент напряженности электромагнитного поля через потенциал (**??**). Согласно теореме 19.4.1 решение уравнений Эйлера–Лагранжа можно подставлять в действие. После подстановки действие (23.256) совпадет с (23.138). □

Рассмотрение в качестве независимых переменных и компонент потенциала электромагнитного поля A_α, и компонент напряженности $F_{\alpha\beta}$ называется *формализмом первого порядка.*

Действие для лагранжиана (23.137) инвариантно относительно общих преобразований координат. Соответствующие вариации полей имеют вид (см. раздел 3.13)

$$\begin{aligned} \delta A_\alpha &= -\partial_\alpha \epsilon^\beta A_\beta - \epsilon^\beta \partial_\beta A_\alpha, \\ \delta g^{\alpha\beta} &= g^{\alpha\gamma} \partial_\gamma \epsilon^\beta + g^{\beta\gamma} \partial_\gamma \epsilon^\alpha - \epsilon^\gamma \partial_\gamma g^{\alpha\beta}. \end{aligned}$$

Отсюда, согласно второй теореме Нетер (19.63), следует зависимость уравнений движения

$$\partial_\beta S,^\beta A_\alpha + S,^\beta F_{\beta\alpha} - \partial_\beta S,^\beta{}_\alpha - \partial_\beta S,_\alpha{}^\beta - S,_{\beta\gamma}\partial_\alpha g^{\beta\gamma} = 0. \tag{23.145}$$

Или, в ковариантном виде,

$$\widetilde{\nabla}_\beta S,^\beta A_\alpha + S,^\beta F_{\beta\alpha} - 2\widetilde{\nabla}^\beta S,_{\beta\alpha} = 0. \tag{23.146}$$

Если выполнены уравнения движения для электромагнитного поля (23.139), то с учетом свернутых тождеств Бианки соотношение (23.146) принимает вид

$$\widetilde{\nabla}_\beta T_{\text{ЕМ}\alpha}{}^\beta = 0. \tag{23.147}$$

Полученное равенство, выполненное для всех решений уравнений Эйлера–Лагранжа, можно интерпретировать, как ковариантное обобщение закона сохранения тензора энергии-импульса.

Действие для электромагнитного поля инвариантно также относительно калибровочных преобразований (**??**), при которых метрика не меняется. Из калибровочной инвариантности в силу второй теоремы Нетер следует зависимость уравнений движения:

$$\widetilde{\nabla}_\alpha S,^\alpha = 0. \tag{23.148}$$

С учетом уравнения Максвелла (23.142) это равенство приводит к закону сохранения электрического тока: $\widetilde{\nabla}_\alpha J^\alpha = 0$. Если ток возникает при варьировании некоторого калибровочно инвариантного действия для полей заряженной материи, то сохранение тока будет выполняться автоматически в силу второй теоремы Нетер. Если же электрический ток вводится в уравнения Максвелла "руками", то условие $\widetilde{\nabla}_\alpha J^\alpha = 0$ необходимо для самосогласованности уравнений.

23.3.6 Электромагнитное поле в общей теории относительности

В настоящем разделе мы ограничимся четырехмерным пространством-временем $\mathbb{M}$, в котором формулы электродинамики имеют свою специфику.

В свободной электродинамике многие формулы принимают особенно простой вид, если использовать язык дифференциальных форм, описанных в главе 4. Как уже было отмечено, электромагнитный потенциал – это локальная 1-форма $\mathbb{U}(1)$-связности,

$$A := dx^\alpha A_\alpha. \tag{23.149}$$

Слово "локальная" в данном случае означает, что форма связности задана на некоторой координатной окрестности базы $\mathbb{U} \subset \mathbb{M}$ главного расслоения $\mathbb{P}\big(\mathbb{M}, \pi, \mathbb{U}(1)\big)$, а не на самом (пятимерном) пространстве расслоения $\mathbb{P}$. Локальная 2-форма кривизны – это внешняя производная от локальной формы связности,

$$F := dA = \frac{1}{2} dx^\alpha \wedge dx^\beta (\partial_\alpha A_\beta - \partial_\beta A_\alpha). \tag{23.150}$$

В дальнейшем слово "локальная" мы, для краткости, будем опускать.

Поскольку форма кривизны является точной, то ее внешняя производная равна нулю (4.32),

$$dF = \frac{1}{2} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \, \partial_\alpha F_{\beta\gamma} = 0. \tag{23.151}$$

Это равенство, очевидно, эквивалентно тождествам Бианки (23.60).

Приведенные выше формулы справедливы независимо от того задана ли на $\mathbb{M}$ аффинная геометрия или нет.

Если на пространстве-времени задана метрика, то можно определить 2-форму, дуальную к форме кривизны:

$$* F := \frac{1}{2} dx^\alpha \wedge dx^\beta (*F)_{\alpha\beta}, \tag{23.152}$$

где компоненты дуальной формы кривизны заданы равенством

$$* F_{\alpha\beta} := \frac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} F^{\gamma\delta}. \tag{23.153}$$

В приведенной формуле $\varepsilon_{\alpha\beta\gamma\delta}(x)$ – полностью антисимметричный тензор четвертого ранга. В определении дуальной формы кривизны метрика встречается дважды. Во-первых, с ее помощью производится подъем и опускание индексов:

$$F^{\gamma\delta} := g^{\gamma\alpha} g^{\delta\beta} F_{\alpha\beta}. \tag{23.154}$$

Во-вторых, метрика входит в определение полностью антисимметричного тензора:

$$\varepsilon_{\alpha\beta\gamma\delta} := \sqrt{|g|}\hat{\varepsilon}_{\alpha\beta\gamma\delta},$$

где $\hat{\varepsilon}_{\alpha\beta\gamma\delta}$ – полностью антисимметричная тензорная плотность четвертого ранга, чьи компоненты равны по модулю единице (см. раздел 3.8).

Преобразование, обратное к преобразованию дуальности (23.263), имеет вид

$$F_{\alpha\beta} = -\frac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} (*F)^{\gamma\delta}, \tag{23.155}$$

где мы использовали формулы (28.66) для сверток антисимметричных тензоров. Полученное равенство эквивалентно следующему свойству оператора Ходжа:

$$F = * * F.$$

Пример 23.3.3. В четырехмерном пространстве Минковского $\mathbb{R}^{1,3}$ компоненты напряженности выражаются через компоненты электрического и магнитного полей (23.74). В этом случае нетрудно получить выражения для компоненты дуальной напряженности:

$$(*F)_{\alpha\beta} = \begin{pmatrix} 0 & H_1 & H_2 & H_3 \\ -H_1 & 0 & -E_3 & E_2 \\ -H_2 & E_3 & 0 & -E_1 \\ -H_3 & -E_2 & E_1 & 0 \end{pmatrix}. \tag{23.156}$$

Это означает, что преобразование дуальности меняет местами электрическое и магнитное поля:

$$\boldsymbol{E} \leftrightarrow -\boldsymbol{H}. \qquad \square$$

Используя дуальную форму кривизны, действие для свободного электромагнитного поля можно переписать в виде интеграла от 4-формы:

$$S_{\text{ем}} := -\frac{1}{4} \int dx \sqrt{|g|} F^{\alpha\beta} F_{\alpha\beta} = -\frac{1}{2} \int * F \wedge F, \tag{23.157}$$

поскольку, согласно общей формуле для внешнего произведения (4.11),

$$*F \wedge F = \frac{1}{4} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \wedge dx^\delta \; * F_{\alpha\beta} F_{\gamma\delta} = \frac{1}{8} dx^\alpha \wedge dx^\beta \wedge dx^\gamma \wedge dx^\delta \, \varepsilon_{\alpha\beta\epsilon\zeta} F^{\epsilon\zeta} F_{\gamma\delta}.$$

Кроме этого мы воспользовались формулой (4.74) для интегрирования форм и сверткой полностью антисимметричных тензоров четвертого ранга (28.67).

Свернув уравнения Максвелла (23.143) с антисимметричным тензором, получим уравнение Максвелла в терминах дуальной формы кривизны:

$$\begin{gathered} \partial_\alpha (*F)_{\beta\gamma} + \partial_\beta (*F)_{\gamma\alpha} + \partial_\gamma (*F)_{\alpha\beta} = J^\delta \varepsilon_{\delta\alpha\beta\gamma}, \\ \nabla_\beta (*F)^{\beta\gamma} = 0. \end{gathered} \tag{23.158}$$

Поскольку преобразование дуальности меняет местами электрическое и магнитное поле, то второе из этих уравнений означает отсутствие в природе магнитных зарядов.

Первую пару уравнений Максвелла можно записать, используя оператор кограницы (4.55), который в рассматриваемом случае имеет вид

$$\delta := *d* .$$

Несложные вычисления позволяют переписать первую пару уравнений Максвелла (23.158) в виде равенства

$$\delta F = \frac{1}{2} J,$$

где $J := dx^\alpha J_\alpha$ – 1-форма электрического тока.

Остановимся на уравнениях электромагнитного поля без источников. Тогда первую пару уравнений Максвелла (23.158) можно записать в виде

$$d * F = 0. \tag{23.159}$$

Это уравнение можно получить путем варьирования действия (23.157) непосредственно в терминах форм. Поскольку $*F$ – это 2-форма, то из свойства (4.31) следует формула

$$d(*F \wedge A) = d(*F) \wedge A + *F \wedge dA.$$

Если вариации всех полей исчезают на границе, то уравнения движения в форме (23.159) легко получить, варьируя непосредственно действие (23.157):

$$\bar\delta S_{\text{ЕМ}} = -\int *F \wedge \bar\delta (dA) = \int d(*F) \wedge \bar\delta A,$$

где мы пометили вариацию формы чертой, чтобы отличить ее от оператора кограницы.

В четырехмерном пространстве-времени можно построить еще одну 4-форму

$$F \wedge F.$$

Эта форма является точной, потому что из тождеств Бианки следует равенство

$$F \wedge dA = d(F \wedge A) - dF \wedge A = d(F \wedge A).$$

Рассмотрим *уравнение самодуальности*

$$F = *F \qquad \Leftrightarrow \qquad F_{\alpha\beta} = *F_{\alpha\beta}. \tag{23.160}$$

Если в правой части этих уравнений стоит знак минус, то уравнение называется *антисамодуальным*.

Предложение 23.3.5. *Если потенциал электромагнитного поля удовлетворяет уравнению (анти-)самодуальности, то он также удовлетворяет уравнениям Эйлера–Лагранжа (23.139).*

Доказательство. Возьмем внешнюю производную от уравнения (23.264):

$$dF = d * F.$$

Левая часть полученного соотношения тождественно равна нулю в силу тождеств Бианки (23.151). Поэтому возникают уравнения Эйлера–Лагранжа в форме (23.159). □

В отличие от уравнений Эйлера–Лагранжа уравнения (анти-)самодуальности имеют первый порядок и проще.

Теперь изучим свойства тензора энергии-импульса в четырехмерном пространстве-времени. Из определения дуального тензора напряженности (23.262) вытекает равенство

$$*F_{\alpha\gamma} * F_\beta{}^\gamma = -\frac{1}{2} g_{\alpha\beta} F^{\gamma\delta} F_{\gamma\delta} + F_{\alpha\gamma} F_\beta{}^\gamma,$$

где мы воспользовались формулами свертки антисимметричных тензоров (28.67). Теперь выражение для тензора энергии-импульса можно переписать в эквивалентном виде

$$T_{\text{ем}\alpha\beta} = -\frac{1}{2} \left[F_{\alpha\gamma} F_\beta{}^\gamma + *F_{\alpha\gamma} * F_\beta{}^\gamma \right]. \tag{23.161}$$

Из приведенного выражения следует, что тензор энергии-импульса электромагнитного поля инвариантен относительно *дуальных* $\mathbb{SO}(2)$ *вращений*:

$$\begin{aligned} F_{\alpha\beta} &\mapsto F'_{\alpha\beta} = F_{\alpha\beta} \cos\omega - *F_{\alpha\beta} \sin\omega, \\ *F_{\alpha\beta} &\mapsto *F'_{\alpha\beta} = F_{\alpha\beta} \sin\omega + *F_{\alpha\beta} \cos\omega, \end{aligned} \tag{23.162}$$

где $\omega \in [0, 2\pi]$ – угол вращения. Конечно, второе из этих уравнений является следствием первого.

Пример 23.3.4. В пространстве Минковского $\mathbb{R}^{1,3}$ компоненты тензора напряженности и его дуального имеют вид (23.74), (23.156). Поэтому для временно́й компоненты тензора энергии-импульса (23.161) получаем хорошо известное выражение

$$T_{\text{ем}00} = \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{H}^2). \tag{23.163}$$

Смешанная компонента контравариантного тензора энергии-импульса электромагнитного поля имеет вид

$$T^{\mu 0}_{\text{ем}} = F^{\mu\nu} F^0{}_\nu.$$

Используя выражение напряженности электромагнитного поля через электрическое и магнитное поля (23.74) нетрудно получить соотношение

$$T^{\mu 0}_{\text{ем}} = \varepsilon^{\mu\nu\rho} E_\nu H_\rho,$$

где $\varepsilon^{\mu\nu\rho}$ – полностью антисимметричный тензор третьего ранга (28.61). Полученное равенство можно записать в виде векторного произведения

$$\boldsymbol{T} = \boldsymbol{E} \times \boldsymbol{H}. \tag{23.164}$$

Этот вектор называется *вектором Пойнтинга* и имеет смысл плотности потока энергии электромагнитного поля. При этом закон сохранения электромагнитной энергии примет вид

$$\partial_0 T^{00}_{\text{ЕМ}} + \partial_\mu T^{\mu 0}_{\text{ЕМ}} = 0. \qquad \square$$

Предложение 23.3.6. *Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временна́я компонента тензора энергии-импульса электромагнитного поля*

$$T_{\text{ЕМ}00} = -\frac{1}{2}\left[F_{0\gamma}F_0{}^{\gamma} + {*F_{0\gamma}} * F_0{}^{\gamma}\right] \tag{23.165}$$

положительно определена и, следовательно, удовлетворяет слабому энергетическому условию.

Доказательство. Ввиду антисимметрии по индексам первое слагаемое в выражении (23.266) имеет вид

$$-\frac{1}{2}F_{0\mu}F_{0\nu}g^{\mu\nu}.$$

При сделанных предположениях матрица $g^{\mu\nu}$ согласно предложению 22.2.1 отрицательно определена. То же верно и для второго слагаемого. $\square$

Следствие. Временна́я компонента тензора энергии-импульса электромагнитного поля удовлетворяет сильному энергетическому условию (21.130).

Доказательство. Поскольку след тензора энергии-импульса равен нулю, то понятие слабого и сильного энергетического условия эквивалентны. $\square$

Предложение 23.3.7. *Действие электромагнитного поля в аффинной геометрии (23.138) инвариантно относительно преобразований Вейля:*

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \qquad A_\alpha \mapsto \bar{A}_\alpha = A_\alpha, \tag{23.166}$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ – произвольная вещественнозначная функция.

Доказательство. Утверждение следует из равенства

$$\sqrt{|g|}g^{\alpha\beta}g^{\gamma\delta} = \sqrt{|\bar{g}|}\bar{g}^{\alpha\beta}\bar{g}^{\gamma\delta}. \qquad \square$$

Замечание. Вейлевская инвариантность действия электромагнитного поля является спецификой четырехмерности пространства-времени. Если размерность пространства-времени отлична от четырех, то след тензора энергии-импульса электромагнитного поля отличен от нуля. $\square$

Равенство нулю следа тензора энергии-импульса электромагнитного поля связано с наличием вейлевской калибровочной инвариантности. Этот факт является общим.

Предложение 23.3.8. *Пусть действие зависит от метрики $g_{\alpha\beta}$ и некоторого набора полей материи φ^{A}*

$$S_{\text{М}} = \int dx L_{\text{М}}(g, \varphi).$$

Если действие инвариантно относительно преобразований Вейля, не затрагивающих поля материи,

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \qquad \varphi^{\text{A}} \mapsto \bar{\varphi}^{\text{A}} = \varphi^{\text{A}},$$

то след тензора энергии-импульса полей материи равен нулю, $T_{\text{М}\alpha}{}^{\alpha} = 0$.

Доказательство. Поскольку параметр преобразования Вейля зависит от точки пространства-времени, то справедлива вторая теорема Нетер. В рассматриваемом случае это означает следующее. Инвариантность действия имеет вид

$$\delta S_{\text{м}} = \int dx \frac{\delta S_{\text{м}}}{\delta g^{\alpha\beta}} \delta g^{\alpha\beta} = 0.$$

Учтем определение тензора энергии-импульса в общей теории относительности (21.10) и выражение для бесконечно малых преобразований Вейля

$$\delta g^{\alpha\beta} = -2\phi g^{\alpha\beta}.$$

Тогда равенство нулю вариации действия полей материи равносильно равенству $g^{\alpha\beta} T_{\text{м}\alpha\beta} = 0$. □

Пример 23.3.5. Бозонная струна описывается набором скалярных безмассовых полей $X^a(x)$, $a = 0, 1, \ldots, \text{D} - 1$, заданных на двумерном пространстве-времени (поверхности) $\mathbb{M}$ и принимающих значения в D-мерном пространстве Минковского $\mathbb{R}^{1,\text{D}-1}$:

$$X: \quad \mathbb{M} \ni x = \{x^\alpha\} \mapsto X(x) = \{X^a(x)\} \in \mathbb{R}^{1,\text{D}-1}.$$

Действие для бозонной струны можно записать в виде

$$S = -\frac{1}{2} \int_{\mathbb{M}} dx \sqrt{|g|} g^{\alpha\beta} \partial_\alpha X^a \partial_\beta X^b \eta_{ab}.$$

Знак этого действия выбран таким образом, чтобы пространственные компоненты X^i, $i = 1, \ldots, \text{D} - 1$, описывались обычным действием (23.13) и давали положительный вклад в гамильтониан системы. Это действие инвариантно относительно преобразований Вейля:

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \qquad X^a \mapsto X^a.$$

Тензор энергии-импульса бозонной струны имеет вид

$$T_{\alpha\beta} := \frac{2}{\sqrt{|g|}} \frac{\delta S}{\delta g^{\alpha\beta}} = -\partial_\alpha X^a \partial_\beta X_a + \frac{1}{2} g_{\alpha\beta} g^{\gamma\delta} \partial_\gamma X^a \partial_\delta X_a,$$

где $X_a := X^b \eta_{ba}$. Его след, очевидно, равен нулю. Вейлевская инвариантность действия для безмассового скалярного поля является спецификой двумерия. В пространстве-времени бóльшего числа измерений она отсутствует. □

23.4 Поле Прокáо

Поле Прокáо $A = \{A_\alpha\}$, $\alpha = 0, 1, \ldots, n - 1$, описывает массивные векторные частицы со спином единица [**?**, **?**]. В пространстве Минковского $\mathbb{R}^{1,n-1}$ ему соответствует лагранжиан

$$L = -\frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} + \frac{1}{2} m^2 A^\alpha A_\alpha, \tag{23.167}$$

где

$$F_{\alpha\beta} := \partial_\alpha A_\beta - \partial_\beta A_\alpha$$

и $m > 0$ – масса поля. Он отличается от лагранжиана электромагнитного поля только массовым членом. Знак перед массовым членом выбран так, чтобы вклад пространственных компонент A_μ, $\mu = 1, \ldots, n-1$, в плотность энергии был положителен.

Уравнения движения для лагранжиана (23.167) имеют вид

$$S,^\alpha := \frac{\delta S}{\delta A_\alpha} = \partial_\beta F^{\beta\alpha} + m^2 A^\alpha = 0. \tag{23.168}$$

Ввиду антисимметрии относительно перестановки индексов, $F^{\alpha\beta} = -F^{\beta\alpha}$, дивергенция этой системы уравнений приводит к дифференциальному уравнению первого порядка на компоненты поля:

$$\partial_\alpha A^\alpha = 0, \tag{23.169}$$

которое должно быть выполнено для любого решения систему уравнений (23.168). Полученное условие имеет вид калибровки Лоренца (23.94) для электромагнитного поля. С учетом условия (23.169) уравнения движения (23.168) принимают вид

$$\Box A^\alpha + m^2 A^\alpha = 0, \qquad \Box := \partial^\alpha \partial_\alpha. \tag{23.170}$$

Таким образом для каждой компоненты векторного поля возникает уравнение Клейна–Гордона–Фока. Однако поле Прокáо описывает не n распространяющихся степеней свободы, а только $n-1$, поскольку на компоненты поля наложено дополнительное условие (23.169). Чтобы убедиться в этом проведем более детальный анализ уравнений движения.

Запишем систему уравнений (23.168) отдельно для временнóй и пространственных компонент:

$$\partial_\mu F^{\mu 0} + m^2 A^0 = 0, \tag{23.171}$$

$$\partial_0 F^{0\mu} + \partial_\nu F^{\nu\mu} + m^2 A^\mu = 0. \tag{23.172}$$

Предположим, что все поля убывают на бесконечности. Тогда согласно теореме 23.3.1 оператор Лапласа обратим и пространственные компоненты поля Прокáо, как и в случае электромагнитного поля (23.79), можно взаимно однозначно разложить на поперечную и продольную составляющие:

$$A_\mu = A^{\text{т}}_\mu + \partial_\mu \varphi. \tag{23.173}$$

Тогда лагранжевы уравнения движения (23.171), (23.172) эквивалентны следующей системе уравнений:

$$-\triangle A_0 + \triangle \dot\varphi + m^2 A_0 = 0, \tag{23.174}$$

$$\partial_\mu \ddot\varphi - \partial_\mu \dot A_0 + m^2 \partial_\mu \varphi = 0, \tag{23.175}$$

$$\Box A^{\text{т}}_\mu + m^2 A^{\text{т}}_\mu = 0, \tag{23.176}$$

где мы расщепили уравнение (23.172) на продольную и поперечную части, а также ввели обозначения для производной по времени и операторов Лапласа и Даламбера:

$$\dot\varphi := \partial_0 \varphi, \qquad \triangle := -\partial^\mu \partial_\mu, \qquad \Box := \partial^\alpha \partial_\alpha = \partial_0^2 - \triangle.$$

Уравнения для поперечных компонент (23.176) – это уравнения Клейна–Гордона–Фока, и они полностью отщепились. Рассмотрим уравнение (23.174) как уравнение на временнýю компоненту:

$$(\triangle - m^2) A_0 = \triangle \dot\varphi. \tag{23.177}$$

Это – уравнение Гельмгольца, решение которого существует и единственно для достаточно широкого класса граничных условий, в частности, для полей, убывающих на бесконечности [**?**]. Запишем его решение в виде

$$A_0 = \frac{1}{\triangle - m^2}\triangle\dot{\varphi}. \tag{23.178}$$

Для полей, убывающих на бесконечности, уравнение (23.175) можно один раз проинтегрировать:

$$\ddot{\varphi} - \dot{A}_0 + m^2\varphi = 0.$$

Подставив сюда решение для A_0 (23.177) и умножив на оператор $\triangle - m^2$, получим уравнение Клейна–Гордона–Фока на функцию φ,

$$\Box\varphi + m^2\varphi = 0. \tag{23.179}$$

Поскольку все проделанные операции обратимы, то мы доказали

Предложение 23.4.1. *Для полей, убывающих на бесконечности, лагранжевы уравнения движения (23.168) для поля Прока́ эквивалентны $n-1$ уравнениям Клейна–Гордона–Фока (23.176), (23.179), дополненных уравнениями для продольной составляющей (23.82) и выражением для временно́й компоненты (23.178).*

Для тех, у кого остались сомнения по поводу следствия уравнений движения (23.169), скажем, что подстановка решения (23.178) в это условие, что нетрудно проверить, также приводит к уравнению Клейна–Гордона–Фока (23.179).

Таким образом, чтобы поставить задачу Коши для поля Прока́, необходимо и достаточно поставить задачу Коши для $n-2$ поперечных компонент $A^{\mathrm{т}}_\mu$ и поля φ, определяющего продольную составляющую. Поэтому поле Прока́ действительно описывает $n-1$ физическую (распространяющуюся) степень свободы.

23.4.1 Гамильтонова формулировка

Векторному полю A_α соответствуют сопряженные импульсы

$$P^\alpha := \frac{\partial L}{\partial(\partial_0 A_\alpha)} = F^{\alpha 0}. \tag{23.180}$$

Скобки Пуассона между координатами и импульсами имеют тот же вид, что и для электромагнитного поля (23.106). Как и для электромагнитного поля, из выражения для импульсов (23.180) следует первичная связь

$$G_1 = P^0 \approx 0. \tag{23.181}$$

Других первичных связей рассматриваемая модель не содержит. Несложные вычисления приводят к следующему выражению для гамильтониана

$$\mathcal{H} = \int d\boldsymbol{x}\left(-\frac{1}{2}P^\mu P_\mu + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - A_0\partial_\mu P^\mu - \frac{1}{2}m^2 A^\alpha A_\alpha + \lambda P^0\right), \tag{23.182}$$

где λ – множитель Лагранжа.

Гамильтоновым уравнениям движения имеют следующий вид:

$$\dot{A}_0 = \lambda, \tag{23.183}$$
$$\dot{A}_\mu = -P_\mu + \partial_\mu A_0, \tag{23.184}$$
$$\dot{P}^0 = \partial_\mu P^\mu + m^2 A_0, \tag{23.185}$$
$$\dot{P}^\mu = \partial_\nu F^{\nu\mu} + m^2 A^\mu. \tag{23.186}$$

Кроме этого, вариация действия по множителю Лагранжа приводит к связи (23.181). Таким образом, полная система гамильтоновых уравнений движения состоит из уравнений (23.183)–(23.186) и связи (23.181). На первый взгляд решения гамильтоновых уравнений содержат функциональный произвол, т.к. множитель Лагранжа λ произволен. Однако этот вывод не верен. Действительно, поскольку связь $P^0 = 0$ для любого решения гамильтоновых уравнений должна выполняться в любой момент времени, то производная по времени от уравнения (23.185) должна быть равна нулю. Это приводит к уравнению

$$m^2 \dot{A}_0 + \partial_\mu \dot{P}^\mu = m^2(\dot{A}_0 + \partial_\mu A^\mu) = 0, \tag{23.187}$$

где мы воспользовались уравнением (23.186). Сравнивая полученное равенство с уравнением (23.183), заключаем, что для любого решения гамильтоновых уравнений множитель Лагранжа имеет вид

$$\lambda = -\partial_\mu A^\mu. \tag{23.188}$$

Теперь можно доказать эквивалентность лагранжевой и гамильтоновой формулировок.

Предложение 23.4.2. *Полная система гамильтоновых уравнений (23.183)–(23.186) и (23.181) эквивалентна лагранжевым уравнениям движения (23.168).*

Доказательство. Пусть выполнены гамильтоновы уравнения (23.181)–(23.186). Из уравнения (23.184) находим импульсы

$$P_\mu = \partial_\mu A_0 - \dot{A}_\mu = F_{\mu 0}.$$

Подстановка этого выражения в формулу (23.186) приводит к уравнению (23.172). Уравнение (23.185) с учетом связи (23.181) дает уравнение (23.171).

Обратно. Пусть выполнены лагранжевы уравнения движения (23.171), (23.172). Введем обозначение $P^\alpha := F^{\alpha 0}$. Отсюда немедленно следует связь (23.181). А для пространственных значений индекса определение $P^\mu = F^{\mu 0}$ эквивалентно уравнению (23.184). Используя новую переменную P^μ, лагранжевы уравнения (23.172) переписываются в виде гамильтонова уравнения (23.186). Соотношение (23.169) является следствием лагранжевых уравнений движения, и оно дает уравнение (23.183) при выполнении условия (23.188). Наконец, лагранжево уравнение (23.171) сводится к уравнению (23.185) при выполнении связи (23.181). □

Замечание. Отметим, что для эквивалентности лагранжевой и гамильтоновой формулировок к гамильтониану необходимо добавить первичную связь (23.181). В противном случае никакой эквивалентности нет. Действительно, равенство $\lambda = 0$ явно противоречит условию (23.188). □

Продолжим построение гамильтонова формализма, согласно общей схеме, рассмотренной в разделе 20.2.

Производная по времени от первичной связи (23.181) равна

$$\dot{G}_1 = [G_1, \mathcal{H}] = \partial_\mu P^\mu + m^2 A^0.$$

Поскольку первичная связь должна сохраняться во времени, то возникает вторичная связь

$$G_2 = \partial_\mu P^\mu + m^2 A^0 \approx 0. \tag{23.189}$$

Нетрудно вычислить производную по времени от вторичной связи:

$$\dot{G}_2 = [G_2, \mathcal{H}] = m^2(\partial_\mu A^\mu + \lambda).$$

Сохранение вторичной связи во времени определяет множитель Лагранжа:

$$\lambda = -\partial_\mu A^\mu. \tag{23.190}$$

Поэтому в теории не возникает других связей.

Равенство (23.190) уже было получено ранее (23.188) при анализе гамильтоновых уравнений движения.

Скобка связей между собой имеет вид

$$[G_1, G_2'] = -m^2 \delta(\boldsymbol{x} - \boldsymbol{x}').$$

Введем обозначение для полного набора связей:

$$\{\Phi^{\text{м}}\} = \{G_1, G_2\}, \qquad \text{м} = 1, 2.$$

Тогда матрица скобок Пуассона для связей примет вид

$$[\Phi^{\text{м}}, \Phi'^{\text{н}}] = \begin{pmatrix} 0 & -m^2 \\ m^2 & 0 \end{pmatrix} \delta(\boldsymbol{x} - \boldsymbol{x}').$$

Она, очевидно, невырождена, и, следовательно, связи G_1 и G_2 являются связями второго рода.

Полный гамильтониан получается после добавления всех связей:

$$\mathcal{H}_{\text{т}} = \int d\boldsymbol{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_0 \partial_\mu P^\mu - \frac{1}{2} m^2 A^\alpha A_\alpha + \lambda P^0 + \mu(\partial_\mu P^\mu + m^2 A^0) \right). \tag{23.191}$$

Поскольку поле Прока́ содержит только связи второго рода, то полный гамильтониан совпадает с расширенным, $\mathcal{H}_{\text{т}} = \mathcal{H}_{\text{е}}$.

Один из множителей Лагранжа уже найден (23.190). Второй множитель Лагранжа определяется из условия сохранения первичной связи во времени,

$$\dot{G}_1 = [G_1, \mathcal{H}_{\text{т}}] = \partial_\mu P^\mu + m^2 A^0 - m^2 \mu = 0.$$

Отсюда вытекает равенство

$$\mu = A^0 + \frac{1}{m^2} \partial_\mu P^\mu. \tag{23.192}$$

Предложение 23.4.3. *Гамильтоновы уравнения движения для полного гамильтониана (23.191) эквивалентны лагранжевым уравнениям (23.168).*

Доказательство. Добавлению к исходному гамильтониану (23.182) вторичной связи приводит к возникновению самой связи из вариационного принципа. Кроме того меняются два гамильтоновых уравнения:

$$\dot{A}_\mu = -P_\mu + \partial_\mu A_0 - \partial_\mu \mu,$$
$$\dot{P}^0 = \partial_\mu P^\mu + m^2 A_0 - \mu m^2.$$

Если выполнена вторичная связь, то из соотношения (23.192) следует равенство $\mu = 0$. Поэтому гамильтоновы уравнения для полного гамильтониана эквивалентны исходным гамильтоновым уравнениям и, в силу предложения 23.4.2, лагранжевым уравнениям (23.168). □

Согласно общей схеме работы, описанной в разделе 20.2.2, решение связей второго рода можно подставлять в действие. В рассматриваемом случае связи легко решаются:

$$A_0 = -\frac{1}{m^2}\partial_\mu P^\mu, \qquad P_0 = 0.$$

После подстановки этого решения в формулу (23.191) получим гамильтониан для физических степеней свободы

$$\mathcal{H}_{\text{ph}} = \int d\boldsymbol{x} \left(-\frac{1}{2} P^\mu P_\mu + \frac{1}{2m^2}(\partial_\mu P^\mu)^2 + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} m^2 A^\mu A_\mu \right). \tag{23.193}$$

Таким образом, временна́я компонента поля Прока́ является нефизической и может быть полностью исключена из теории путем решения пары связей второго рода. Пространственные компоненты A_μ и сопряженные импульсы P^μ являются физическими (т.е. распространяющимися). Их динамика полностью определяется гамильтонианом для физических степеней свободы (23.193). Мы видим, что поле Прока́ описывает $n-1$ степень свободы, что на единицу превышает количество степеней свободы электромагнитного поля.

Гамильтонова плотность для физических степеней свободы поля Прока́ (23.193), очевидно, положительно определена. В то же время гамильтоновы плотности исходного (23.182) и полного (23.191) гамильтонианов не являются положительно определенными, т.к. содержат множители Лагранжа.

Гамильтоновы уравнения движения для физических степеней свободы, следующие из гамильтониана (23.193) имеют вид

$$\dot{A}_\mu = -P_\mu - \frac{1}{m^2}\partial_\mu \partial_\nu P^\nu, \tag{23.194}$$
$$\dot{P}^\mu = \partial_\nu F^{\nu\mu} + m^2 A^\mu. \tag{23.195}$$

Здесь проявляется важная особенность теории поля: в отличие от гамильтонового описания динамики точечных частиц, уравнение (23.194) для импульсов является не алгебраическим, а дифференциальным. Поэтому выразить импульсы через скорости, чтобы вернуться к лагранжеву описанию, не удается.

Чтобы сравнить гамильтонову и лагранжеву динамику физических степеней свободы, разложим координаты и импульсы на поперечную и продольную составляющие. Разложение пространственных компонент уже было введено (23.173). Для импульсов мы пишем

$$P_\mu = P^{\text{т}}_\mu + \partial_\mu \chi.$$

Тогда гамильтоновы уравнения (23.194), (23.195) также раскладываются на поперечные и продольные части:

$$\dot{A}^{\mathrm{T}}_{\mu} = -P^{\mathrm{T}}_{\mu}, \tag{23.196}$$

$$\partial_\mu \dot{\varphi} = -\partial_\mu \chi + \frac{\triangle}{m^2}\partial_\mu \chi, \tag{23.197}$$

$$\dot{P}^{\mathrm{T}}_{\mu} = -\triangle A^{\mathrm{T}}_{\mu} + m^2 A^{\mathrm{T}}_{\mu}, \tag{23.198}$$

$$\partial_\mu \dot{\chi} = m^2 \partial_\mu \varphi. \tag{23.199}$$

Гамильтоновы уравнения для поперечных компонент (23.196), (23.198) отщепились и, как легко видеть, эквивалентны уравнениям Клейна–Гордона–Фока (23.176).

Для полей, убывающих на бесконечности, уравнения (23.197) и (23.199) можно один раз проинтегрировать:

$$\dot{\varphi} = -\chi + \frac{\triangle}{m^2}\chi,$$
$$\dot{\chi} = m^2 \varphi.$$

Первое уравнение можно решить относительно χ,

$$\chi = \frac{m^2}{\triangle - m^2}\dot{\varphi},$$

и подставить данное решение во второе. В результате получим уравнение Клейна–Гордона–Фока (23.179). Поскольку все проделанные операции были обратимы, то мы доказали

Предложение 23.4.4. *Для полей, убывающих на бесконечности, гамильтоновы уравнения движения для физических степеней свободы (23.194), (23.195) эквивалентны лагранжевым уравнениям для физических степеней свободы (23.176), (23.179).*

Сейчас проявилось очень важное преимущество гамильтонова подхода, как более гибкого. А именно, выделение физических степеней свободы на гамильтоновом языке было проведено явно, и получено локальное выражение для соответствующего гамильтониана. В то же время, выделение физических степеней свободы в лагранжевом формализме нелокально, т.к. для нахождения поля φ необходимо решить дифференциальное уравнение (23.82).

23.5 Поле Янга–Миллса

Поля Янга–Миллса (калибровочные поля) играют важную роль в математической физике, потому что лежат в основе современных моделей, объединяющих электромагнитные, слабые и сильные взаимодействия. Экспериментальное их обнаружение пока остается под вопросом.

С геометрической точки зрения у нас есть следующая конструкция. Во-первых, мы предполагаем, что пространство-время представляет собой многообразие $\mathbb{M}$. Для общности будем рассматривать пространство-время произвольной размерности n. Затем строится главное расслоение $\mathbb{P}(\mathbb{M}, \pi, \mathbb{G})$ с некоторой структурной группой Ли $\mathbb{G}$ размерности $\mathsf{dim}\,\mathbb{G} = \mathrm{N}$, о которой мы поговорим несколько позже, и проекцией π. Мы предполагаем, что на главном расслоении задана связность, т.е. инвариантное

распределение горизонтальных подпространств (см. раздел 14.1). Каждая связность взаимно однозначно определяется формой связности, заданной на пространстве главного расслоения $\mathbb{P}$. Если задано координатное покрытие базы $\mathbb{M} = \cup_i \mathbb{U}_i$, и семейство локальных сечений $\sigma_i : \mathbb{U}_i \to \mathbb{P}$, то форма связности взаимно однозначно определяет семейство локальных форм связности, заданных на каждой координатной окрестности $\mathbb{U}_i$, которые определяются формой связности с помощью возвратов отображений σ_i. Локальная форма связности – это 1-форма на координатной окрестности $\mathbb{U}_i$ со значениями в алгебре Ли $\mathfrak{g}$ группы ли $\mathbb{G}$:

$$\omega = dx^\alpha A_\alpha{}^{\text{A}} L_{\text{A}}, \tag{23.200}$$

где L_{A}, $\text{A} = 1, \ldots, \text{N}$, – базис алгебры Ли (см. раздел 9.1). Компоненты локальной формы связности $A_\alpha{}^{\text{A}}(x)$ это и есть поля Янга–Миллса (после добавления соответствующих уравнений движения, что будет сделано позже). Базис алгебры Ли удовлетворяет некоторым коммутационным соотношениям

$$[L_{\text{A}}, L_{\text{B}}] = f_{\text{AB}}{}^{\text{C}} L_{\text{C}},$$

где $f_{\text{AB}}{}^{\text{C}}$ – структурные константы группы Ли $\mathbb{G}$.

При изменении локального сечения $\sigma \mapsto \sigma'$ (или при вертикальном автоморфизме) компоненты локальной формы связности меняются $A_\alpha{}^{\text{A}} \mapsto A'_\alpha{}^{\text{A}}$. Правило преобразования компонент связности известно (14.13), однако оно явно содержит функцию композиции для группы Ли $\mathbb{G}$ и поэтому неудобно. Инфинитезимальные калибровочные преобразования с параметрами $\epsilon^{\text{A}}(x)$ имеют вид

$$A'_\alpha{}^{\text{A}} = A_\alpha{}^{\text{A}} + \nabla_\alpha \epsilon^{\text{A}}, \tag{23.201}$$

где

$$\nabla_\alpha \epsilon^{\text{A}} := \partial_\alpha \epsilon^{\text{A}} - A_\alpha{}^{\text{B}} \epsilon^{\text{C}} f_{\text{BC}}{}^{\text{A}}.$$

Чтобы записать конечные калибровочные преобразования в удобном виде, рассмотрим произвольное точное представление группы Ли $\mathbb{G}$. Для определенности рассмотрим присоединенное представление. В этом представлении каждому элементу базиса алгебры Ли $\mathfrak{g}$ соответствует матрица: $L_{\text{A}} \mapsto f_{\text{AB}}{}^{\text{C}}$, где первый индекс A нумерует базисные векторы, а второй и третий рассматриваются, как матричные. Тогда каждой локальной форме связности можно поставить в соответствие матричнозначную 1-форму с компонентами

$$A_{\alpha\text{B}}{}^{\text{C}} := -A_\alpha{}^{\text{A}} f_{\text{AB}}{}^{\text{C}}. \tag{23.202}$$

(Знак минус не имеет принципиального значения.) Тогда поле Янга–Миллса при изменении локального сечения преобразуется по-правилу (14.15)

$$A_\alpha \mapsto A'_\alpha = S A_\alpha S^{-1} + \partial_\alpha S S^{-1}, \tag{23.203}$$

где $S(x) = \{S_{\text{A}}{}^{\text{B}}(x)\}$ – матрица присоединенного представления, и мы, для краткости, опустили матричные индексы. Это – хорошо известное калибровочное преобразование полей Янга–Миллса.

Каждой локальной форме связности соответствует локальная форма кривизны (см. раздел 14.2). Она представляет собой 2-форму на координатной окрестности $\mathbb{U}_i$ со значениями в алгебре Ли:

$$R = dx^\alpha \wedge dx^\beta F_{\alpha\beta}{}^{\text{A}} L_{\text{A}}, \tag{23.204}$$

где компоненты имеют вид (14.34)

$$F_{\alpha\beta}{}^{\mathrm{A}} = \partial_\alpha A_\beta{}^{\mathrm{A}} - \partial_\beta A_\alpha{}^{\mathrm{A}} - A_\alpha{}^{\mathrm{B}} A_\beta{}^{\mathrm{C}} f_{\mathrm{BC}}{}^{\mathrm{A}}. \tag{23.205}$$

Это есть напряженность поля Янга–Миллса. Ранее было показано, что при калибровочном преобразовании (23.203) компоненты напряженности преобразуются ковариантным образом (14.35)

$$F_{\alpha\beta}{}^{\mathrm{A}} \mapsto F'_{\alpha\beta}{}^{\mathrm{A}} = F_{\alpha\beta}{}^{\mathrm{B}} S^{-1}{}_{\mathrm{B}}{}^{\mathrm{A}}.$$

Если ввести матричнозначную форму кривизны

$$F_{\alpha\beta\mathrm{B}}{}^{\mathrm{C}} := -F_{\alpha\beta}{}^{\mathrm{A}} f_{\mathrm{AB}}{}^{\mathrm{C}}, \tag{23.206}$$

то ее закон преобразования будет выглядеть немного иначе:

$$F_{\alpha\beta} \mapsto F'_{\alpha\beta} = S F_{\alpha\beta} S^{-1}, \tag{23.207}$$

где мы, для краткости, опустили матричные индексы.

Матричнозначные компоненты напряженности можно выразить через матричнозначные поля Янга–Миллса:

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha - [A_\alpha, A_\beta], \tag{23.208}$$

где $[A_\alpha, A_\beta]$ – коммутатор матриц, и мы опустили матричные индексы. Чтобы получить данное выражение для компонент напряженности из формулы (23.205), использованы тождества Якоби для структурных констант (9.10).

Согласно теореме 14.2.2 каждая форма кривизны удовлетворяет тождествам Бианки. Тождества Бианки для локальной формы кривизны были получены ранее (14.2.6):

$$\nabla_\alpha F_{\beta\gamma}{}^{\mathrm{A}} + \nabla_\beta F_{\gamma\alpha}{}^{\mathrm{A}} + \nabla_\gamma F_{\alpha\beta}{}^{\mathrm{A}} = 0, \tag{23.209}$$

где

$$\nabla_\alpha F_{\beta\gamma}{}^{\mathrm{A}} = \partial_\alpha F_{\beta\gamma}{}^{\mathrm{A}} - A_\alpha{}^{\mathrm{B}} F_{\beta\gamma}{}^{\mathrm{C}} f_{\mathrm{BC}}{}^{\mathrm{A}} \tag{23.210}$$

– ковариантная производная (по отношению к калибровочным преобразованиям) от компонент напряженности поля Янга–Миллса.

Предложение 23.5.1. *Для существования такого калибровочного преобразования, чтобы преобразованное поле Янга–Миллса обратилось бы в нуль в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ необходимо и достаточно, чтобы локальная форма кривизны связности была равна нулю в $\mathbb{U}$.*

Доказательство. Если локальная форма связности равна нулю, то локальная форма кривизны (23.205) также обращается в нуль.

Обратно. Рассмотрим правило преобразования полей Янга–Миллса (23.203) как дифференциальное уравнение на матрицу S. Если преобразованное поле Янга–Миллса равно нулю, $A'_\alpha = 0$, то матрица преобразования должна удовлетворять системе дифференциальных уравнений:

$$\partial_\alpha S = -S A_\alpha. \tag{23.211}$$

Эта система локально разрешима тогда и только тогда, когда выполнены условия разрешимости:

$$\partial_{[\beta}\partial_{\alpha]} S = 0,$$

где квадратные скобки обозначают антисимметризацию по индексам. Вторая частная производная от матрицы преобразования имеет вид

$$\partial_\beta\partial_\alpha S = -\partial_\beta S A_\alpha - S\partial_\beta A_\alpha = S(A_\beta A_\alpha - \partial_\beta A_\alpha),$$

где в первом слагаемом мы воспользовались исходным уравнением (23.211). Антисимметризация полученного выражения по индексам α и β приводит к равенству

$$F_{\alpha\beta} = 0.$$

Таким образом, равенство нулю локальной формы кривизны является необходимым и достаточным условием разрешимости системы уравнений (23.211) в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$. □

Определение. Из доказанного утверждения следует, что если локальная форма кривизны равна нулю, то в некоторой окрестности $\mathbb{U} \subset \mathbb{M}$ компоненты соответствующей локальной формы связности имеют вид

$$A_\alpha = -S^{-1}\partial_\alpha S = \partial_\alpha S^{-1} S. \tag{23.212}$$

Такое поле Янга–Миллса называется *чистой калибровкой.* □

Легко проверить, что напряженность поля Янга–Миллса для поля (23.212) действительно обращается в нуль.

Описанная выше геометрическая конструкция не зависит от того задана ли на базе $\mathbb{M}$ какая либо геометрия, т.е. метрика и аффинная связность, или нет. Главное расслоение, связность и кривизна определяются в дифференциальной геометрии самостоятельно. Закон преобразования компонент локальных форм связности и кривизны при изменении сечения следует из определения. Для этого не надо вводить какие либо дополнительные поля и ковариантные производные. Не смотря на отсутствие аффинной связности в частных производных при определении локальной формы кривизны (23.205) и в тождествах Бианки (23.210), все выписанные соотношения ковариантны относительно общих преобразований координат.

Следующий шаг построения моделей состоит в построении действия. Мы считаем, что действие должно быть инвариантно относительно общих преобразований координат и калибровочных преобразований (23.203). Чтобы построить соответствующий инвариант нам понадобятся две метрики: метрика $g_{\alpha\beta}$ на пространственно-временно́м многообразии и двусторонне инвариантная метрика η_{AB} на групповом многообразии. Соответствующее действие имеет вид

$$S_{\text{YM}} = \int_{\mathbb{M}} dx\, L_{\text{YM}}, \qquad L_{\text{YM}} := -\frac{1}{4}\sqrt{|g|}\, g^{\alpha\gamma} g^{\beta\delta} F_{\alpha\beta}{}^{\text{A}} F_{\gamma\delta}{}^{\text{B}} \eta_{\text{AB}}. \tag{23.213}$$

В квантовой теории поля в качестве пространства-времени выбирается пространство Минковского $\mathbb{R}^{1,n-1}$, группа общих преобразования координат сужается до группы Пуанкаре $\mathbb{IO}(1,n-1)$, и в качестве пространственно-временно́й метрики выбирается метрика Лоренца, $g_{\alpha\beta} \mapsto \eta_{\alpha\beta}$.

Если калибровочная группа $\mathbb{G}$ является абелевой, как, например, в электродинамике, то в качестве двусторонне инвариантной метрики подойдет любая постоянная матрица. Действительно, при калибровочных преобразованиях (23.203) компоненты напряженности поля Янга–Миллса для абелевой группы вообще не преобразуются.

Для неабелевых групп в качестве двустронне инвариантной метрики выберем форму Киллинга–Картана (см. раздел 9.45). Согласно теореме 9.5.1 форма Киллинга–Картана невырождена тогда и только тогда, когда группа Ли $\mathbb{G}$ полупроста. Поэтому мы предполагаем, что структурная группа Ли $\mathbb{G}$ полупроста.

Согласно теореме 9.11.5 любая полупростая группа Ли единственным образом представляется в виде прямого произведения простых групп. В этом случае для каждого сомножителя можно написать отдельный инвариант вида (23.213) со своей константой связи. Поэтому в дальнейшем, для простоты, мы ограничимся рассмотрением простых групп Ли. Если мы научимся работать с простыми группами Ли, то построить модель, соответствующую их прямому произведению, не составит труда.

В дальнейшем мы покажем, что канонический гамильтониан для физических степеней свободы, которые описываются действием (23.213), положительно определен тогда и только тогда, когда метрика η_{AB} положительно определена. Согласно теореме 9.5.2 связная полупростая группа Ли компактна тогда и только тогда, когда ее форма Киллинга–Картана

$$\eta_{\mathrm{AB}} := -f_{\mathrm{AC}}{}^{\mathrm{D}} f_{\mathrm{BD}}{}^{\mathrm{C}} \tag{23.214}$$

положительно определена. Поэтому мы предполагаем, что структурная группа Ли $\mathbb{G}$ связна, проста и компактна.

Базис алгебры Ли L_{A} определен с точностью до невырожденных линейных преобразований

$$L_{\mathrm{A}} \mapsto S_{\mathrm{A}}{}^{\mathrm{B}} L_{\mathrm{B}}, \qquad S \in \mathbb{GL}(\mathrm{N}, \mathbb{R}).$$

При этом форма Киллинга–Картана преобразуется по тензорному закону

$$\eta_{\mathrm{AB}} \mapsto S_{\mathrm{A}}{}^{\mathrm{C}} S_{\mathrm{B}}{}^{\mathrm{D}} \eta_{\mathrm{CD}}.$$

Согласно теореме 28.1.9 произвольную симметричную матрицу можно привести к диагональному виду с помощью ортогональной матрицы. Дальнейшей растяжкой координат можно добиться того, что на диагонали формы Киллинга–Картана будут стоять единицы. Таким образом, не теряя общности, можно считать, что для связной компактной простой группы Ли форма Киллинга–Картана совпадает с единичной матрицей

$$\eta_{\mathrm{AB}} = \delta_{\mathrm{AB}}. \tag{23.215}$$

Пример 23.5.1. В единых моделях теории поля используются, как правило, унитарные $\mathbb{SU}(\mathrm{N})$ и ортогональные $\mathbb{SO}(\mathrm{N})$ группы. Эти группы связны, компактны и просты. □

Если вместо полей Янга–Миллса $A_\alpha{}^{\mathrm{A}}$ рассматривать их точное матричное представление, то лагранжиан (23.213) можно записать, используя след матриц,

$$L_{\mathrm{YM}} := \frac{1}{4}\sqrt{|g|}\, g^{\alpha\gamma} g^{\beta\delta}\, \mathsf{tr}\left(F_{\alpha\beta} F_{\gamma\delta}\right).$$

Обратим внимание на изменение общего знака, что связано с определением формы Киллинга–Картана (23.214).

Для построения физически содержательных моделей математической физики, необходимо введение дополнительных полей, например, скаляров и спиноров. Мы предполагаем, что эти поля являются локальными сечениями векторного ассоциированного расслоения $\mathbb{E}(\mathbb{M}, \pi_{\mathbb{E}}, \mathbb{V}, \mathbb{G}, \mathbb{P})$, типичным слоем которого является векторное пространство $\mathbb{V}$, в котором задано представление структурной группы $\mathbb{G}$. Отображение $\pi_{\mathbb{E}} : \mathbb{E} \to \mathbb{M}$ – это проекция. Каждое сечение ассоциированного расслоения

задается набором компонент $\varphi = \{\varphi^i\}$, $i = 1, \ldots, \dim \mathbb{V}$. При калибровочных преобразованиях поля преобразуются по некоторому представлению структурной группы:

$$\varphi \mapsto \varphi' = \varphi S^{-1} = \{\varphi^j S^{-1}{}_j^i\}, \tag{23.216}$$

где $S_j{}^i$ – матрица представления. Ковариантная производная имеет вид (14.24)

$$\nabla_\alpha \varphi^i = \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i, \tag{23.217}$$

в котором

$$A_{\alpha j}{}^i := -A_\alpha{}^{\mathrm{A}} L_{\mathrm{A}j}{}^i,$$

где $L_{\mathrm{A}j}{}^i$ – представление базиса алгебры Ли L_{A} в векторном пространстве $\mathbb{V}$. Нетрудно убедиться в том, что при калибровочных преобразованиях (23.203), (23.216) ковариантная производная ведет себя ковариантно:

$$\nabla_\alpha \varphi^i \mapsto \nabla'_\alpha \varphi'^i = \nabla_\alpha \varphi^j S^{-1}{}_j^i,$$

где

$$\nabla'_\alpha \varphi'^i = \partial_\alpha \varphi'^i + \varphi'^j A'_{\alpha j}{}^i,$$

Определение. В физике векторное пространство $\mathbb{V}$ называется *изотопическим*, а калибровочные преобразования (23.216) – *изотопическими вращениями*. □

При построении моделей математической физики мы предполагаем, что уравнения движения для полей φ, не взаимодействующих с полями Янга–Миллса, следуют из некоторого действия

$$S = \int_{\mathbb{M}} dx\, L(\varphi, \partial\varphi),$$

где лагранжиан L зависит от полей $\varphi = \{\varphi^i\}$ и их частных производных $\partial\varphi = \{\partial_\alpha \varphi^i\}$. Как и для электромагнитного поля, взаимодействие с полями Янга–Миллса вводится путем *минимальной подстановки*, которая заключается в замене всех частных производных на ковариантные:

$$\partial_\alpha \varphi^i \mapsto \nabla_\alpha \varphi^i := \partial_\alpha \varphi^i + \varphi^j A_{\alpha j}{}^i.$$

Кроме этого, к лагранжиану добавляется лагранжиан поля Янга–Миллса

$$L(\varphi, \partial\varphi) \mapsto -\frac{1}{4e^2} F^{\alpha\beta\mathrm{A}} F_{\alpha\beta\mathrm{A}} + L(\varphi, \nabla\varphi),$$

который приводит к динамическим уравнениям движения для самого поля Янга–Миллса. Выше e – константа связи и

$$F^{\alpha\beta\mathrm{A}} := g^{\alpha\gamma} g^{\beta\delta} \eta^{\mathrm{AB}} F_{\gamma\delta\mathrm{B}}.$$

Для построения теории возмущений удобно перейти к новым переменным $A_\alpha{}^{\mathrm{A}} \mapsto eA_\alpha{}^{\mathrm{A}}$. Тогда лагранжиан примет вид

$$-\frac{1}{4} F^{\alpha\beta\mathrm{A}} F_{\alpha\beta\mathrm{A}} + L(\varphi, \nabla\varphi),$$

где

$$\begin{aligned} F_{\alpha\beta}{}^{\mathrm{A}} &= \partial_\alpha A_\beta{}^{\mathrm{A}} - \partial_\beta A_\alpha{}^{\mathrm{A}} - eA_\alpha{}^{\mathrm{B}} A_\beta{}^{\mathrm{C}} f_{\mathrm{BC}}{}^{\mathrm{A}}, \\ \nabla_\alpha \varphi^i &= \partial_\alpha \varphi^i + e\varphi^j A_{\alpha j}{}^i. \end{aligned}$$

В таком виде удобно строить теорию возмущений при малой константе связи e. В дальнейшем, для упрощения обозначений, мы положим $e = 1$.

23.5.1 Лагранжева формулировка

В квантовой теории поля мы предполагаем, что пространством-временем является пространство Минковского, $\mathbb{M} \approx \mathbb{R}^{1,n-1}$, которое топологически тривиально, т.е. диффеоморфно $\mathbb{R}^n$. В этом случае согласно теореме 13.1.1 расслоение тривиально

$$\mathbb{P}(\mathbb{R}^{1,n-1}, \pi, \mathbb{G}) \approx \mathbb{R}^{1,n-1} \times \mathbb{G}.$$

Если выбрать систему координат, например, декартову, покрывающую все пространство-время, то связность на главном расслоении будет находиться во взаимно однозначном соответствии с локальной формой связности. По этой причине в физике часто говорят, что поля Янга–Миллса это и есть связность.

Действие для полей Янга–Миллса в пространстве Минковского $\mathbb{R}^{1,n-1}$ имеет вид

$$S_{\text{YM}} = \int dx L_{\text{YM}}, \qquad L_{\text{YM}} = -\frac{1}{4} F^{\alpha\beta_{\text{A}}} F_{\alpha\beta_{\text{A}}}. \tag{23.218}$$

где подъем и опускание греческих индексов производится с помощью метрики Минковского $\eta_{\alpha\beta}$, а латинских – с помощью формы Киллинга–Картана η_{AB}. Из этого действия следуют нелинейные уравнения движения

$$\nabla_\beta F^{\beta\alpha_{\text{A}}} := \partial_\beta F^{\beta\alpha_{\text{A}}} - A_\beta{}^{\text{B}} F^{\beta\alpha_{\text{C}}} f_{\text{BC}}{}^{\text{A}} = 0. \tag{23.219}$$

Добавление к лагранжиану Янга–Миллса массового члена $\frac{1}{2} m^2 A^{\alpha_{\text{A}}} A_{\alpha_{\text{A}}}$, как легко видеть, нарушает калибровочную инвариантность.

В отличие от электродинамики в пространстве Минковского для неабелевых калибровочных групп $\mathbb{G}$ уравнения (23.219) нелинейны. Поэтому говорить о "свободных" полях Янга–Миллса не имеет смысла. Взаимодействие полей Янга–Миллса есть всегда. В линейном приближении уравнения (23.219) имеют то же вид, что и набор из N электромагнитных полей. При этом инфинитезимальные калибровочные преобразования (23.203) в линейном приближении по полям принимают вид $\mathbb{U}(1)$ преобразований:

$$A'_\alpha{}^{\text{A}} = A_\alpha{}^{\text{A}} + \partial_\alpha \epsilon^{\text{A}} + \mathsf{o}_\alpha{}^{\text{A}}(\epsilon, A),$$

где $\epsilon^{\text{A}}(x)$, $\text{A} = 1, \ldots, \text{N}$, – параметры преобразований. Таким образом, в линейном приближении свободное поле Янга–Миллса описывает N безмассовых векторных полей.

По построению, действие калибровочно инвариантно. Согласно второй теореме Нетер в таком случае между уравнениями движения существует N линейных зависимостей по числу параметров преобразования. Для нахождения этих зависимостей достаточно знать вид инфинитезимальных калибровочных преобразований (23.201). Из определения инвариантности следует равенство

$$\delta S_{\text{YM}} = \int dx \frac{\delta S_{\text{YM}}}{\delta A_\alpha{}^{\text{A}}} \nabla_\alpha \epsilon^{\text{A}} = 0.$$

После интегрирования по частям, получим зависимость уравнений движения:

$$\nabla_\alpha \nabla_\beta F^{\beta\alpha}{}_{\text{A}} = 0, \qquad \text{A} = 1, \ldots, \text{N}. \tag{23.220}$$

В справедливости полученного равенства можно убедиться с помощью прямых вычислений. Действительно, в силу антисимметрии напряженности, $F^{\alpha\beta}{}_{\text{A}} = -F^{\beta\alpha}{}_{\text{A}}$, равенство (23.220) можно переписать в виде

$$\frac{1}{2}(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) F^{\beta\alpha}{}_{\text{A}} = \frac{1}{2} F_{\alpha\beta}{}^{\text{C}} F^{\beta\alpha\text{B}} f_{\text{CAB}} = 0,$$

где мы воспользовались свойством коммутатора ковариантных производных (14.42), симметрией тензора $F_{\alpha\beta}{}^{\mathrm{C}}F^{\beta\alpha}{}_{\mathrm{B}}$ по индексам C и B и антисимметрией структурных констант (предложение 9.5.1).

Если в правой части уравнений движения (23.219) стоит ток,

$$\nabla_\beta F^{\beta\alpha}{}_{\mathrm{A}} = J^\alpha{}_{\mathrm{A}}, \tag{23.221}$$

То из тождества (23.220) следует закон ковариантного сохранения тока:

$$\nabla_\alpha J^\alpha{}_{\mathrm{A}} = 0, \tag{23.222}$$

выполнение которого необходимо для самосогласованности уравнений движения. В свою очередь, если ток получен в результате варьирования некоторого калибровочно инвариантного действия для других полей, то равенство (23.222) будет автоматически выполнено на уравнениях движения, опять же, в силу второй теоремы Нетер. Если же ток $J^\alpha{}_{\mathrm{A}}$ вводится в уравнения движения "руками", то на него необходимо наложить условие сохранения (23.222).

Эффективным инструментом исследования свободных уравнений движения в электродинамике явилось разложение пространственных компонент электромагнитного потенциала на поперечную и продольную составляющие. Для каждого значения индекса A поле Янга–Миллса также можно разложить на поперечную и продольные части. Однако это не будет столь же эффективно, т.к. уравнения "свободного" поля Янга–Миллса (23.219) нелинейны. Такое разложение эффективно в квантовой теории поля при построении теории возмущений.

Теперь рассмотрим законы сохранения, вытекающие из первой теоремы Нетер. По-построению, действие Янга–Миллса инвариантно относительно глобального действия группы Пуанкаре $\mathbb{IO}(1,n-1)$. Из инвариантности относительно трансляций, согласно первой теореме Нетер, следует закон сохранения энергии-импульса (см. раздел 19.2.1),

$$\partial_\beta \widehat{T}_\alpha{}^\beta = 0,$$

где

$$\widehat{T}_{\alpha\beta} = -\partial_\alpha A_{\gamma_{\mathrm{A}}} F_\beta{}^{\gamma_{\mathrm{A}}} + \frac{1}{4}\eta_{\alpha\beta}F^2$$

– канонический тензор энергии-импульса поля Янга–Миллса. Это выражение страдает двумя недостатками. Во-первых, оно не инвариантно относительно калибровочных преобразований. Во-вторых, отсутствует симметрия относительно перестановки индексов. Для их устранения используем произвол в выборе канонического тензора энергии-импульса. Заметим, что выполнено равенство

$$\partial_\gamma(A_{\alpha_{\mathrm{A}}}F_\beta{}^{\gamma_{\mathrm{A}}}) = \partial_\gamma A_{\alpha_{\mathrm{A}}}F_\beta{}^{\gamma_{\mathrm{A}}} + A_{\alpha_{\mathrm{A}}}\nabla_\gamma F_\beta{}^{\gamma_{\mathrm{A}}} + A_{\alpha_{\mathrm{A}}}A_\gamma{}^{\mathrm{C}} f_{\mathrm{CB}}{}^{\mathrm{A}} F_\beta{}^{\gamma_{\mathrm{B}}}.$$

Если выполнены уравнения движения, что предполагается в первой теореме Нетер, то второе слагаемое обращается в нуль, и можно переопределить тензор энергии-импульса

$$T^{\mathrm{s}}_{\alpha\beta} := \widehat{T}_{\alpha\beta} + \partial_\gamma(A_{\alpha_{\mathrm{A}}}F_\beta{}^{\gamma_{\mathrm{A}}}) = -F_{\alpha\gamma_{\mathrm{A}}}F_\beta{}^{\gamma_{\mathrm{A}}} + \frac{1}{4}\eta_{\alpha\beta}F^2. \tag{23.223}$$

Этот тензор энергии-импульса, очевидно, калибровочно инвариантен и симметричен. На уравнениях движения он сохраняется:

$$\partial_\beta T^{\mathrm{s}}{}_\alpha{}^\beta = 0. \tag{23.224}$$

При лоренцевых вращениях поле Янга–Миллса преобразуется, как и электромагнитное поле (23.98). Поэтому для спинового момента получаем следующее выражение

$$S_{\alpha\beta}{}^{\gamma} = A_{\alpha\mathrm{A}}F_{\beta}{}^{\gamma\mathrm{A}} - A_{\beta\mathrm{A}}F_{\alpha}{}^{\gamma\mathrm{A}}. \tag{23.225}$$

Полный тензор момента количества движения (19.51) состоит из орбитального и спинового моментов:

$$\widehat{J}_{\alpha\beta}{}^{\gamma} = x_{\beta}\widehat{T}_{\alpha}{}^{\gamma} - x_{\alpha}\widehat{T}_{\beta}{}^{\gamma} + S_{\alpha\beta}{}^{\gamma}. \tag{23.226}$$

Используя выражения для симметричного тензора энергии-импульса (23.223) и спинового момента (23.225) его можно переписать в виде

$$\widehat{J}_{\alpha\beta}{}^{\gamma} = M_{\alpha\beta}{}^{\gamma} + \partial_{\delta}(x_{\alpha}A_{\beta\mathrm{A}}F^{\gamma\delta\mathrm{A}} - x_{\beta}A_{\alpha\mathrm{A}}F^{\gamma\delta\mathrm{A}}), \tag{23.227}$$

где

$$M_{\alpha\beta}{}^{\gamma} := x_{\beta}T^{\mathrm{s}}{}_{\alpha}{}^{\gamma} - x_{\alpha}T^{\mathrm{s}}{}_{\beta}{}^{\gamma}$$

– орбитальный момент, построенный по симметричному тензору энергии-импульса (23.223). Поскольку второе слагаемое в выражении (23.227) имеет вид (19.39), то тензор момента количества движения можно переопределить, положив

$$J_{\alpha\beta}{}^{\gamma} = M_{\alpha\beta}{}^{\gamma} = x_{\beta}T^{\mathrm{s}}{}_{\alpha}{}^{\gamma} - x_{\alpha}T^{\mathrm{s}}{}_{\beta}{}^{\gamma}, \tag{23.228}$$

который также сохраняется

$$\partial_{\gamma}J_{\alpha\beta}{}^{\gamma} = 0.$$

Отметим полное отсутствие спинового момента в этом выражении.

Поскольку действие для поля Янга–Миллса инвариантно относительно калибровочных преобразований, то оно также инвариантно относительно глобальных преобразований из калибровочной группы. Поэтому можно воспользоваться первой теоремой Нетер. С калибровочной инвариантностью действия для поля Янга–Миллса связан закон сохранения заряда

$$\partial_{\alpha}J^{\alpha}{}_{\mathrm{A}} = 0, \tag{23.229}$$

где

$$J^{\alpha}{}_{\mathrm{A}} = \nabla_{\beta}F^{\beta\alpha}{}_{\mathrm{A}} \tag{23.230}$$

– ток, соответствующий первой теореме Нетер, то есть, когда параметр калибровочных преобразований не зависит от точки пространства-времени. При отсутствии полей материи этот закон тривиален, т.к. ток равен нулю в силу уравнений движения (23.219).

23.5.2 Гамильтонова формулировка

Рассмотрим гамильтонову формулировку поля Янга–Миллса в пространстве Минковского $\mathbb{R}^{1,n-1}$ с действием (23.218). По-определению, импульсы, сопряженные компонентам $A_{\alpha}{}^{\mathrm{A}}$, имеют вид

$$P^{\alpha}{}_{\mathrm{A}} := \frac{\partial L_{\text{YM}}}{\partial \dot{A}_{\alpha}{}^{\mathrm{A}}} = F^{0\alpha}{}_{\mathrm{A}}, \tag{23.231}$$

где $\dot{A}_{\alpha}{}^{\mathrm{A}} := \partial_0 A_{\alpha}{}^{\mathrm{A}}$. Выпишем отличные от нуля одновременны́е скобки Пуассона:

$$[A_{\alpha}{}^{\mathrm{A}}, P'^{\beta}{}_{\mathrm{B}}] := [A_{\alpha}{}^{\mathrm{A}}(x^0, \boldsymbol{x}), P^{\beta}{}_{\mathrm{B}}(x^0, \boldsymbol{x}')] = \delta^{\beta}_{\alpha}\delta^{\mathrm{A}}_{\mathrm{B}}\delta(\boldsymbol{x} - \boldsymbol{x}').$$

Как и в случае электромагнитного поля в теории возникают первичные связи. Из определения импульсов (23.218) следует N первичных связей:

$$G_{1\mathrm{A}} = P^0{}_{\mathrm{A}} = 0, \qquad \mathrm{A} = 1, \ldots, \mathrm{N}. \tag{23.232}$$

Других первичных связей в модели нет. Скобки Пуассона первичных связей между собой, очевидно, равны нулю:

$$[G_{1\mathrm{A}}, G'_{1\mathrm{B}}] = 0.$$

Несложные вычисления приводят к гамильтониану

$$\mathcal{H} = \int d\boldsymbol{x} \left(-\frac{1}{2} P^{\mu\mathrm{A}} P_{\mu\mathrm{A}} + \frac{1}{4} F^{\mu\nu\mathrm{A}} F_{\mu\nu\mathrm{A}} - A_0{}^{\mathrm{A}} \nabla_\mu P^\mu{}_{\mathrm{A}} + \lambda^{\mathrm{A}} P^0{}_{\mathrm{A}} \right), \tag{23.233}$$

где мы добавили первичные связи с множителями Лагранжа λ^{A} и ковариантная производная от импульсов равна

$$\nabla_\mu P^\nu{}_{\mathrm{A}} := \partial_\mu P^\nu{}_{\mathrm{A}} - A_{\mu\mathrm{A}}{}^{\mathrm{B}} P^\nu{}_{\mathrm{B}}.$$

Гамильтоновы уравнения движения для построенного гамильтониана имеют вид

$$\dot{A}_0{}^{\mathrm{A}} = \lambda^{\mathrm{A}}, \tag{23.234}$$

$$\dot{A}_\mu{}^{\mathrm{A}} = -P_\mu{}^{\mathrm{A}} + \nabla_\mu A_0{}^{\mathrm{A}}, \tag{23.235}$$

$$\dot{P}^0{}_{\mathrm{A}} = \nabla_\mu P^\mu{}_{\mathrm{A}}, \tag{23.236}$$

$$\dot{P}^\mu{}_{\mathrm{A}} = \nabla_\nu F^{\nu\mu}{}_{\mathrm{A}} + A_{0\mathrm{A}}{}^{\mathrm{B}} P^\mu{}_{\mathrm{B}}, \tag{23.237}$$

где

$$\nabla_\mu A_0{}^{\mathrm{A}} := \partial_\mu A_0{}^{\mathrm{A}} + A_{\mu\mathrm{B}}{}^{\mathrm{A}} A_0{}^{\mathrm{B}},$$
$$\nabla_\nu F^{\rho\mu}{}_{\mathrm{A}} := \partial_\nu F^{\rho\mu}{}_{\mathrm{A}} - A_{\nu\mathrm{A}}{}^{\mathrm{B}} F^{\rho\mu}{}_{\mathrm{B}}.$$

Конечно, эти уравнения необходимо дополнить уравнениями первичных связей (23.232), которые возникают при варьировании соответствующего действия по множителям Лагранжа.

Гамильтоновы уравнения движения (23.234)–(23.248), содержат множители Лагранжа λ^{A}, которые в исходном действии рассматриваются в качестве произвольных функций. Однако, уравнения движения, в принципе, могут наложить на них некоторые ограничения. Чтобы проверить совместность гамильтоновых уравнений, рассмотрим эволюцию первичных связей во времени. Поскольку $P^0{}_{\mathrm{A}} = 0$, то из уравнения (23.247) следуют равенства

$$\nabla_\mu P^\mu{}_{\mathrm{A}} = 0, \tag{23.238}$$

которые должны быть выполнены для любого решения гамильтоновых уравнений. С учетом уравнения (23.248) производная по времени от полученных равенств имеет вид

$$\partial_0 (\nabla_\mu P^\mu{}_{\mathrm{A}}) = \nabla_\mu \nabla_\nu F^{\nu\mu}{}_{\mathrm{A}} + \nabla_\mu A_{0\mathrm{A}}{}^{\mathrm{B}} P^\mu{}_{\mathrm{B}} + A_{0\mathrm{A}}{}^{\mathrm{B}} \nabla_\mu P^\mu{}_{\mathrm{B}} - \dot{A}_{\mu\mathrm{A}}{}^{\mathrm{B}} P^\mu{}_{\mathrm{B}}. \tag{23.239}$$

Первое слагаемое обращается в нуль:

$$\frac{1}{2} (\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) F^{\nu\mu}{}_{\mathrm{A}} = \frac{1}{2} F_{\mu\nu}{}^{\mathrm{C}} F^{\nu\mu\mathrm{B}} f_{\mathrm{CAB}} = 0,$$

где мы воспользовались свойством коммутатора ковариантных производных (14.42), симметрией тензора $F_{\mu\nu}{}^{\text{C}}F^{\nu\mu\text{B}}$ по индексам C и B и антисимметрией структурных констант (предложение 9.5.1). Второе и четвертое слагаемые сокращаются с учетом уравнения (23.235). В результате возникает равенство

$$\partial_0(\nabla_\mu P^\mu{}_{\text{A}}) = A_{0\text{A}}{}^{\text{B}}\nabla_\mu P^\mu{}_{\text{B}} \approx 0, \tag{23.240}$$

т.к. правая часть равна нулю на уравнениях движения (23.238). На этом анализ заканчивается, и мы видим, что никаких ограничений на множители Лагранжа гамильтоновы уравнения движения не накладывают.

Продолжим общий гамильтонов анализ. Производная по времени от первичных связей равна

$$\dot G_1 = [G_1, \mathcal{H}] = \nabla_\mu P^\mu{}_{\text{A}}.$$

Это приводит к вторичным связям:

$$G_{2\text{A}} = \nabla_\mu P^\mu{}_{\text{A}} \approx 0. \tag{23.241}$$

Скобка Пуассона первичных и вторичных связей между собой равна нулю,

$$[G_{1\text{A}}, G'_{2\text{B}}] = 0.$$

Несложно вычислить скобки Пуассона вторичных связей между собой,

$$[G_{2\text{A}}, G'_{2\text{B}}] = -f_{\text{AB}}{}^{\text{C}}G_{2\text{C}}. \tag{23.242}$$

Правая часть, очевидно, обращается в нуль на поверхности связей.

Выше мы уже показали, что вторичные связи сохраняются во времени (23.240).

Таким образом, полная система связей

$$\{G_a\} := \{G_{1\text{A}}, G_{2\text{A}}\}, \qquad a = 1, \dots, 2\text{N}, \tag{23.243}$$

находится в инволюции, и модель содержит 2N связей первого рода.

Полный гамильтониан получается после добавления всех связей:

$$\mathcal{H}_{\text{T}} = \int d\boldsymbol{x} \left(-\frac{1}{2}P^{\mu\text{A}}P_{\mu\text{A}} + \frac{1}{4}F^{\mu\nu\text{A}}F_{\mu\nu\text{A}} + \lambda^{\text{A}}P^0{}_{\text{A}} + \mu^{\text{A}}\nabla_\mu P^\mu{}_{\text{A}} \right),$$

где слагаемые $A_0{}^{\text{A}}\nabla_\mu P^\mu{}_{\text{A}}$ мы абсорбировали, переопределив множители Лагранжа μ^{A}. Ему соответствует полное действие

$$S_{\text{T}} = \int dx \left(P^\alpha{}_{\text{A}}\dot A_\alpha{}^{\text{A}} + \frac{1}{2}P^{\mu\text{A}}P_{\mu\text{A}} - \frac{1}{4}F^{\mu\nu\text{A}}F_{\mu\nu\text{A}} - \lambda^{\text{A}}P^0{}_{\text{A}} - \mu^{\text{A}}\nabla_\mu P^\mu{}_{\text{A}} \right). \tag{23.244}$$

Это действие является более общим, чем исходное (23.218), т.к. совпадает с ним только при $\lambda^{\text{A}} = 0$ и $\mu^{\text{A}} = -A_0{}^{\text{A}}$.

Гамильтоновы уравнения движения для действия (23.244) имеют вид:

$$\frac{\delta S_{\text{т}}}{\delta P^0{}_{\text{A}}} = \dot{A}_0{}^{\text{A}} - \lambda^{\text{A}} = 0, \tag{23.245}$$

$$\frac{\delta S_{\text{т}}}{\delta P^{\mu}{}_{\text{A}}} = \dot{A}_{\mu}{}^{\text{A}} + P_{\mu}{}^{\text{A}} + \nabla_{\mu}\mu^{\text{A}} = 0, \tag{23.246}$$

$$\frac{\delta S_{\text{т}}}{\delta A_0{}^{\text{A}}} = -\dot{P}^0{}_{\text{A}} = 0, \tag{23.247}$$

$$\frac{\delta S_{\text{т}}}{\delta A_{\mu}{}^{\text{A}}} = -\dot{P}^{\mu}{}_{\text{A}} + \nabla_{\nu}F^{\nu\mu}{}_{\text{A}} - \mu^{\text{B}}f_{\text{AB}}{}^{\text{C}}P^{\mu}{}_{\text{C}} = 0, \tag{23.248}$$

$$\frac{\delta S_{\text{т}}}{\delta \lambda^{\text{A}}} = -P^0{}_{\text{A}} = 0, \tag{23.249}$$

$$\frac{\delta S_{\text{т}}}{\delta \mu^{\text{A}}} = -\nabla_{\mu}P^{\mu}{}_{\text{A}} = 0. \tag{23.250}$$

Согласно теореме 20.2.3 каждой связи первого рода соответствует калибровочное преобразование, относительно которого полное действие инвариантно. Первичным связям (23.232) соответствует генератор калибровочных преобразований

$$T_1 := \int d\boldsymbol{x}\, \epsilon_1^{\text{A}} P^0{}_{\text{A}}$$

с некоторыми малыми параметрами $\epsilon_1^{\text{A}}(x)$. При этом преобразуются только временны́е компоненты поля Янга–Миллса $A_0{}^{\text{A}}$ и множители Лагранжа λ^{A}:

$$\delta_1 A_0{}^{\text{A}} = [A_0{}^{\text{A}}, T_1] = \epsilon_1^{\text{A}}, \qquad \delta_1\lambda^{\text{A}} = \dot{\epsilon}_1^{\text{A}}. \tag{23.251}$$

Все остальные поля остаются без изменения. Согласно второй теореме Нетер из инвариантности действия относительно калибровочного преобразования (23.251) следует зависимость уравнений движения:

$$\frac{\delta S_{\text{т}}}{\delta A_0{}^{\text{A}}} - \partial_t \frac{\delta S_{\text{т}}}{\delta \lambda^{\text{A}}} = 0.$$

В справедливости этого тождества нетрудно убедиться прямой проверкой.

Вторичным связям первого рода (23.241) также соответствует генератор преобразований

$$T_2 := \int d\boldsymbol{x}\, \epsilon_2^{\text{A}} \nabla_{\mu} P^{\mu}{}_{\text{A}},$$

где $\epsilon_2^{\text{A}}(x)$ – малые параметры второго калибровочного преобразования. В этом случае преобразуются только пространственные компоненты поля Янга–Миллса $A_{\mu}{}^{\text{A}}$, сопряженные импульсы $P^{\mu}{}_{\text{A}}$ и множители Лагранжа μ^{A}:

$$\begin{aligned}
\delta_2 A_{\mu}{}^{\text{A}} &= [A_{\mu}{}^{\text{A}}, T_2] = -\nabla_{\mu}\epsilon_2^{\text{A}},\\
\delta_2 P^{\mu}{}_{\text{A}} &= [P^{\mu}{}_{\text{A}}, T_2] = -\epsilon_2^{\text{B}} f_{\text{AB}}{}^{\text{C}} P^{\mu}{}_{\text{C}},\\
\delta_2 \mu^{\text{A}} &= \dot{\epsilon}_2^{\text{A}} - \epsilon_2^{\text{B}}\mu^{\text{C}} f_{\text{BC}}{}^{\text{A}}.
\end{aligned}$$

Согласно второй теореме Нетер, из локальной инвариантности действия следует зависимость уравнений Эйлера–Лагранжа:

$$\nabla_{\mu}\frac{\delta S_{\text{т}}}{\delta A_{\mu}{}^{\text{A}}} - \frac{\delta S_{\text{т}}}{\delta P^{\mu}{}_{\text{B}}} f_{\text{AB}}{}^{\text{C}} P^{\mu}{}_{\text{C}} - \partial_t \frac{\delta S_{\text{т}}}{\delta \mu^{\text{A}}} - \mu^{\text{B}} f_{\text{AB}}{}^{\text{C}} \frac{\delta S_{\text{т}}}{\delta \mu^{\text{C}}} = 0,$$

что также легко проверить. Таким образом, как и в электродинамике, число параметров калибровочных преобразований полного действия $S_{\text{т}}$ в гамильтоновом формализме по сравнению с лоренцевой формулировкой удвоилось.

Генератором калибровочных преобразований в лагранжевом формализме (**??**) является линейная комбинация связей первого рода:

$$T = \int dx \, \left[(\dot{\epsilon}^{\text{A}} - A_0{}^{\text{B}} \epsilon^{\text{C}} f_{\text{BC}}{}^{\text{A}}) P^0{}_{\text{A}} - \epsilon^{\text{A}} \nabla_\mu P^\mu{}_{\text{A}}) \right].$$

Действительно,

$$\begin{aligned}
\delta A_0{}^{\text{A}} &= [A_0^{\text{A}}, T] = \dot{\epsilon}^{\text{A}} - A_0{}^{\text{B}} \epsilon^{\text{C}} f_{\text{BC}}{}^{\text{A}} = \nabla_0 \epsilon^{\text{A}}, \\
\delta A_\mu{}^{\text{A}} &= [A_\mu{}^{\text{A}}, T] = \nabla_\mu \epsilon^{\text{A}}, \\
\delta P^0{}_{\text{A}} &= \epsilon^{\text{B}} f_{\text{AB}}{}^{\text{C}} P^0{}_{\text{C}}, \\
\delta P^\mu{}_{\text{A}} &= \epsilon^{\text{B}} f_{\text{AB}}{}^{\text{C}} P^\mu{}_{\text{C}}.
\end{aligned}$$

По сравнению с электродинамикой возникли нетривиальные преобразования импульсов. Это связано с тем, что в электродинамике компоненты напряженности инвариантны, а для полей Янга–Миллса – ковариантны.

Поля $A_\mu{}^{\text{A}}$ и $P^\mu{}_{\text{A}}$, достаточно быстро убывающие на бесконечности, для каждого значения индекса A можно разложить на поперечную и продольную части. В отличие от электродинамики это разложение неэффективно, т.к. гамильтоновы уравнения движения "свободного" поля Янга–Миллса (23.245)–(23.250) нелинейны. Это разложение используют при построении теории возмущений в квантовой теории поля.

Решения гамильтоновых уравнений движения содержат произвол, соответствующий множителям Лагранжа λ^{A} и μ^{A}. Для того, чтобы устранить этот произвол необходимо наложить 2N калибровочных условий. В квантовой теории поля обычно используют такие же калибровочные условия (23.94) как и в электродинамике для каждого значения индекса A. Если зафиксировать, например, кулоновскую калибровку $\partial_\mu A^\mu = 0$, то это даст N калибровочных условий. Недостающие N калибровочных условий в гамильтоновом формализме можно получить, рассматривая эволюцию этих калибровочных условий во времени. Из уравнения (23.235) следует

$$\partial_t (\partial_\mu A^\mu{}_{\text{A}}) = -\partial_\mu P^\mu{}_{\text{A}} + \partial^\mu \nabla_\mu A_{0\text{A}}.$$

Тогда кулоновскую калибровку для потенциала можно дополнить N калибровочными условиями

$$-\partial_\mu P^\mu{}_{\text{A}} + \partial^\mu \nabla_\mu A_{0\text{A}} = 0.$$

В этом случае калибровочные условия будут согласованы с гамильтоновыми уравнениями для исходного гамильтониана (23.233). Вместо этих калибровочных можно выбрать другие, например, $A_{0\text{A}} = 0$.

Гамильтонова формулировка поля Янга–Миллса на произвольном многообразии $\mathbb{M}$ с метрикой g проводится, по существу, так же как и в пространстве Минковского. Необходимо произвести замену $\eta_{\alpha\beta} \mapsto g_{\alpha\beta}$ и рассматривать компоненты $g_{\alpha\beta}$ как внешнее поле. Почти все формулы останутся при этом без изменения, только в ковариантных производных появятся символы Кристоффеля, возникающие при интегрировании по частям.

23.5.3 Поля Янга–Миллса в аффинной геометрии

Как уже было отмечено, для построения моделей математической физики на многообразии $\mathbb{M}$ (пространстве-времени) кроме поля Янга–Миллса желательно иметь метрику и аффинную связность, т.е. аффинную геометрию. Поскольку поле Янга–Миллса – это компоненты локальной формы связности, то действие для "свободного" поля Янга–Миллса (23.213) зависит только от метрики, но не от связности.

Уравнения движения и тензор энергии-импульса поля Янга–Миллса получаются простым варьированием действия (23.213):

$$\sqrt{|g|}S_{\text{YM}},^{\alpha}{}_{\text{A}} := \frac{\delta S_{\text{YM}}}{\delta A_{\alpha}{}^{\text{A}}} = \partial_\beta(\sqrt{|g|}F^{\beta\alpha}{}_{\text{A}}) = \sqrt{|g|}\widetilde{\nabla}_\beta F^{\beta\alpha}{}_{\text{A}} = 0, \tag{23.252}$$

$$\sqrt{|g|}S_{\text{YM},\alpha\beta} := \frac{\delta S_{\text{YM}}}{\delta g^{\alpha\beta}} = \frac{1}{2}\sqrt{|g|}T_{\text{YM}\alpha\beta} = -\frac{1}{2}\sqrt{|g|}F_{\alpha\gamma\text{A}}F_\beta{}^{\gamma\text{A}} + \frac{1}{8}\sqrt{|g|}g_{\alpha\beta}F^2, \tag{23.253}$$

где

$$\widetilde{\nabla}_\beta F^{\beta\alpha}{}_{\text{A}} := \partial_\beta F^{\beta\alpha}{}_{\text{A}} + F^{\gamma\alpha}{}_{\text{A}}\widetilde{\Gamma}_{\beta\gamma}{}^{\beta} - A_{\beta\text{A}}{}^{\text{B}}F^{\beta\alpha}{}_{\text{B}}.$$

Слагаемое с символами Кристоффеля $\widetilde{\Gamma}$ возникло при интегрировании по частям (см. раздел 7.6). Отсюда вытекает выражение для тензора энергии-импульса поля Янга–Миллса

$$T_{\text{YM}\alpha\beta} = -F_{\alpha\gamma\text{A}}F_\beta{}^{\gamma\text{A}} + \frac{1}{4}g_{\alpha\beta}F^2. \tag{23.254}$$

Полученное выражение является ковариантным обобщением симметричного канонического тензора энергии-импульса (23.223), полученного в пространстве Минковского из первой теоремы Нетер.

Если поле Янга–Миллса взаимодействует с другими полями, то в правой части уравнений движения (23.252) возникают токи:

$$\widetilde{\nabla}_\beta F^{\beta\alpha}{}_{\text{A}} = J^{\alpha}{}_{\text{A}}. \tag{23.255}$$

Как и в случае электродинамики, для поля Янга–Миллса можно воспользоваться формализмом первого порядка.

Предложение 23.5.2. *Уравнения движения для действия (23.213) эквивалентны уравнениям Эйлера–Лагранжа для действия*

$$S = -\int_{\mathbb{M}} dx\sqrt{|g|}\left[\frac{1}{2}F^{\alpha\beta\text{A}}(\partial_\alpha A_{\beta\text{A}} - \partial_\beta A_\alpha) - \frac{1}{4}F^{\alpha\beta\text{A}}F_{\alpha\beta\text{A}}\right], \tag{23.256}$$

в котором компоненты $F^{\alpha\beta\text{A}}$ и $A_{\alpha\text{A}}$ рассматриваются в качестве независимых переменных.

Доказательство. Повторяет доказательство предложения 23.3.4 для электродинамики. □

Действие для поля Янга–Миллса (23.213) инвариантно относительно общих преобразований координат. Соответствующие вариации полей имеют вид (см. раздел 3.13)

$$\delta A_{\alpha}{}^{\text{A}} = -\partial_\alpha\epsilon^\beta A_\beta{}^{\text{A}} - \epsilon^\beta\partial_\beta A_{\alpha}{}^{\text{A}},$$
$$\delta g^{\alpha\beta} = g^{\alpha\gamma}\partial_\gamma\epsilon^\beta + g^{\beta\gamma}\partial_\gamma\epsilon^\alpha - \epsilon^\gamma\partial_\gamma g^{\alpha\beta}.$$

Отсюда, согласно второй теореме Нетер (19.63), следует зависимость уравнений движения

$$\partial_\beta S,^{\beta\text{A}} A_{\alpha\text{A}} + S,^{\beta\text{A}} F_{\beta\alpha\text{A}} - \partial_\beta S,^{\beta}{}_{\alpha} - \partial_\beta S,_{\alpha}{}^{\beta} - S,_{\beta\gamma}\partial_\alpha g^{\beta\gamma} = 0. \tag{23.257}$$

Или, в ковариантном виде,

$$\widetilde{\nabla}_\beta S,^{\beta\text{A}} A_{\alpha\text{A}} + S,^{\beta\text{A}} F_{\beta\alpha\text{A}} - 2\widetilde{\nabla}^\beta S,_{\beta\alpha} = 0. \tag{23.258}$$

Если выполнены уравнения движения для поля Янга–Миллса (23.252), то с учетом свернутых тождеств Бианки соотношение (23.258) принимает вид

$$\widetilde{\nabla}_\beta T_{\text{YM}\alpha}{}^{\beta} = 0. \tag{23.259}$$

Полученное равенство, выполненное для всех решений уравнений Эйлера–Лагранжа, можно интерпретировать, как ковариантное обобщение закона сохранения тензора энергии-импульса (23.224).

Действие для поля Янга–Миллса инвариантно также относительно калибровочных преобразований (23.201). Этой инвариантности, согласно второй теореме Нетер, соответствует зависимость уравнений движения:

$$\widetilde{\nabla}_\alpha S,^{\alpha}{}_{\text{A}} = 0. \tag{23.260}$$

С учетом уравнения (23.255) это равенство приводит к закону сохранения тока: $\widetilde{\nabla}_\alpha J^\alpha{}_{\text{A}} = 0$. Если ток возникает при варьировании некоторого калибровочно инвариантного действия для полей материи, то сохранение тока будет выполняться автоматически в силу второй теоремы Нетер. Если же ток вводится в уравнения для поля Янга–Миллса "руками", то условие $\widetilde{\nabla}_\alpha J^\alpha{}_{\text{A}} = 0$ необходимо для самосогласованности уравнений.

23.5.4 Поле Янга–Миллса в общей теории относительности

Если ограничиться рассмотрением четырехмерного пространства-времени, то в теории поля Янга–Миллса возникают определенные специфические свойства. Как и в электродинамике поле Янга–Миллса определяет локальные формы связности

$$A^{\text{A}} := dx^\alpha A_\alpha{}^{\text{A}}, \qquad \text{A} = 1, \ldots, \text{N}.$$

Теперь локальная форма кривизны (23.204) не является внешней производной от локальной формы связности. Однако ее можно записать с помощью внешней ковариантной производной, введенной в разделе 14.2,

$$R^{\text{A}} = DA^{\text{A}} := dA^{\text{A}} - \frac{1}{2} A^{\text{B}} \wedge A^{\text{C}} f_{\text{BC}}{}^{\text{A}}. \tag{23.261}$$

Тождества Бианки (23.209) также можно записать в компактном виде с помощью оператора внешнего ковариантного дифференцирования (14.43):

$$DR^{\text{A}} = 0.$$

На этом, собственно, преимущество использования обозначений дифференциальных форм в теории полей Янга–Миллса заканчивается.

В общей теории относительности для каждого значения изотопического индекса A можно ввести 2-формы, дуальные к форме локальной кривизны:

$$*\,F^{\mathrm{A}} := \frac{1}{2} dx^{\alpha} \wedge dx^{\beta} (*F^{\mathrm{A}})_{\alpha\beta}, \tag{23.262}$$

где компоненты дуальной формы кривизны заданы равенством

$$*\,F_{\alpha\beta}{}^{\mathrm{A}} := \frac{1}{2} \varepsilon_{\alpha\beta\gamma\delta} F^{\gamma\delta\,\mathrm{A}}. \tag{23.263}$$

В приведенной формуле $\varepsilon_{\alpha\beta\gamma\delta}(x)$ – полностью антисимметричный тензор четвертого ранга.

Дальнейшая перезапись основных уравнений в точности следует анализу, проведенному для электромагнитного поля в разделе 23.3.6, надо лишь добавить изотопический индекс A там, где это необходимо. В частности, *уравнение (анти-) самодуальности* имеет вид

$$F_{\alpha\beta}{}^{\mathrm{A}} = \pm\, *\, F_{\alpha\beta}{}^{\mathrm{A}}. \tag{23.264}$$

Справедливо также

Предложение 23.5.3. *Если поле Янга–Миллса удовлетворяет уравнению (анти-) самодуальности, то оно также удовлетворяет уравнениям Эйлера–Лагранжа (23.252).*

Решения уравнений (анти-) самодуальности важны в теории полей Янга–Миллса. Для евклидовой сигнатуры метрики соответствующие решения в евклидовом пространстве $\mathbb{R}^4$ называются *(анти-)инстантонами.*

Используя дуальную форму кривизны, тензор энергии-импульса поля Янга–Миллса (23.254) можно переписать в виде

$$T_{\text{ум}\alpha\beta} = -\frac{1}{2} \left[F_{\alpha\gamma\mathrm{A}} F_{\beta}{}^{\gamma\mathrm{A}} + *F_{\alpha\gamma\mathrm{A}} * F_{\beta}{}^{\gamma\mathrm{A}} \right]. \tag{23.265}$$

Он, очевидно, инвариантен относительно дуальных вращений (23.162) для каждого значения индекса A. Если базис алгебры Ли выбран таким образом, что форма Киллинга–Картана совпадает с символом Кронекера, то эту инвариантность можно расширить до группы $\mathbb{SO}(2\mathrm{N})$.

Так же как и в электродинамике, доказывается

Предложение 23.5.4. *Если метрика на многообразии $\mathbb{M}$ имеет лоренцеву сигнатуру и координата x^0 является временем, то временна́я компонента тензора энергии-импульса поля Янга–Миллса*

$$T_{\text{ем}00} = -\frac{1}{2} \left[F_{0\gamma\mathrm{A}} F_{0}{}^{\gamma\mathrm{A}} + *F_{0\gamma\mathrm{A}} * F_{0}{}^{\gamma\mathrm{A}} \right] \tag{23.266}$$

положительно определена и, следовательно, удовлетворяет слабому энергетическому условию. Она также удовлетворяет сильному энергетическому условию.

Ситуация с полем Янга–Миллса во многих отношениях аналогична ситуации в электродинамике.

Предложение 23.5.5. *Действие для полей Янга–Миллса в аффинной геометрии (23.213) инвариантно относительно преобразований Вейля, не затрагивающих само поле Янга–Миллса:*

$$g_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} = \mathsf{e}^{2\phi} g_{\alpha\beta}, \qquad A_\alpha{}^{\mathrm{A}} \mapsto \bar{A}_\alpha{}^{\mathrm{A}} = A_\alpha{}^{\mathrm{A}}, \tag{23.267}$$

где $\phi(x) \in \mathcal{C}^2(\mathbb{M})$ *– произвольная вещественнозначная функция.*

Доказательство. Утверждение следует из равенства

$$\sqrt{|g|}\, g^{\alpha\beta} g^{\gamma\delta} = \sqrt{|\bar{g}|}\, \bar{g}^{\alpha\beta} \bar{g}^{\gamma\delta}. \qquad \square$$

Как следствие общего утверждения 23.3.8, получаем, что след тензора энергии-импульса поля Янга–Миллса (23.254) для четырехмерного пространства-времени равен нулю, $T_{\text{ум}\alpha}{}^{\alpha} = 0$. Это, конечно, легко проверить.

Глава 24

Геометрия поверхностей

Геометрия поверхностей является классическим разделом дифференциальной геометрии. Она относительно проста, и многие вопросы можно решить аналитически, что позволяет глубже понять проблемы, возникающие в более высоких размерностях. Кроме этого геометрия поверхностей в последние годы приобрела особую актуальность в связи с большим числом физических приложений и, в первую очередь, в связи с интересом к теории струн, мировая поверхность которых представляет собой двумерное многообразие с метрикой лоренцевой сигнатуры.

24.1 Геометрия Римана–Картана на поверхности

Минимальная размерность многообразия, на котором возможно существование нетривиального кручения и кривизны, равна двум. Настоящая глава посвящена именно этому случаю.

Пусть на поверхности задана геометрия Римана–Картана, т.е. задана метрика и кручение, а тензор неметричности тождественно равен нулю, $Q_{\alpha\beta\gamma} = 0$. Мы предполагаем, что на поверхности задана метрика g либо лоренцевой, либо евклидовой сигнатуры.

В двумерном случае многие формулы дифференциальной геометрии значительно упрощаются, благодаря наличию антисимметричного (псевдо)тензора второго ранга

$$\varepsilon_{ab} = -\varepsilon_{ba}, \qquad \varepsilon_{01} = \varepsilon_{12} = 1. \tag{24.1}$$

Приставка “псевдо” означает, что при пространственном отражении или обращении времени ε_{ab} меняет знак. В евклидовом случае, конечно, есть только пространственное отражение одной из координат. Напомним, что в псевдоевклидовом и евклидовом случаях латинские индексы принимают значения $\{0, 1\}$ и $\{1, 2\}$, соответственно. Контравариантные компоненты этого тензора зависят от сигнатуры метрики и отличаются знаком

$$\varepsilon^{01} = -1, \qquad \varepsilon^{12} = 1. \tag{24.2}$$

Свойства полностью антисимметричного тензора второго ранга приведены в приложении 28.4. Существование антисимметричного тензора второго ранга означает, что произвольное антисимметричное тензорное поле второго ранга, $X_{ab} = -X_{ba}$, взаимно однозначно определяется псевдоскалярным полем X^*:

$$X_{ab} = \oplus\, \frac{1}{2}\varepsilon_{ab} X^*, \quad \text{где} \quad X^* = \varepsilon^{ab} X_{ab}. \tag{24.3}$$

Здесь и далее верхний знак внутри окружности $\oplus$ или $\ominus$ относится к римановой поверхности, а нижний – к лоренцевой.

Благодаря антисимметрии, тензор кривизны в геометрии Римана–Картана всегда можно представить в виде

$$R_{abcd} = \mathop{\oplus}\frac{1}{2}\varepsilon_{ab}\varepsilon_{cd}R = \frac{1}{2}(\eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc})R, \tag{24.4}$$

где R – скалярная кривизна. Здесь и далее в этом разделе римановы и лоренцевы поверхности будут рассматриваться параллельно. Поэтому символ η_{ab} будет обозначать и символ Кронекера, и метрику Минковского. Из формулы (24.4) следует, что у тензора кривизны только одна неприводимая компонента – скалярная кривизна – отлична от нуля. Соотношение (24.4) справедливо как в римановой геометрии, так и в геометрии Римана–Картана при нетривиальном кручении. В частности, тензор Риччи всегда симметричен

$$R_{ab} = \frac{1}{2}\eta_{ab}R. \tag{24.5}$$

Отсюда следует, что вакуумные уравнения Эйнштейна (21.1) при нулевой космологической постоянной в двумерном пространстве-времени удовлетворяются тождественно.

Все квадратичные инварианты тензора кривизны выражаются через квадрат скалярной кривизны. Например,

$$R_{abcd}R^{abcd} = R^2, \tag{24.6}$$

$$R_{ab}R^{ab} = \frac{1}{2}R^2. \tag{24.7}$$

В двумерном случае упрощается также общий вид тензора кручения. Нетрудно проверить, что он взаимно однозначно представляется либо своим следом:

$$T_{abc} = \eta_{ac}T_b - \eta_{bc}T_a, \qquad T_b := T_{ab}{}^a, \tag{24.8}$$

либо псевдоследом:

$$T_{abc} = \mathop{\oplus}\frac{1}{2}\varepsilon_{ab}T^*_c, \qquad T^*_c := T_{abc}\varepsilon^{ab}. \tag{24.9}$$

Псевдовектор T^*_c и след T_a тензора кручения взаимно однозначно связаны между собой простыми соотношениями:

$$T^*_a = 2\varepsilon_{ab}T^b, \qquad T_a = \frac{1}{2}\varepsilon_{ab}T^{*b}.$$

Квадратичный инвариант тензора кручения имеет вид

$$T_{abc}T^{abc} = 2T_aT^a = \mathop{\oplus}\frac{1}{2}T^*_aT^{*a}. \tag{24.10}$$

При проведении вычислений часто бывает удобно переходить от ковариантной производной $\widetilde{\nabla}_\alpha$, определяемой символами Кристоффеля и возникающей при интегрировании по частям, к ковариантной производной ∇_α, определяемой полной метрической связностью с кручением. Эта связь осуществляется с помощью соотношения

$$\Gamma_{\alpha\beta\gamma} = \widetilde{\Gamma}_{\alpha\beta\gamma} + g_{\alpha\gamma}T_\beta - g_{\alpha\beta}T_\gamma,$$

которое следует из (7.19), (24.8).

Поскольку тождества Бианки антисимметричны по трем индексам, то они автоматически удовлетворяются в двух измерениях и не накладывают никаких ограничений на тензор кривизны и кручения.

Теперь обсудим переменные Картана: репер $e_\alpha{}^a$ и лоренцеву связность $\omega_\alpha{}^{ab}$. Наличие антисимметричного (псевдо)тензора второго ранга позволяет воспользоваться следующей параметризацией лоренцевой связности в геометрии Римана–Картана:

$$\omega_\alpha{}^{ab} = \omega_\alpha \varepsilon^{ab}. \tag{24.11}$$

1-форму ω_α также будем называть лоренцевой связностью. В евклидовом случае вместо термина "лоренцева связность" следует говорить "$\mathbb{SO}(2)$-связность". Для определенности, мы будем говорить лоренцева связность, имея в виду оба случая. При пространственных отражениях или обращении времени эта форма, также как и антисимметричный тензор ϵ^{ab}, меняет знак.

Из определения (24.11) следует, что

$$\omega_a = \bigoplus \frac{1}{2}\omega_{abc}\varepsilon^{bc} = \bigoplus \varepsilon_a{}^b(c_b + T_b), \tag{24.12}$$

где c_b – след коэффициентов неголономности (7.91).

Закон преобразования связности ω_α при локальных лоренцевых вращениях неоднороден (см. раздел **??**). Прямые вычисления показывают, что при вращении на угол $\omega(x)$, лоренцева связность преобразуется так же, как и электромагнитное поле

$$\omega_\alpha \mapsto \omega'_\alpha = \omega_\alpha + \partial_\alpha \omega. \tag{24.13}$$

Тензор кривизны поверхности в переменных Картана представляется в виде

$$R_{\alpha\beta}{}^{ab} = F_{\alpha\beta}\varepsilon^{ab}, \qquad F_{\alpha\beta} := \partial_\alpha \omega_\beta - \partial_\beta \omega_\alpha. \tag{24.14}$$

Отсутствие квадратичных слагаемых в тензоре кривизны связано с тем, что в двумерном случае группы Лоренца $\mathbb{SO}(1,1)$ и вращений $\mathbb{SO}(2)$ являются абелевыми. При этом скалярная кривизна имеет вид

$$R = 2\varepsilon^{\alpha\beta}\partial_\alpha\omega_\beta = \varepsilon^{\alpha\beta}F_{\alpha\beta}, \qquad F_{\alpha\beta} = \bigoplus \frac{1}{2}\varepsilon_{\alpha\beta}R. \tag{24.15}$$

где введен антисимметричный тензор

$$\varepsilon^{\alpha\beta} = e^\alpha{}_a e^\beta{}_b \varepsilon^{ab}, \qquad \nabla_\alpha \varepsilon^{\beta\gamma} = 0. \tag{24.16}$$

Выше введен репер $e_\alpha{}^a$, который в двумерном случае называется *диадой*, и определяется соотношением

$$g_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b \eta_{ab}$$

с точностью до локальных лоренцевых вращений.

Приведем также явное выражение для псевдоследа тензора кручения в картановых переменных, которое часто используется в приложениях,

$$T^{*a} = 2\varepsilon^{\alpha\beta}(\partial_\alpha e_\beta{}^a + \omega_\alpha e_\beta{}^b \varepsilon_b{}^a). \tag{24.17}$$

В римановом пространстве кручение равно нулю, и из уравнения (24.17) следует выражение лоренцевой связности через производные от диады:

$$\widetilde{\omega}_\alpha = -e_\alpha{}^a \varepsilon^{\beta\gamma}\partial_\beta e_{\gamma a}. \tag{24.18}$$

В приложениях часто используется также полностью антисимметричная тензорная плотность единичного веса

$$\hat{\varepsilon}_{\alpha\beta} = \frac{1}{\sqrt{|g|}}\varepsilon_{\alpha\beta}, \qquad \hat{\varepsilon}^{\alpha\beta} = \sqrt{|g|}\varepsilon^{\alpha\beta}. \tag{24.19}$$

Ее компоненты во всех системах координат равны по модулю единице, и частная производная обращается в нуль:

$$\partial_\alpha \hat{\varepsilon}_{\beta\gamma} = \partial_\alpha \hat{\varepsilon}^{\beta\gamma} = 0. \tag{24.20}$$

Определитель диады можно записать в различном виде:

$$\sqrt{|g|} = \det e_\alpha{}^a = e_0{}^a e_1{}^b \varepsilon_{ab} = \bigoplus \hat{\varepsilon}^{\alpha\beta} e_\alpha{}^0 e_\beta{}^1 = \bigoplus \frac{1}{2}\hat{\varepsilon}^{\alpha\beta} e_\alpha{}^a e_\beta{}^b \varepsilon_{ab}. \tag{24.21}$$

Для вычислений, проводимых с мономами третьего и более высокого порядка по лоренцевым векторам, полезны следующие тождества:

$$\varepsilon_{ab}\varepsilon_{cd} + \varepsilon_{bc}\varepsilon_{ad} + \varepsilon_{ca}\varepsilon_{bd} = 0, \tag{24.22}$$

$$\eta_{ab}\varepsilon_{cd} + \eta_{da}\varepsilon_{bc} + \eta_{cd}\varepsilon_{ab} + \eta_{bc}\varepsilon_{da} = 0. \tag{24.23}$$

Тождество (24.22), в котором слагаемые отличаются циклической перестановкой первых трех индексов, следует из того, что в двумерном пространстве не существует тензора, антисимметричного по трем индексам. Тождество (24.23) аналогично тождествам Фирца и проверяется прямой подстановкой.

24.2 Неголономный базис в двух измерениях

В двумерном пространстве-времени Римана–Картана удобно использовать неголономный базис в касательном пространстве, определяемый диадой:

$$e_a := \partial_a := e^\alpha{}_a \partial_\alpha. \tag{24.24}$$

Многие геометрические объекты выражаются через коэффициенты неголономности $c_{ab}{}^c$, которые определяются коммутатором базисных векторных полей,

$$[e_a, e_b] := c_{ab}{}^c e_c. \tag{24.25}$$

Из выражения (24.24) следует, что коэффициенты неголономности имеют вид

$$\begin{aligned} c_{ab}{}^c &= \left(e^\alpha{}_a \partial_\alpha e^\beta{}_b - e^\alpha{}_b \partial_\alpha e^\beta{}_a\right) e_\beta{}^c \\ &= -e^\alpha{}_a e^\beta{}_b \left(\partial_\alpha e_\beta{}^c - \partial_\beta e_\alpha{}^c\right). \end{aligned} \tag{24.26}$$

В двух измерениях они взаимно однозначно определяются следом

$$c_{ab}{}^c = \delta^c_a c_b - \delta^c_b c_a, \qquad c_b := c_{ab}{}^a. \tag{24.27}$$

Из уравнения (24.26) вытекает, что коэффициенты неголономности инвариантны относительно общих преобразований координат. При локальном лоренцевом повороте на угол $\omega(x)$ они преобразуются, как лоренцева связность:

$$\delta c_b = -\omega \varepsilon_b{}^c c_c - \varepsilon_b{}^c \partial_c \omega. \tag{24.28}$$

Приведем два полезных тождества:

$$\varepsilon^{ab}\partial_a e^{\alpha}{}_b = \varepsilon^{\alpha b}c_b, \qquad \varepsilon^{\alpha\beta}\partial_\alpha e_\beta{}^a = -\varepsilon^{ab}c_b. \tag{24.29}$$

Лоренцева связность в неголономном базисе определяется соотношением

$$\omega_a = e_a{}^\alpha \omega_\alpha. \tag{24.30}$$

Скалярная кривизна R и след тензора кручения $T_b := T_{ab}{}^a$ в неголономном базисе имеют вид

$$R = 2\varepsilon^{ab}(\partial_a\omega_b + c_a\omega_b), \tag{24.31}$$
$$T_b = -c_b + \varepsilon_b{}^c\omega_c. \tag{24.32}$$

В геометрии Римана–Картана диада и лоренцева связность являются независимыми переменными. В то же время по заданной диаде всегда можно построить вторую лоренцеву связность и другие геометрические объекты, которые соответствуют нулевому кручению. В этом случае, т.е. в (псевдо)римановой геометрии, лоренцева связность взаимно однозначно определяется коэффициентами неголономности:

$$\tilde{\omega}_a = \varepsilon_a{}^b c_b. \tag{24.33}$$

Разность между двумя лоренцевыми связностями определяется тензором кручения:

$$\omega_a = \tilde{\omega}_a + \epsilon_a{}^b T_b. \tag{24.34}$$

Из уравнения (24.31) следует, что скалярная кривизна в римановой геометрии определяется коэффициентами неголономности следующей простой формулой

$$\tilde{R} = 2\partial_a c^a + 2c_a c^a \tag{24.35}$$

24.3 Выбор системы координат

В двумерном пространстве метрика содержит три независимые компоненты, при этом две из них можно фиксировать за счет выбора системы координат (фиксирования калибровки). Таким образом после фиксирования координат многие геометрические характеристики выражаются через одну функцию, что существенно упрощает анализ двумерных моделей. В настоящем разделе приведен ряд формул в различных системах координат, которые часто встречаются в приложениях.

24.3.1 Диагональная калибровка

В *диагональной калибровке*, которую используют в вычислениях, метрика и ее обратная имеют диагональный вид:

$$g_{\alpha\beta} = \begin{pmatrix} g_{00} & 0 \\ 0 & g_{11} \end{pmatrix}, \qquad g^{\alpha\beta} = \begin{pmatrix} 1/g_{00} & 0 \\ 0 & 1/g_{11} \end{pmatrix}. \tag{24.36}$$

Эта калибровка фиксирует только одну из двух произвольных функций, параметризующих преобразования координат. Оставшийся функциональный произвол можно

использовать для фиксирования других переменных, входящих в модель. Символы Кристоффеля (7.22) для метрики (24.36) имеют следующий вид:

$$\begin{aligned}
\widetilde{\Gamma}_{00}{}^{0} &= \frac{1}{2}\frac{\dot{g}_{00}}{g_{00}}, \qquad \widetilde{\Gamma}_{01}{}^{0} = \widetilde{\Gamma}_{10}{}^{0} = \frac{1}{2}\frac{g'_{00}}{g_{00}}, \qquad \widetilde{\Gamma}_{11}{}^{0} = -\frac{1}{2}\frac{\dot{g}_{11}}{g_{00}}, \\
\widetilde{\Gamma}_{00}{}^{1} &= -\frac{1}{2}\frac{g'_{00}}{g_{11}}, \qquad \widetilde{\Gamma}_{01}{}^{1} = \widetilde{\Gamma}_{10}{}^{1} = \frac{1}{2}\frac{\dot{g}_{11}}{g_{11}}, \qquad \widetilde{\Gamma}_{11}{}^{1} = \frac{1}{2}\frac{g'_{11}}{g_{11}},
\end{aligned} \tag{24.37}$$

Где точка и штрих обозначают дифференцирование по x^0 и x^1 соответственно. У тензора кривизны со всеми опущенными индексами имеется только одна нетривиальная линейно независимая компонента

$$R_{0101} = \frac{1}{2}(\ddot{g}_{11} + g''_{00}) - \frac{1}{4}\left(\frac{\dot{g}_{11}^2}{g_{11}} + \frac{g_{00}'^2}{g_{00}} + \frac{\dot{g}_{00}\dot{g}_{11}}{g_{00}} + \frac{g'_{00}g'_{11}}{g_{11}}\right). \tag{24.38}$$

Остальные отличные от нуля компоненты получаются простой перестановкой индексов. Тензор Риччи в рассматриваемой калибровке диагонален:

$$R_{00} = \frac{1}{g_{11}}R_{0101}, \qquad R_{11} = \frac{1}{g_{00}}R_{0101}. \tag{24.39}$$

Скалярная кривизна равна

$$R = \frac{\ddot{g}_{11} + g''_{00}}{g_{00}g_{11}} - \frac{1}{2g_{00}g_{11}}\left(\frac{\dot{g}_{11}^2}{g_{11}} + \frac{g_{00}'^2}{g_{00}} + \frac{\dot{g}_{00}\dot{g}_{11}}{g_{00}} + \frac{g'_{00}g'_{11}}{g_{11}}\right). \tag{24.40}$$

24.3.2 Конформная калибровка для римановых поверхностей

Для двумерной римановой метрики часто используют *конформную калибровку*, т.е. такую систему координат, в которой метрика является вейлевски евклидовой:

$$ds^2 = \Phi(dx^2 + dy^2), \tag{24.41}$$

где $\Phi = \Phi(x, y)$ – некоторая функция от координат $x = x^1$ и $y = x^2$. *Конформный множитель* Φ предполагается всюду отличным от нуля, иначе метрика будет вырождена, а соответствующая скалярная кривизна может обратиться в бесконечность (это зависит от того, как функция Φ стремится к нулю). Координаты в которых метрика имеет вид (24.41) называются *изотермическими*. Для положительно определенной метрики известна локальная теорема существования изотермических координат.

Теорема 24.3.1. *Если метрика принадлежит классу $\mathcal{C}^3(\mathbb{M})$, тогда для любой точки $x \in \mathbb{M}$ существует окрестность, в которой можно выбрать изотермические координаты.*

Доказательство. См., например, [?]. □

Заметим, что конформный множитель не является скалярным полем, т.к. в других системах координат метрика в общем случае не будет иметь вейлевски евклидова вида.

Если метрика имеет вид (24.41), то удобно ввести комплексные координаты (2.76). Тогда интервал запишется в виде

$$ds^2 = \Phi dz d\bar{z}.$$

После конформного преобразования $z \to \zeta(z)$ интервал сохраняет вейлевски евклидову форму

$$ds^2 = \Phi \frac{dz}{d\zeta}\frac{d\bar{z}}{d\bar{\zeta}} d\zeta d\bar{\zeta}.$$

То есть изотермические координаты определены с точностью до конформных преобразований. По этой причине калибровка (24.41) называется конформной. В изотермических координатах отличны от нуля только два символа Кристоффеля:

$$\widetilde{\Gamma}_{zz}{}^{z} = \frac{\partial_z \Phi}{\Phi}, \qquad \widetilde{\Gamma}_{\bar{z}\bar{z}}{}^{\bar{z}} = \frac{\partial_{\bar{z}} \Phi}{\Phi}.$$

Отличные от нуля линейно независимые компоненты тензора кривизны имеют вид

$$\widetilde{R}_{z\bar{z}\bar{z}}{}^{\bar{z}} = \widetilde{R}_{\bar{z}zz}{}^{z} = \frac{\Phi \partial^2_{z\bar{z}} \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^2} \tag{24.42}$$

У тензора Риччи отличны от нуля только две компоненты:

$$\widetilde{R}_{z\bar{z}} = \widetilde{R}_{\bar{z}z} = \frac{\Phi \partial^2_{z\bar{z}} \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^2}. \tag{24.43}$$

Это приводит к следующему выражению для скалярной кривизны

$$\widetilde{R} = 4\frac{\Phi \partial^2_{z\bar{z}} \Phi - \partial_z \Phi \partial_{\bar{z}} \Phi}{\Phi^3} = \frac{4}{\Phi} \partial^2_{z\bar{z}} \ln|\Phi| \tag{24.44}$$

Часто для конформного множителя используют следующую параметризацию

$$\Phi = \mathrm{e}^{2\phi}. \tag{24.45}$$

Тогда

$$\widetilde{R} = 2\,\mathrm{e}^{-2\phi} \triangle \phi, \tag{24.46}$$

где $\triangle\phi := 4\partial^2_{z\bar{z}}\phi$ – обычный лапласиан.

Для поверхностей постоянной кривизны (см. раздел 25) скалярная кривизна постоянна,

$$\widetilde{R} = -2K = \mathrm{const}, \tag{24.47}$$

где K – гауссова кривизна. Это уравнение в изотермических координатах сводится к *уравнению Лиувилля*

$$\triangle\phi + Ke^{2\phi} = 0. \tag{24.48}$$

Уравнение Лиувилля является интегрируемым, и его общее решение приведено в разделе 25.4.

24.3.3 Конформная калибровка для лоренцевых поверхностей

На псевдоримановой поверхности $\mathbb{M}$ интервал в конформной калибровке имеет вид

$$ds^2 = \Phi(dt^2 - dx^2) = \Phi du dv \tag{24.49}$$

где $\Phi = \Phi(u, v)$ – конформный множитель и u и v – координаты светового конуса (2.82). Соответствующая метрика и ее обратная в конусных координатах имеют вид

$$g_{\alpha\beta} = \frac{\Phi}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad g^{\alpha\beta} = \frac{2}{\Phi}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{24.50}$$

Теорема 24.3.2. *Если метрика дважды дифференцируема $g_{\alpha\beta} \in \mathcal{C}^2(\mathbb{M})$, то каждая точка $\mathbb{M}$ имеет координатную окрестность, в которой метрика является вейлевски лоренцевой (24.49).*

Доказательство. Докажем, что преобразованием координат произвольную обратную метрику можно привести к виду (24.50). Рассмотрим преобразование координат $x^\alpha \mapsto y^\alpha(x)$ в произвольной координатной окрестности. Тогда компоненты обратной метрики преобразуются по правилу

$$g^{\alpha\beta} \mapsto \tilde{g}^{\alpha\beta} = g^{\gamma\delta}\frac{\partial y^\alpha}{\partial x^\gamma}\frac{\partial y^\beta}{\partial x^\delta}.$$

Для упрощения формул введем обозначение

$$g^{\alpha\beta} = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

где $a, b, c, d \in \mathcal{C}^2(\mathbb{M})$. Не ограничивая общности, можно считать, что исходная система координат выбрана таким образом, что $a > 0$. Потребуем, чтобы преобразованные компоненты метрики имели вид (24.50). Тогда функции перехода должны удовлетворять двум уравнениям:

$$\begin{aligned} a\left(\frac{\partial y^0}{\partial x^0}\right)^2 + 2b\frac{\partial y^0}{\partial x^0}\frac{\partial y^0}{\partial x^1} + c\left(\frac{\partial y^0}{\partial x^1}\right)^2 = 0, \\ a\left(\frac{\partial y^1}{\partial x^0}\right)^2 + 2b\frac{\partial y^1}{\partial x^0}\frac{\partial y^1}{\partial x^1} + c\left(\frac{\partial y^1}{\partial x^1}\right)^2 = 0. \end{aligned} \tag{24.51}$$

Эти квадратные уравнения относительно $\partial y^0/\partial x^0$ и $\partial y^1/\partial x^0$ легко решаются. В результате получаем два эквивалентных линейных уравнения:

$$\frac{\partial y^0}{\partial x^0} + \lambda_0\frac{\partial y^0}{\partial x^1} = 0, \quad \frac{\partial y^1}{\partial x^0} + \lambda_1\frac{\partial y^1}{\partial x^1} = 0, \tag{24.52}$$

где

$$\lambda_0 := \frac{b - \sqrt{b^2 - ac}}{a}, \quad \lambda_1 := \frac{b + \sqrt{b^2 - ac}}{a}.$$

Поскольку метрика лоренцева, то $b^2 - ac > 0$.

Предположим, что существуют решения уравнений (24.52) такие, что $\{\partial_\alpha y^0\} \neq 0$ и $\{\partial_\alpha y^1\} \neq 0$. Тогда кривые

$$y^0(x) = C^0, \quad y^1(x) = C^1, \quad C^{0,1} = \mathsf{const}, \tag{24.53}$$

определяют два семейства характеристик (17.51) для волнового уравнения на лоренцевой поверхности. Действительно, уравнения (24.51) это и есть уравнения характеристик.

Для дальнейшего нам понадобится

Лемма 24.3.1. *Пусть функция $W(x) \in \mathcal{C}^2$ такова, что $\partial W/\partial x^1 \neq 0$. Для того, чтобы семейство кривых $W = C = \mathsf{const}$ давало семейство характеристик волнового уравнения (17.49), необходимо и достаточно, чтобы выражение $W = C$ было общим интегралом одного из обыкновенных дифференциальных уравнений*

$$\frac{dx^1}{dx^0} = \lambda_0(x^0, x^1), \quad \frac{dx^1}{dx^0} = \lambda_1(x^0, x^1). \tag{24.54}$$

Доказательство. Из теоремы о неявной функции следует, что уравнение $W = C$ при $\partial_1 W \neq 0$ определяет неявную функцию $x^1 = x^1(x^0, C)$, при этом

$$\frac{dx^1}{dx^0} = -\frac{\partial_0 W}{\partial_1 W}.$$

Если функция $W - C$ является характеристикой волнового уравнения, то она удовлетворяет квадратному уравнению

$$a\partial_0 W^2 + 2b\partial_0 W \partial_1 W + c\partial_1 W^2 = 0,$$

которое приводит к уравнению для неявной функции

$$a\left(\frac{dx^1}{dx^0}\right)^2 - 2b\frac{dx^1}{dx^0} + c = 0,$$

которое сводится к одному из двух уравнений (24.54).

Легко проверить также обратное утверждение. □

Согласно общей теории дифференциальных уравнений решения обыкновенных дифференциальных уравнений (24.54) существуют в, возможно, меньшей окрестности. Согласно доказанной лемме, это означает существование двух семейств характеристик $y^0(x) = C_0$ и $y^1(x) = C_1$ класса $\mathcal{C}^1$. Поскольку $\lambda_{0,1} \in \mathcal{C}^2$, то и $y^{0,1} \in \mathcal{C}^2$.

Далее, из уравнений (24.52) следует неравенство

$$\det \frac{\partial y}{\partial x} := \partial_0 y^0 \partial_1 y^1 - \partial_1 y^0 \partial_0 y^1 = (\lambda_1 - \lambda_0)\partial_1 y^0 \partial_1 y^1 = \frac{2\sqrt{b^2 - ac}}{a}\partial_1 y^0 \partial_1 y^1 \neq 0.$$

Это означает, что преобразование координат $x \mapsto y$ невырождено. □

Замечание. Приведенное доказательство существования системы координат, в которой метрика имеет вейлевски лоренцев вид (24.49), можно обобщить на евклидов случай (см., например, [**?**]). Однако при этом нужно потребовать аналитичность компонент метрики. □

Система координат, в которой метрика является вейлевски лоренцевой, определена неоднозначно. Преобразования светоподобных координат

$$u, v \quad \mapsto \quad u'(u), v'(v), \tag{24.55}$$

которые приводят только к изменению конформного множителя, будем, как и в евклидовом случае, называть конформными.

Выражения для символов Кристоффеля и тензора кривизны аналогичны соответствующим формулам в римановом случае. В конусных координатах отличны от нуля только два символа Кристоффеля:

$$\widetilde{\Gamma}_{uu}{}^{u} = \frac{\partial_u \Phi}{\Phi}, \qquad \widetilde{\Gamma}_{vv}{}^{v} = \frac{\partial_v \Phi}{\Phi}. \tag{24.56}$$

Линейно независимые отличные от нуля компоненты тензора кривизны имеют вид

$$\widetilde{R}_{uvv}{}^{v} = \widetilde{R}_{vuu}{}^{u} = \frac{\Phi\partial^2_{uv}\Phi - \partial_u\Phi\partial_v\Phi}{\Phi^2}. \tag{24.57}$$

У тензора Риччи отличны от нуля только недиагональные компоненты:

$$\widetilde{R}_{uv} = \widetilde{R}_{vu} = \frac{\Phi\partial^2_{uv}\Phi - \partial_u\Phi\partial_v\Phi}{\Phi^2}. \tag{24.58}$$

Это приводит к следующему выражению для скалярной кривизны

$$\widetilde{R} = 4\frac{\Phi\partial^2_{uv}\Phi - \partial_u\Phi\partial_v\Phi}{\Phi^3} = \frac{4}{\Phi}\partial^2_{uv}\ln|\Phi|. \tag{24.59}$$

Если конформный множитель Φ положителен, то его можно параметризовать с помощью экспоненты

$$\Phi = e^{2\phi}, \tag{24.60}$$

где $\phi = \phi(u, v)$ – некоторая функция координат. Тогда скалярная кривизна (24.59) принимает вид

$$\widetilde{R} = 2e^{-2\phi}\Box\phi, \tag{24.61}$$

где $\Box$ – оператор Даламбера на плоскости Минковского (2.112).

Для поверхностей постоянной кривизны (см. раздел 25) скалярная кривизна постоянна,

$$\widetilde{R} = -2K = \mathsf{const}, \tag{24.62}$$

где K – гауссова кривизна. Это уравнение в конформной калибровке сводится к *уравнению Лиувилля*

$$\Box\phi + Ke^{2\phi} = 0. \tag{24.63}$$

Уравнение Лиувилля является интегрируемым, и его общее решение приведено в разделе 25.4.

24.4 Координаты светового конуса

В трех предыдущих разделах мы использовали преобразования координат для уменьшения числа неизвестных компонент метрики. В настоящем разделе мы рассмотрим удобный способ введения базиса в касательном пространстве для *лоренцевой* поверхности, оставляя свободу в выборе координат на многообразии.

Если на поверхности задана метрика лоренцевой сигнатуры, то при проведении вычислений бывает удобно использовать координаты светового конуса для векторов касательного пространства, которые совпадают с координатами светового конуса (2.82) на плоскости Минковского. Преимущество этих координат связано с тем, что локальное лоренцево вращение не перемешивает компоненты векторного поля, т.к. каждая из них преобразуется по неприводимому одномерному представлению связной компоненты единицы группы $\mathbb{SO}(1,1)$.

Преобразование компонент вектора $X = X^a e_a$ относительно ортонормального базиса в касательном пространстве при переходе от декартовых координат к *координатам светового конуса* определим следующим образом:

$$X^{\pm} := \frac{1}{\sqrt{2}}(X^0 \pm X^1), \tag{24.64}$$

Якобиан преобразования $X^0, X^1 \mapsto X^+, X^-$ равен единице. Обратное преобразование имеет вид

$$X^0 = \frac{1}{\sqrt{2}}(X^+ + X^-), \qquad X^1 = \frac{1}{\sqrt{2}}(X^+ - X^-). \tag{24.65}$$

В частности, конусные координаты репера (диады) равны

$$e_\alpha{}^{\pm} = \frac{1}{\sqrt{2}}(e_\alpha{}^0 \pm e_\alpha{}^1). \tag{24.66}$$

Нетрудно проверить, что определитель репера можно записать в виде

$$\sqrt{|g|} = \det e_\alpha{}^a = e_0{}^- e_1{}^+ - e_0{}^+ e_1{}^- = \hat{\varepsilon}^{\alpha\beta} e_\alpha{}^+ e_\beta{}^-, \tag{24.67}$$

что эквивалентно тождеству

$$\varepsilon^{\alpha\beta} e_\alpha{}^+ e_\beta{}^- = 1. \tag{24.68}$$

При проведении расчетов часто приходится обращать диаду. Выпишем соответствующие формулы:

$$\begin{aligned} e^0{}_+ &= -\frac{1}{\sqrt{|g|}} e_1{}^-, \qquad & e^1{}_+ &= \frac{1}{\sqrt{|g|}} e_0{}^- \\ e^0{}_- &= \frac{1}{\sqrt{|g|}} e_1{}^+, \qquad & e^1{}_- &= -\frac{1}{\sqrt{|g|}} e_0{}^+. \end{aligned} \tag{24.69}$$

Скалярное произведение векторов в координатах светового конуса имеет вид

$$X^a Y^b \eta_{ab} = X^- Y^+ + X^+ Y^-. \tag{24.70}$$

Замечание. Важным является то обстоятельство, что лоренцев квадрат вектора,

$$X^2 = 2X^+ X^-, \tag{24.71}$$

линеен по компонентам, а не квадратичен, как в декартовой системе координат. Это помогает в ряде случаев записать явно решение некоторых уравнений, не извлекая квадратного корня. □

Скалярное произведение можно записывать в виде $X^{\pm} Y^{\pm} \eta_{\pm\pm}$, где индекс $\pm = \{+, -\}$ пробегает два значения, а лоренцева метрика и ее обратная в координатах светового конуса имеют одинаковый вид,

$$\eta_{\pm\pm} = \eta^{\pm\pm} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{24.72}$$

Ко- и контравариантные компоненты векторов связаны простыми соотношениями:

$$X^+ = X_-, \qquad X^- = X_+, \tag{24.73}$$

при этом соотношения между ковариантными компонентами такие же, как и для контравариантных векторов (24.64):

$$X_\pm = \frac{1}{\sqrt{2}}(X_0 \pm X_1).$$

Псевдоскаляр, построенный из двух векторов, в координатах светового конуса принимает вид

$$X^a Y^b \varepsilon_{ab} = X^{\pm} Y^{\pm} \varepsilon_{\pm\pm} = X^- Y^+ - X^+ Y^-, \tag{24.74}$$

где

$$\varepsilon_{\pm\pm} = -\varepsilon^{\pm\pm} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{24.75}$$

$$\varepsilon^{\pm}{}_{\pm} = -\varepsilon_{\pm}{}^{\pm} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{24.76}$$

Используя выражение для псевдоследа тензора кручения (24.17), получаем его компоненты в координатах светового конуса

$$T^{*\pm} = 2\varepsilon^{\alpha\beta}(\partial_\alpha \mp \omega_\alpha)e_\beta{}^{\pm}. \tag{24.77}$$

Отсюда следует, что каждая компонента репера $e_\alpha{}^{\pm}$ определяет одну компоненту псевдоследа. Тождество (**??**) в координатах светового конуса принимает вид

$$R = \widetilde{R} + 2\nabla_a T^a - 4T^+T^-, \tag{24.78}$$

где

$$\nabla_\alpha T^{*\pm} = \partial_\alpha T^{*\pm} \mp \omega_\alpha T^{*\pm}.$$

Координаты светового конуса сыграют большую роль в общем решении уравнений движения двумерной гравитации в разделе **??**.

Глава 25

Поверхности постоянной кривизны

Поверхности постоянной кривизны являются относительно простыми объектами, которые хорошо изучены и классифицированы. Универсальными накрывающими для любой полной римановой поверхности постоянной кривизны являются либо сфера $\mathbb{S}^2$, либо евклидова плоскость $\mathbb{R}^2$, либо одна компонента связности двуполостного гиперболоида $\mathbb{H}^2$, соответственно для положительной, нулевой и отрицательной гауссовой кривизны. Все полные римановы поверхности постоянной кривизны получаются из этих трех как фактор пространства по действию свободной и собственно разрывной группы преобразований (теорема 12.2.4).

В случае псевдоримановых поверхностей, метрика которых имеет лоренцеву сигнатуру, можно не различать поверхности постоянной положительной и отрицательной кривизны, т.к. они связаны между собой простым изменением знака метрики: $g_{\alpha\beta} \mapsto -g_{\alpha\beta}$, не меняющим сигнатуру. Для нулевой кривизны, универсальной накрывающей является плоскость Минковского $\mathbb{R}^{1,1}$. Если кривизна отлична от нуля, то универсальной накрывающей поверхности постоянной кривизны является универсальная накрывающая однополостного гиперболоида $\mathbb{L}^2$.

В настоящей главе мы подробно опишем все универсальные накрывающие поверхности постоянной кривизны, за исключением евклидовой плоскости $\mathbb{R}^2$ и двумерного пространства Минковского $\mathbb{R}^{1,1}$, которые достаточно полно были рассмотрены во Введении.

25.1 Сфера $\mathbb{S}^2$

Рассмотрим трехмерное евклидово пространство $\mathbb{R}^3$ с декартовыми координатами x, y, z. Двумерная сфера $\mathbb{S}^2_a$ радиуса $a > 0$ с центром в начале системы координат задается уравнением (см. рис. 25.1, *a*)

$$x^2 + y^2 + z^2 = a^2. \tag{25.1}$$

Сфера, очевидно, является связным и односвязным многообразием.

Вложение сферы в трехмерное евклидово пространство $\mathbb{S}^2_a \hookrightarrow \mathbb{R}^3$, заданное уравнением (25.1), индуцирует на ней положительно определенную метрику с помощью возврата отображения (см. раздел 3.9). Явный вид индуцированной метрики проще всего выглядит в сферической системе координат r, θ, φ (см. раздел 8.1). Чтобы получить его, в выражение для евклидовой метрики в $\mathbb{R}^3$

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$$

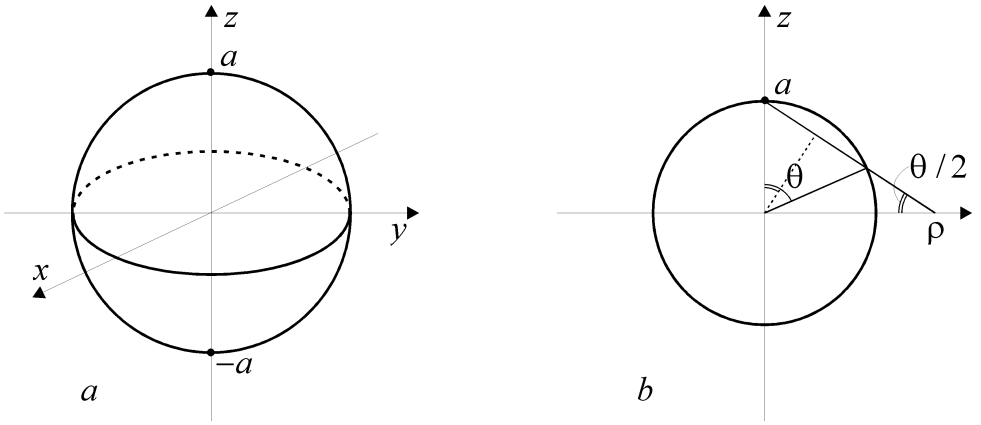


Рис. 25.1: Сфера радиуса a с центром в начале координат (*a)*. Стереографическая проекция точки сферы на плоскость x, y (*b)*.

достаточно подставить уравнение сферы $r = a$. В результате метрика, индуцированная на сфере, в сферической системе координат примет вид

$$d\Omega := a^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{25.2}$$

где координаты на сфере θ и φ называются, соответственно, *азимутальным* и *полярным* углами. Координаты θ, φ покрывают всю сферу, но не определяют глобальную систему координат на сфере, т.к. необходимо провести отождествление, например, $\varphi \sim \varphi + 2\pi$. Соответствующую метрику и ее обратную можно записать в матричных обозначениях:

$$g_{\alpha\beta} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \qquad g^{\alpha\beta} = \frac{1}{a^2} \begin{pmatrix} 1 & 0 \\ 0 & \sin^{-2}\theta \end{pmatrix}.$$

Из символов Кристоффеля (7.22) отличны от нуля только три:

$$\Gamma_{\theta\varphi}{}^{\varphi} = \Gamma_{\varphi\theta}{}^{\varphi} = \frac{\cos\theta}{\sin\theta}, \qquad \Gamma_{\varphi\varphi}{}^{\theta} = -\sin\theta\cos\theta.$$

Тензор кривизны имеет только две нетривиальные линейно независимые компоненты:

$$R_{\theta\varphi\theta}{}^{\varphi} = -1, \qquad R_{\theta\varphi\varphi}{}^{\theta} = \sin^2\theta. \tag{25.3}$$

Тензор кривизны со всеми опущенными индексами имеет только одну независимую нетривиальную компоненту

$$R_{\theta\varphi\theta\varphi} = -a^2\sin^2\theta.$$

Тензор Риччи диагонален:

$$R_{\theta\theta} = -1, \qquad R_{\varphi\varphi} = -\sin^2\theta.$$

Соответствующая скалярная кривизна (7.78) равна константе

$$R = -2K = -\frac{2}{a^2}.$$

Отсюда следует, что согласно нашему определению тензора кривизны, тензора Риччи и скалярной кривизны, скалярная кривизна сферы единичного радиуса равна -2. Поэтому мы также используем *гауссову кривизну* K, которая отличается от скалярной кривизны множителем -2. Для сферы единичного радиуса гауссова кривизна равна единице, $K = 1$.

Экстремали (линии наименьшей длины) на сфере определяются следующей теоремой.

Теорема 25.1.1. *Сечения сферы плоскостями, проходящими через ее центр и только они являются экстремалями. Все экстремали полны.*

Доказательство. См., например, [**?**]. □

Таким образом все экстремали на сфере являются большими окружностями и замкнуты. Ясно, что они полны, т.к. их можно проходить бесконечное число раз. Любые две точки, не расположенные на противоположных полюсах, можно соединить двумя экстремалями. Противоположные полюса соединяются бесконечно большим числом экстремалей.

Метрика евклидова пространства $\mathbb{R}^3$ инвариантна относительно действия группы трехмерных вращений $\mathbb{SO}(3)$. Это означает, что метрика, индуцированная на сфере, допускает по крайней мере три вектора Киллинга по числу параметров группы Ли преобразований. С другой стороны, поверхность постоянной кривизны допускает существование не более трех линейно независимых векторов Киллинга (теорема 16.2.1). Поэтому сфера представляет собой максимально симметричную поверхность постоянной положительной кривизны. Любой вектор Киллинга на сфере представляет собой линейную комбинацию векторов Киллинга, соответствующих вращениям. В декартовой системе координат $\{x^i\} = \{x, y, z\}$, $i = 1, 2, 3$, векторы Киллинга евклидова пространства $\mathbb{R}^3$ имеют вид

$$K_i = \varepsilon_{ij}{}^k x^j \partial_k. \tag{25.4}$$

где ε_{ijk} – полностью симметричный тензор третьего ранга. Они удовлетворяют коммутационным соотношениям (2.119) алгебры $\mathfrak{so}(3)$:

$$[K_i, K_j] = -\varepsilon_{ij}{}^k K_k. \tag{25.5}$$

Используя формулы перехода от декартовых координат к сферическим (8.1), получаем явный вид векторных полей Киллинга на сфере $\mathbb{S}^2_a$ в сферических координатах:

$$\begin{aligned} K_1 &= -\sin\varphi\partial_\theta - \frac{\cos\theta\cos\varphi}{\sin\theta}\partial_\varphi, \\ K_2 &= \quad \cos\varphi\partial_\theta - \frac{\cos\theta\sin\varphi}{\sin\theta}\partial_\varphi, \\ K_3 &= \quad \partial_\varphi. \end{aligned} \tag{25.6}$$

Можно проверить, что они удовлетворяют той же алгебре (25.5), что и векторы Киллинга евклидова пространства.

В приложениях часто используются *стереографические координаты* ρ, φ на сфере $\mathbb{S}^2_a$, в которых индуцированная метрика (25.2) является вейлевски евклидовой. Они строятся следующим образом. Спроектируем точку (θ, φ), не находящуюся на северном полюсе, на плоскость x, y из северного полюса, как показано на рис. 25.1,*b*. При этом точки нижней и верхней полусферы отображаются соответственно на внутренность и дополнение к кругу радиуса a на плоскости x, y. Тогда полярный радиус ρ на плоскости x, y связан с азимутальным углом θ следующим соотношением

$$\frac{a}{\rho} = \operatorname{tg}\frac{\theta}{2} = \frac{1-\cos\theta}{\sin\theta}.$$

Отсюда находим радиус

$$\rho = a\frac{\sin\theta}{1-\cos\theta}. \tag{25.7}$$

Подставляя соотношение между дифференциалами

$$d\rho = -a\frac{1}{1-\cos\theta}d\theta$$

в выражение для индуцированной метрики (25.2), получаем выражение для метрики на сфере в стереографических координатах

$$ds^2 = \frac{4}{(1+\rho^2/a^2)^2}(d\rho^2+\rho^2 d\varphi^2), \quad 0\le\rho<\infty, \quad 0\le\varphi<2\pi. \tag{25.8}$$

Отсюда следует, что в стереографических координатах метрика на сфере является вейлевски евклидовой.

Векторы Киллинга в стереографических координатах имеют вид:

$$\begin{aligned} K_1 &= \frac{\rho^2+a^2}{2a}\sin\varphi\partial_\rho - \frac{\rho^2-a^2}{2a\rho}\cos\varphi\partial_\varphi,\\ K_2 &= -\frac{\rho^2+a^2}{2a}\cos\varphi\partial_\rho - \frac{\rho^2-a^2}{2a\rho}\sin\varphi\partial_\varphi,\\ K_3 &= \partial_\varphi. \end{aligned} \tag{25.9}$$

Стереографические координаты покрывают все точки сферы, за исключением северного полюса, который отображается в бесконечность на плоскости x, y. Аналогично можно построить проекцию точек сферы из южного полюса. Тогда сфера будет покрыта двумя картами.

Рассмотрим плоскость x, y как комплексную плоскость $z = x + iy \in \mathbb{C}$. В комплексных координатах метрика (25.8) примет вид

$$ds^2 = \frac{4}{(1+z\bar z/a^2)^2}dzd\bar z. \tag{25.10}$$

При стереографической проекции северный полюс сферы отображается в бесконечно удаленную точку на комплексной плоскости. Комплексная плоскость, объединенная с бесконечно удаленной точкой, называется *расширенной комплексной плоскостью* и обозначается $\overline{\mathbb{C}}$. Комплексная плоскость $\mathbb{C}\approx\mathbb{R}^2$ и расширенная комплексная плоскость $\overline{\mathbb{C}}\approx\mathbb{S}^2$, как было отмечено в примере 2.2.17, имеют разные топологии. Сфера $\mathbb{S}^2$, на которую взаимно однозначно отображается расширенная комплексная плоскость с помощью стереографической проекции называется *сферой Римана*.

Отображение расширенной комплексной плоскости $\overline{\mathbb{C}}$ на сферу $\mathbb{S}^2$ можно записать в комплексном виде. Обозначим декартовы координаты точки, расположенной на сфере, через ξ, η, t. По-определению, они удовлетворяют уравнению

$$\xi^2+\eta^2+t^2 = a^2,$$

т.е. в качестве координат на сфере можно выбрать координаты ξ и η или их комплексную переменную $\zeta := \xi + i\eta$. Они связаны с комплексными координатами на плоскости $z = x + iy$ соотношением

$$\frac{\zeta}{z} = \frac{a-t}{a}.$$

Исключив из этих соотношений t, получим равенство

$$\zeta = \frac{2z}{1 + z\bar{z}/a^2}.$$

Выпишем также формулы, для координаты t и обратного преобразования $\xi, \eta, t \to z$

$$t = \frac{z\bar{z} - a^2}{z\bar{z} + a^2}a, \qquad z = \frac{a}{a - t}\zeta.$$

25.2 Двуполостный гиперболоид $\mathbb{H}^2$

Рассмотрим трехмерное пространство Минковского $\mathbb{R}^{1,2}$ с декартовыми координатами t, x, y. Метрика Лоренца на этом многообразии определяется квадратичной формой

$$ds^2 = dt^2 - dx^2 - dy^2. \tag{25.11}$$

Двуполостный гиперболоид "радиуса" $a > 0$, вложенный в пространство Минковского $\mathbb{R}^{1,2}$, задается уравнением

$$t^2 - x^2 - y^2 = a^2, \qquad a > 0, \tag{25.12}$$

и показан на рис. 25.2,*a*. Это уравнение имеет решения только при $|t| \geq a$. Верхняя и нижняя полы гиперболоида соответствуют значениям $t \geq a$ и $t \leq -a$. Как многообразие двуполостный гиперболоид состоит из двух компонент связности: верхней и нижней полы. Каждая пола является связным и односвязным многообразием. Для определенности выберем верхнюю полу гиперболоида и обозначим ее $\mathbb{H}^2$. В дальнейшем, говоря про двуполостный гиперболоид, будем подразумевать именно его верхнюю полу $\mathbb{H}^2$. Верхняя (или нижняя) пола двуполостного гиперболоида называется также *плоскостью Лобачевского* или *гиперболической плоскостью*.

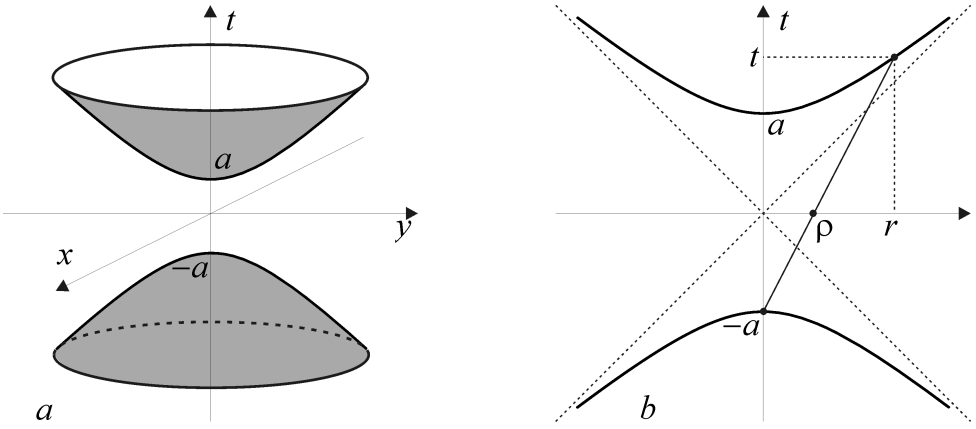


Рис. 25.2: Двуполостный гиперболоид $\mathbb{H}^2$, вложенный в трехмерное пространство Минковского $\mathbb{R}^{1,2}$ (a). Стереографические координаты ρ, φ на двуполостном гиперболоиде (b).

Вложение $\mathbb{H}^2 \hookrightarrow \mathbb{R}^{1,2}$ двуполостного гиперболоида в пространство Минковского (25.12) индуцирует на нем метрику евклидовой сигнатуры (риманову метрику). Это очевидно, поскольку все касательные векторы к поверхности (25.12) пространственноподобны.

Вычисление геометрических характеристик двуполостного гиперболоида наиболее просто проводится в гиперболических координатах:

$$\begin{aligned} x &= r\,\mathsf{sh}\,\theta\,\mathsf{cos}\,\varphi,\\ y &= r\,\mathsf{sh}\,\theta\,\mathsf{sin}\,\varphi,\\ t &= \pm r\,\mathsf{ch}\,\theta, \end{aligned} \tag{25.13}$$

которые мы обозначим теми же буквами r, θ, φ, что и сферические координаты. Преобразования координат (25.13) определены при всех значениях координат r, θ, φ. В приведенных формулах знак $\pm$ в формуле для t относится к гиперболическим координатам, соответственно, в верхнем $t > 0$ и нижнем $t < 0$ полупространствах.

Якобиан преобразования легко вычислить

$$J = \frac{\partial(t, x, y)}{\partial(r, \theta, \varphi)} = \pm r^2\,\mathsf{sh}\,\theta.$$

Таким образом, преобразование координат вырождено при $r = 0$ и $\theta = 0$, т.е. на оси времени t.

Обратные преобразования запишем в следующем виде

$$\begin{aligned} r &= \sqrt{t^2 - x^2 - y^2},\\ \theta &= \pm\,\mathsf{arcth}\,\frac{\sqrt{x^2 + y^2}}{t},\\ \varphi &= \mathsf{arctg}\,\frac{y}{x}. \end{aligned} \tag{25.14}$$

Они определены при $t^2 \geq x^2 + y^2$, и знак $\pm$ относится соответственно к верхней и нижней поле гиперболоида. Мы видим, что внутренность светового конуса будущего, $t > 0$, с вершиной в начале координат и удаленной полуплоскостью $y = 0$, $x \geq 0$ отображается на область

$$0 < r < \infty, \qquad 0 < \theta < \infty, \qquad 0 < \varphi < 2\pi. \tag{25.15}$$

При этом внутренность светового конуса прошлого, $t < 0$, с удаленной полуплоскостью также отображается на область (25.15). Заметим, что полуплоскость необходимо удалить из-за полярного угла $0 < \varphi < 2\pi$.

В гиперболических координатах двуполостный гиперболоид (25.12) задается особенно просто: $r = a = \mathsf{const}$. Поэтому углы θ и φ можно рассматривать в качестве координат на двуполостном гиперболоиде $\mathbb{H}^2$.

Лоренцева метрика (25.11) в гиперболических координатах (25.13) принимает вид

$$ds^2 = dr^2 - r^2(d\theta^2 + \mathsf{sh}^{\,2}\theta\, d\varphi^2).$$

Поэтому индуцированная метрика на двуполостном гиперболоиде $\mathbb{H}^2$, $r = a = \mathsf{const}$, задается квадратичной формой

$$dl^2 = a^2(d\theta^2 + \mathsf{sh}^{\,2}\theta\, d\varphi^2). \tag{25.16}$$

При этом мы изменили общий знак двумерной индуцированной метрики для того, чтобы она стала положительно определенной. Соответствующая метрика и ее обратная в матричных обозначениях имеют вид

$$g_{\alpha\beta} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \mathsf{sh}^{\,2}\theta \end{pmatrix} \qquad g^{\alpha\beta} = a^{-2} \begin{pmatrix} 1 & 0 \\ 0 & \mathsf{sh}^{\,-2}\theta \end{pmatrix}. \tag{25.17}$$

Замечание. Если бы мы рассматривали вложение (25.12) двуполостного гиперболоида в трехмерное евклидово пространство $\mathbb{R}^3$, то индуцированная метрика имела бы другой вид. □

Прямые вычисления показывают, что только три символа Кристоффеля (7.22) отличны от нуля:

$$\Gamma_{\theta\varphi}{}^{\varphi} = \Gamma_{\varphi\theta}{}^{\varphi} = \frac{\mathsf{ch}\,\theta}{\mathsf{sh}\,\theta}, \qquad \Gamma_{\varphi\varphi}{}^{\theta} = -\,\mathsf{sh}\,\theta\,\mathsf{ch}\,\theta.$$

Полный тензор кривизны имеет две отличные от нуля независимые компоненты:

$$R_{\theta\varphi\theta}{}^{\varphi} = 1, \qquad R_{\theta\varphi\varphi}{}^{\theta} = -\,\mathsf{sh}^{\,2}\theta \tag{25.18}$$

При этом тензор кривизны со всеми опущенными индексами имеет только одну отличную от нуля независимую компоненту

$$R_{\theta\varphi\theta\varphi} = a^2\,\mathsf{sh}^{\,2}\theta.$$

Соответствующий тензор Риччи диагонален:

$$R_{\theta\theta} = 1, \qquad R_{\varphi\varphi} = \mathsf{sh}^{\,2}\theta,$$

и скалярная кривизна постоянна

$$\widetilde{R} = -2K = \frac{2}{a^2}.$$

Для гиперболоида единичного радиуса, $a = 1$, гауссова кривизна $K = -1$. Таким образом, двуполостный гиперболоид является поверхностью постоянной отрицательной гауссовой кривизны.

Очевидно, что группа Лоренца $\mathbb{O}(1,2)$, действующая глобально в пространстве Минковского $\mathbb{R}^{1,2}$, является группой симметрии и лоренцевой метрики (25.11), и двуполостного гиперболоида (25.12). Поэтому группой симметрии индуцированной метрики на двуполостном гиперболоиде (25.16) является группа Ли псевдоортогональных вращений (группа Лоренца) $\mathbb{SO}(1,2)$. Соответствующая алгебра $\mathfrak{so}(1,2)$ в пространстве Минковского $\mathbb{R}^{1,2}$ задается тремя векторами Киллинга:

$$\begin{aligned} K_0 &= -x\partial_y + y\partial_x = -\partial_\varphi, \\ K_1 &= \;\; y\partial_t + t\partial_y = \;\; \sin\varphi\partial_\theta + \frac{\mathsf{ch}\,\theta\cos\varphi}{\mathsf{sh}\,\theta}\partial_\varphi, \\ K_2 &= -t\partial_x - x\partial_t = -\cos\varphi\partial_\theta + \frac{\mathsf{ch}\,\theta\sin\varphi}{\mathsf{sh}\,\theta}\partial_\varphi, \end{aligned} \tag{25.19}$$

которые мы записали в декартовых и гиперболических координатах. Векторы Киллинга в гиперболических координатах не содержат компоненты вдоль ∂_r. Это означает, что они касательны к двуполостному гиперболоиду. Коммутационные соотношения для векторных полей Киллинга имеют вид

$$[K_0, K_1] = K_2, \quad [K_2, K_0] = K_1, \quad [K_1, K_2] = -K_0. \tag{25.20}$$

Отметим, что инверсия времени $t \to -t$ преобразует $K_{1,2} \to -K_{1,2}$ и не меняет K_0. При этом алгебра Ли полей Киллинга остается неизменной.

Введем обозначение $\{K_i\} = \{K_0, K_1, K_2\}$, $i = 0, 1, 2$. Тогда алгебру $\mathfrak{so}(1,2)$ можно записать в компактном виде:

$$[K_i, K_j] = -\varepsilon_{ij}{}^k K_k,$$

где ε_{ijk} – полностью антисимметричный тензор третьего ранга, $\varepsilon_{012} = 1$, а подъем индексов, в отличие от алгебры вращений $\mathfrak{so}(3)$, осуществляется с помощью обратной метрики Минковского $\eta^{ij} = \mathsf{diag}\,(+--)$.

Поскольку в общем случае поверхность допускает не более трех векторов Киллинга, то двуполостный гиперболоид представляет собой максимально симметричную поверхность.

Замечание. Внимательный читатель, конечно, заметил, что формулы для двуполостного гиперболоида и сферы, рассмотренной в разделе 25.1, очень похожи. Соответствующие выражения отличаются частью знаков и заменой тригонометрических функций от θ на гиперболические. □

Группа $\mathbb{SO}(1,2)$ некомпактна и ее алгебра Ли $\mathfrak{so}(1,2)$ имеет нетривиальную неабелеву подалгебру, натянутую на векторные поля K_1 и $K_0 + K_2$. Действительно, из коммутационных соотношений (25.20) следует равенство

$$[K_1, K_0 + K_2] = -(K_0 + K_2). \tag{25.21}$$

Соответствующая двумерная неабелева группа Ли подробно рассмотрена в разделе 9.7, где $L_x = -\alpha K_1$, $L_y = K_0 + K_2$. Отметим, что алгебра вращений $\mathfrak{so}(3)$ не имеет неабелевых подгрупп.

Легко найти интегральные кривые полей Киллинга в пространстве Минковского $\mathbb{R}^{1,2}$. Например, интегральными кривыми поля K_2 в плоскости t, x являются гиперболы

$$t^2 - x^2 = \pm c^2, \qquad c \in \mathbb{R}.$$

Нетрудно проверить, что двуполостные гиперболоиды, параметризуемые "радиусом" a, являются интегральными подмногообразиями полей Киллинга (25.19).

Нетрудно проверить, что векторы Киллинга линейно зависимы:

$$tK_0 + xK_1 + yK_2 = 0. \tag{25.22}$$

С точки зрения теоремы Фробениуса 3.11.1 векторные поля K_1 и $K_0 + K_2$ линейно независимы и определяют распределение в пространстве Минковского $\mathbb{R}^{1,2}$. Они находятся в инволюции и, следовательно, определяют семейство интегральных подмногообразий, которыми являются двуполостные гиперболоиды.

Двуполостный гиперболоид является римановым многообразием постоянной отрицательной гауссовой кривизны. Выше было показано, что эта поверхность естественным образом вкладывается в трехмерное пространство Минковского $\mathbb{R}^{1,2}$, что связано с $\mathbb{SO}(1,2)$ симметрией метрики (25.16). Можно доказать, что не существует $\mathcal{C}^2$ вложения полной поверхности постоянной отрицательной гауссовой кривизны в трехмерное евклидово пространство [**?**]. В то же время в классе $\mathcal{C}^1$ вложение полной поверхности постоянной отрицательной гауссовой кривизны уже возможно [**?**].

Теперь обсудим экстремали.

Теорема 25.2.1. *Сечения двуполостного гиперболоида $\mathbb{H}^2$ плоскостями, проходящими через начало координат, и только они является экстремалями. Все экстремали на двуполостном гиперболоиде полны.*

Доказательство. Уравнения для экстремалей интегрируются в явном виде. □

Из рис. 25.2,*a* ясно, что все экстремали незамкнуты. Через две произвольные точки на гиперболоиде $\mathbb{H}^2$ и начало координат проходит одна и только одна плоскость. Это означает, что две произвольные точки на гиперболоиде можно соединить единственной экстремалью. При преобразовании Лоренца плоскость переходит в плоскость. Поэтому экстремаль переходит в экстремаль. Это следовало ожидать, поскольку метрика двуполостного гиперболоида (25.16) инвариантна относительно преобразований Лоренца.

Для того, чтобы записать метрику двуполого гиперболоида (25.16) в вейлевски евклидовом виде, введем стереографические координаты ρ, φ. Это делается аналогично введению стереографических координат на сфере. Сначала введем обычные полярные координаты r, φ на плоскости x, y. Выберем точку $(t, r, \varphi) \in \mathbb{H}^2$ на верхней полé гиперболоида и спроектируем ее из вершины нижней полы $x = y = 0$, $t = -a$ на плоскость x, y. На плоскости x, y ей будет соответствовать точка с координатами ρ, φ (см. рис. 25.2,*b*). Из рисунка ясно, что справедливо соотношение

$$\frac{r}{\rho} = \frac{t+a}{a}.$$

Подставив сюда время $t = \sqrt{a^2 + r^2}$ из уравнения гиперболоида (25.12), получим связь радиусов:

$$r = \frac{2a^2\rho}{a^2 - \rho^2}.$$

Это приводит к соотношению дифференциалов

$$dr = 2a^2 \frac{a^2 + \rho^2}{(a^2 - \rho^2)^2} d\rho.$$

Чтобы записать метрику двуполостного гиперболоида в стереографических координатах, запишем ее сначала в полярных координатах (после изменения общего знака):

$$ds^2 = \frac{a^2}{a^2 + r^2} dr^2 + r^2 d\varphi^2. \tag{25.23}$$

Для получения этой формулы надо в евклидовой метрике (25.11) произвести замену координат $x, y \mapsto r, \varphi$ и подставить выражение для времени $t = \sqrt{a^2 + r^2}$. Замена $r \mapsto \rho$ приводит к метрике в стереографических координатах

$$ds^2 = \frac{4}{(1 - \rho^2/a^2)^2} (d\rho^2 + \rho^2 d\varphi^2). \tag{25.24}$$

Эта метрика является вейлевски евклидовой и определена при

$$0 \leq \rho < a, \qquad 0 \leq \varphi < 2\pi.$$

Она отличается от метрики сферы (25.8) только знаком перед ρ^2, и этим определяется различие в области определения стереографических координат для сферы $\mathbb{S}^2$ и двуполостного гиперболоида $\mathbb{H}^2$.

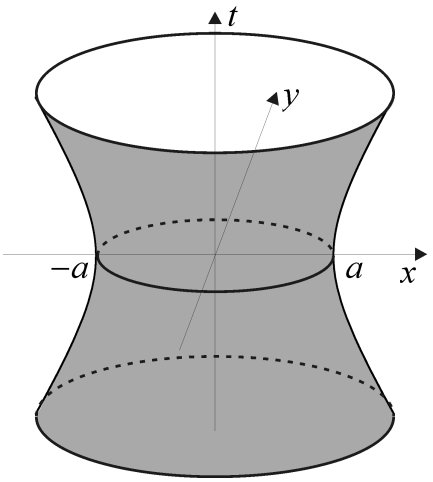


Рис. 25.3: Однополостный гиперболоид $\mathbb{L}^2$, вложенный в трехмерное пространство Минковского $\mathbb{R}^{1,2}$.

25.3 Однополостный гиперболоид $\mathbb{L}^2$

Как и раньше, будем считать, что в трехмерном пространстве Минковского $\mathbb{R}^{1,2}$ заданы декартовы координаты t, x, y и метрика Лоренца (25.11). Рассмотрим *однополостный гиперболоид* $\mathbb{L}^2$, вложенный в пространство-время Минковского $\mathbb{R}^{1,2}$. Он задается уравнением (см. рис. 25.3).

$$t^2 - x^2 - y^2 = -a^2, \qquad a > 0, \tag{25.25}$$

которое отличается от уравнения для двуполостного гиперболоида (25.12) знаком правой части. Это уравнение имеет решения только при $x^2 + y^2 \geq a^2$.

С топологической точки зрения однополостный гиперболоид является связным, но не односвязным многообразием. Его фундаментальная группа совпадает с фундаментальной группой окружности, $\pi(\mathbb{L}^2) = \mathbb{Z}$, а универсальная накрывающая, очевидно, диффеоморфна евклидовой плоскости $\mathbb{R}^2$.

Из рисунка ясно, что в любой точке гиперболоида касательное пространство содержит как времениподобные, так и пространственноподобные векторы по отношению к метрике (**??**). Это значит, что вложение $\mathbb{L}^2 \hookrightarrow \mathbb{R}^{1,2}$ индуцирует на поверхности $\mathbb{L}^2$ метрику лоренцевой сигнатуры.

Для проведения вычислений удобно воспользоваться гиперболическими координатами r, θ, φ:

$$\begin{aligned} x &= r\,\mathsf{ch}\,\theta\cos\varphi, \\ y &= r\,\mathsf{ch}\,\theta\sin\varphi, \\ t &= r\,\mathsf{sh}\,\theta. \end{aligned} \tag{25.26}$$

Эти формулы определены при всех значениях r, θ, φ. Якобиан преобразования координат (25.26) равен

$$J = \frac{\partial(t, x, y)}{\partial(r, \theta, \varphi)} = r^2\,\mathsf{ch}\,\theta.$$

Таким образом, преобразование координат вырождено только в начале координат $r = 0$. Поскольку якобиан преобразования координат неотрицателен, то переход к новым координатам сохраняет ориентацию.

Обратные преобразования запишем в виде

$$\begin{aligned} r &= \sqrt{x^2+y^2-t^2}, \\ \theta &= \operatorname{arccth}\frac{\sqrt{x^2+y^2}}{t}, \\ \varphi &= \operatorname{arctg}\frac{y}{x}. \end{aligned} \tag{25.27}$$

Они определены при $x^2+y^2 \geq t^2$. Мы видим, что внешность светового конуса с вершиной в начале координат и удаленной полуплоскостью $y=0$, $x \geq 0$ отображается на область

$$0 < r < \infty, \qquad -\infty < \theta < \infty, \qquad 0 < \varphi < 2\pi.$$

В гиперболических координатах однополостный гиперболоид определяется уравнением $r = a = \operatorname{const}$. Поэтому углы θ и φ можно выбрать в качестве координат на однополостном гиперболоиде $\mathbb{L}^2$.

Метрика Лоренца (25.11) в гиперболических координатах (25.27) принимает вид

$$ds^2 = -dr^2 + r^2(d\theta^2 - \operatorname{ch}^2\theta d\varphi^2).$$

Следовательно, индуцированная метрика на однополостном гиперболоиде $r = a = \operatorname{const}$ имеет вид

$$ds^2 = a^2(d\theta^2 - \operatorname{ch}^2\theta d\varphi^2). \tag{25.28}$$

Она имеет, очевидно, лоренцеву сигнатуру.

В матричных обозначениях метрика и ее обратная имеют вид

$$g_{\alpha\beta} = a^2\begin{pmatrix} 1 & 0 \\ 0 & -\operatorname{ch}^2\theta \end{pmatrix} \qquad g^{\alpha\beta} = a^{-2}\begin{pmatrix} 1 & 0 \\ 0 & -\operatorname{ch}^{-2}\theta \end{pmatrix}. \tag{25.29}$$

Из символов Кристоффеля только три отличны от нуля:

$$\Gamma_{\theta\varphi}{}^{\varphi} = \Gamma_{\varphi\theta}{}^{\varphi} = \frac{\operatorname{sh}\theta}{\operatorname{ch}\theta}, \qquad \Gamma_{\varphi\varphi}{}^{\theta}\operatorname{sh}\theta\operatorname{ch}\theta.$$

У полного тензора кривизны только две линейно независимые компоненты отличны от нуля:

$$R_{\theta\varphi\theta}{}^{\varphi} = 1, \qquad R_{\theta\varphi\varphi}{}^{\theta} = \operatorname{ch}^2\theta.$$

Тензор кривизны со всеми опущенными индексами имеет только одну линейно независимую компоненту

$$R_{\theta\varphi\theta\varphi} = -a^2\operatorname{ch}^2\theta.$$

Тензор Риччи диагонален:

$$R_{\theta\theta} = 1, \qquad R_{\varphi\varphi} = -\operatorname{ch}^2\theta.$$

Наконец, скалярная кривизна равна

$$R = -2K = \frac{2}{a^2}.$$

Таким образом, однополостный гиперболоид $\mathbb{L}^2$ является поверхностью постоянной отрицательной гауссовой кривизны, $K = -1/a^2$. Чтобы получить поверхность постоянной положительной гауссовой кривизны с метрикой лоренцевой сигнатуры,

можно рассмотреть тот же однополостный гиперболоид, но вложенный в трехмерное пространство с метрикой

$$ds^2 = -dt^2 + dx^2 + dy^2,$$

что приведет к изменению знака метрики на поверхности, а, значит, и знака скалярной кривизны.

Метрика трехмерного пространства Минковского (25.11) и уравнение гиперболоида (25.25) инвариантны относительно трехмерной группы лоренцевых вращений $\mathbb{SO}(1,2)$. Это означает, что преобразования из этой группы являются изометриями. Группа Лоренца $\mathbb{SO}(1,2)$ трехмерна, и ее действие в пространстве Минковского $\mathbb{R}^{1,2}$ задается тремя векторами Киллинга, которые в декартовых координатах имеют тот же вид, что и в случае двуполостного гиперболоида (25.19):

$$\begin{aligned} K_0 &= -x\partial_y + y\partial_x, \\ K_1 &= y\partial_t + t\partial_y, \\ K_2 &= -t\partial_x - x\partial_t. \end{aligned} \tag{25.30}$$

Конечно, они удовлетворяют той же алгебре Ли $\mathfrak{so}(1,2)$ (25.20), поскольку пространство Минковского $\mathbb{R}^{1,2}$ не изменилось. Изменится только их вид в гиперболических координатах:

$$\begin{aligned} K_0 &= -\partial_\varphi, \\ K_1 &= \sin\varphi\partial_\theta + \frac{\operatorname{sh}\theta\cos\varphi}{\operatorname{ch}\theta}\partial_\varphi, \\ K_2 &= -\cos\varphi\partial_\theta + \frac{\operatorname{sh}\theta\sin\varphi}{\operatorname{ch}\theta}\partial_\varphi. \end{aligned} \tag{25.31}$$

Поскольку максимальное число полей Киллинга для поверхности равно трем, то для однополостного гиперболоида векторы Киллинга (25.31) определяют максимальный набор изометрий.

Как и в предыдущем разделе, векторы Киллинга (25.31) на однополостном гиперболоиде линейно зависимы (25.22). Векторы K_1 и $K_0 + K_2$ образуют подалгебру в $\mathfrak{so}(1,2)$ и определяют двумерное инволютивное распределение в пространстве Минковского $\mathbb{R}^{1,2}$. Согласно теореме Фробениуса 3.11.1 это распределение определяет семейство интегральных подмногообразий, которыми являются однополостные гиперболоиды.

Теперь исследуем экстремали (17.17) на однополостном гиперболоиде $\mathbb{L}^2$.

Теорема 25.3.1. *Сечения однополостного гиперболоида $\mathbb{L}^2$ плоскостями, проходящими через начало координат, и только они является экстремалями. Все экстремали на однополостном гиперболоиде полны.*

Доказательство. См., например, в [**?**]. □

Обозначим угол между секущей плоскостью, проходящей через начало координат, и плоскостью x, y через ψ. В зависимости от величины этого угла все экстремали делятся на три класса:

$$\begin{aligned} \psi \in [0, \pi/4) \quad &- \text{ пространственноподобные}, \\ \psi = \pi/4 \quad &- \text{ светоподобные}, \\ \psi \in (\pi/4, \pi/2] \quad &- \text{ времениподобные}. \end{aligned}$$

Пространственноподобные экстремали замкнуты, а остальные – нет.

Через любые две различные точки на гиперболоиде $\mathbb{L}^2$ и начало координат можно провести одну и только одну плоскость. Это значит, что две произвольные точки однополостного гиперболоида можно соединить по крайней мере одной экстремалью. Если экстремали, соединяющие две точки пространственноподобны, то таких экстремалей две. В остальных случаях экстремаль, соединяющая две точки, единственна.

При преобразовании Лоренца плоскость переходит в плоскость. Следовательно, экстремаль отображается в экстремаль.

Вычислим длину пространственноподобной замкнутой экстремали. Рассмотрим плоскость

$$t = kx, \qquad 0 < k < 1,$$

которая параллельна оси y и определяет пространственноподобную экстремаль. В гиперболических координатах это уравнение принимает вид

$$\mathsf{th}\,\theta = k\cos\varphi.$$

Отсюда следует равенство для дифференциалов

$$d\theta = -\frac{k\sin\varphi}{1-k^2\cos^2\varphi}d\varphi.$$

Подставив это выражение в метрику (25.28), получим квадрат элемента длины пространственноподобной экстремали

$$ds^2 = -a^2\frac{1-k^2}{(1-k^2\cos^2\theta)^2}d\varphi^2.$$

Чтобы вычислить длину экстремали, надо извлечь квадрат и проинтегрировать:

$$L := a\sqrt{1-k^2}\int_0^{2\pi}\frac{d\varphi}{1-k^2\cos^2\varphi} = 2\pi a.$$

Поскольку задача симметрична относительно вращений вокруг оси t, то отсюда следует, что длина произвольной пространственноподобной экстремали не зависит от угла, под которым плоскость пересекает двуполостный гиперболоид, и равна длине окружности, по которой гиперболоид пересекает плоскость x, y.

Теперь обсудим *стереографические координаты*, для однополостного гиперболоида, в которых она является конформно плоской. Их введение в данном случае не так наглядно, как для сферы, и сложнее. Определим стереографические координаты τ, σ для гиперболоида, как координаты точки на плоскости t, x, которая является центральной проекцией точки гиперболоида из точки $t = x = 0$, $y = -a$.

Чтобы явно построить функции перехода требуется некоторое геометрическое построение. Пусть точка на гиперболоиде имеет координаты t, x, y. По-определению, они удовлетворяют равенству (25.25), т.е. в качестве координат на гиперболоиде можно выбрать, например, декартовы координаты t, x, которые покрывают половину гиперболоида. Соединим эту точку с точкой $t = x = 0$, $y = -a$ прямой. Соответствующие проекции на плоскости t, y и x, y показаны на рис.25.4 *a* и *b*. Из рисунков следуют равенства:

$$\frac{t}{\tau} = \frac{y+a}{a}, \qquad \frac{x}{\sigma} = \frac{y+a}{a}. \tag{25.32}$$

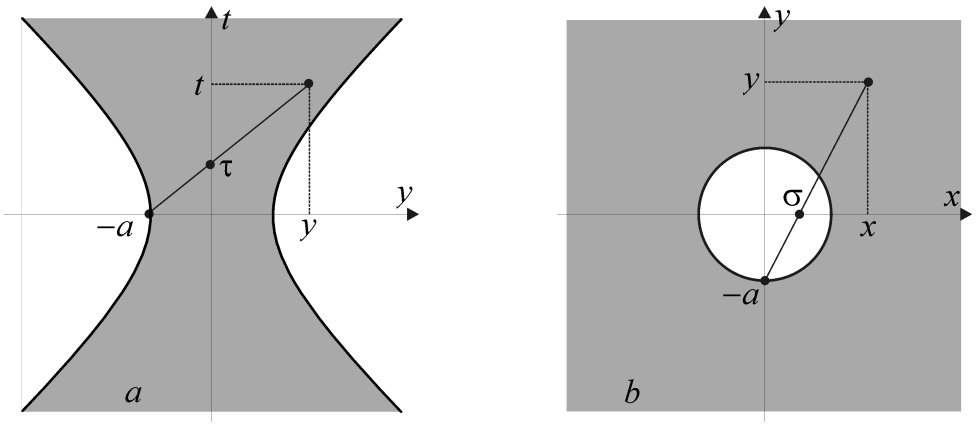


Рис. 25.4: Стереографические координаты τ, σ на однополостном гиперболоиде. Проекции на плоскости t, y (a) и x, y (b).

Из этих уравнений выразим t и x и подставим в уравнение гиперболоида (25.25). В результате возникнет квадратное уравнение на y, которое можно решить:

$$y = a\frac{a^2 + \tau^2 - \sigma^2}{a^2 - \tau^2 + \sigma^2},$$

где мы выбрали больший корень. Тогда

$$\frac{y + a}{a} = \frac{2a^2}{a^2 - \tau^2 + \sigma^2}.$$

Теперь уравнения (25.32) дают функции перехода:

$$t = \tau\frac{2a^2}{a^2 - \tau^2 + \sigma^2}, \qquad x = \sigma\frac{2a^2}{a^2 - \tau^2 + \sigma^2}. \tag{25.33}$$

Якобиан преобразования координат $\tau, \sigma \mapsto t, x$ легко вычислить

$$J = \frac{\partial(t, x)}{\partial(\tau, \sigma)} = 4a^4\frac{a^2 + \tau^2 - \sigma^2}{(a^2 - \tau^2 + \sigma^2)^3}.$$

Он равен нулю и обращается в бесконечность, соответственно, на гиперболах $\tau^2 - \sigma^2 = -a^2$ и $\tau^2 - \sigma^2 = a^2$ и имеет разный знак по разные стороны гипербол. Таким образом, стереографические координаты пробегают всю плоскость $-\infty < \tau, \sigma < \infty$ за исключением четырех ветвей гипербол.

На гиперболах $\tau^2 - \sigma^2 = -a^2$ из формул преобразования координат (25.33) следуют равенства $t = \tau$ и $x = \sigma$, и, следовательно, $t^2 - x^2 = -a^2$. То есть гиперболы $\tau^2 - \sigma^2 = -a^2$ соответствует сечениям гиперболоида плоскостью $y = 0$.

Отображение однополостного гиперболоида на плоскость τ, σ показано на рис.25.5. Для наглядности изображен гиперболоид конечной высоты. Мы предполагаем, что его верхняя и нижняя окружности бесконечно удалены. Также для наглядности мы сжали плоскость τ, σ вдоль светоподобных направлений

$$\xi = \tau + \sigma, \qquad \eta = \tau - \sigma, \tag{25.34}$$

и изобразили ее в виде квадрата. Этого можно добиться с помощью невырожденной замены координат, например,

$$\xi \mapsto \operatorname{arctg}\xi, \qquad \eta \mapsto \operatorname{arctg}\eta.$$

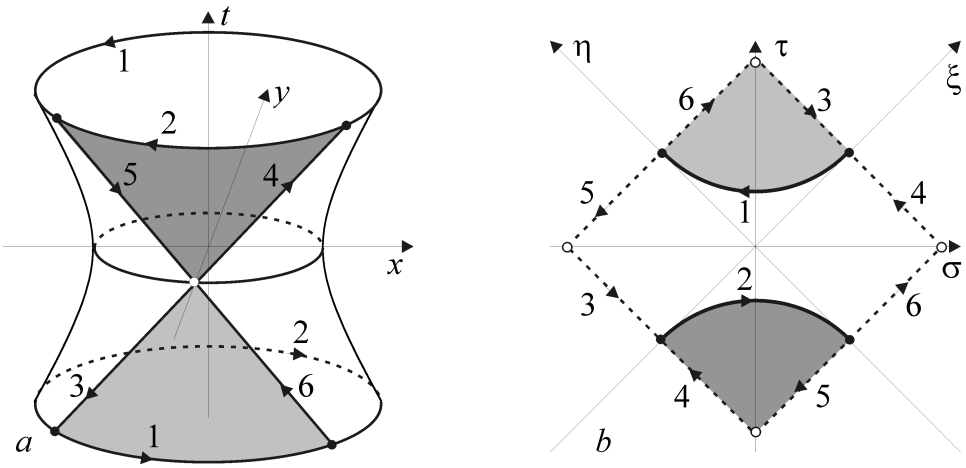


Рис. 25.5: Стереографические координаты на однополостном гиперболоиде. Центральная проекция гиперболоида с координатами t, x (a). Стереографические координаты τ, σ (b).

Мы также приняли следующие обозначения. Закрашенные черным кружки обозначают точки, находящиеся на бесконечном расстоянии, а полые – на конечном. Точке $t = x = 0$, $y = -a$, из которой ведется проектирование, соответствуют четыре незакрашенные вершины квадрата. Кривые между выделенными точками на гиперболоиде и их образы обозначены одинаковыми цифрами. Стрелки показывают выбранную ориентацию кривых.

При отображении на плоскость τ, σ весь гиперболоид разбивается на три области, ограниченные кривыми (1,5,3,2,6,4), (2,5,4) и (1,6,3). То есть гиперболоид разрезается вдоль двух прямых $t = \pm x$, $y = -a$. Центральной точке $t = x = 0$, $y = -a$ при стереографической проекции соответствуют четыре точки – вершины квадрата – и их необходимо отождествить. Проще всего проследить отображение областей, рассмотрев отображение их границ. Область, ограниченная кривыми (1,5,3,2,6,4) отображается на среднюю часть плоскости τ, σ между гиперболами $\tau^2 - \sigma^2 = a^2$. Области, ограниченные кривыми (2,5,4) и (1,6,3), отображаются, соответственно, в области на плоскости τ, σ, расположенные ниже гиперболы $\tau = -\sqrt{a^2 + \sigma^2}$ и выше гиперболы $\tau = \sqrt{a^2 + \sigma^2}$.

Каждая точка гиперболоида из областей (1,5,3,2,6,4), (2,5,4) и (1,6,3) имеет единственные стереографические координаты. При этом две светоподобные прямые $t^2 = x^2,\ y = -a$ (линии 3,4 и 5,6 на рис. 25.5) отображаются в бесконечность на плоскости τ, σ. Точки прямых $t^2 = x^2,\ y = -a$ лежат на конечном расстоянии, и поэтому стороны квадрата на плоскости τ, σ нарисованы пунктиром. Напротив, бесконечно удаленная граница гиперболоида отображается на гиперболы $\tau^2 - \sigma^2 = a^2$, и поэтому нарисованы жирной сплошной линией.

Противоположные стороны квадрата на τ, σ плоскости лежат на конечном расстоянии и отображаются в одну прямую на гиперболоиде. Это означает, что их необходимо отождествить (за исключением бесконечно удаленных точек $\xi = 0$, $\eta = \pm\infty$ и $\xi = \pm\infty$, $\eta = 0$).

Чтобы получить метрику однополостного гиперболоида в стереографических координатах, сначала запишем ее в координатах t, x. Это легко сделать путем исключения координаты y из метрики Лоренца (25.11) с помощью уравнения (25.25). В

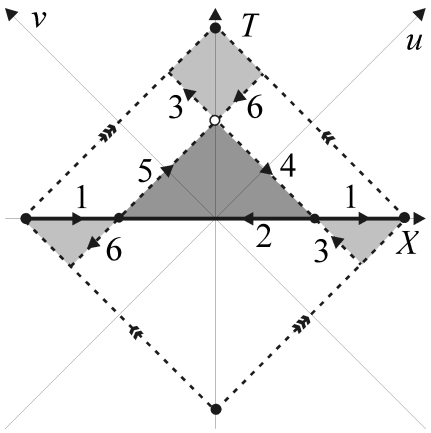


Рис. 25.6: Координаты на однополостном гиперболоиде, в которых метрика зависит только от времени.

результате метрика примет вид

$$ds^2 = \frac{(a^2-x^2)dt^2 + 2txdtdx - (a^2+t^2)dx^2}{a^2+t^2-x^2}. \tag{25.35}$$

Подставив сюда выражения для дифференциалов из формул преобразования координат (25.33), мы получим, что в стереографических координатах метрика однополостного гиперболоида является вейлевски лоренцевой:

$$ds^2 = \frac{4}{\left(1-\frac{\tau^2-\sigma^2}{a^2}\right)^2}(d\tau^2 - d\sigma^2). \tag{25.36}$$

В светоподобных координатах (25.34) метрика примет вид

$$ds^2 = \frac{4}{\left(1-\xi\eta/a^2\right)^2}d\xi d\eta. \tag{25.37}$$

Произведем еще одну замену координат так, чтобы интервал зависел только от времениподобной координаты. В этой системе координат двумерное пространство-время является однородным и анализ экстремалей значительно упрощается. Сначала с помощью дробно линейного преобразования преобразуем светоподобные координаты $\xi, \eta \mapsto u, v$:

$$\xi := a\frac{u+a}{u-a}, \qquad \eta := a\frac{v+a}{v-a}. \tag{25.38}$$

На рис. 25.6 показан однополостный гиперболоид в новых координатах. Цифрами обозначены образы соответствующих кривых на рис. 25.3. Ось абсцисс нарисована жирной линией, что означает ее полноту (бесконечную удаленность). Верхняя и нижняя вершины квадрата также лежат на бесконечном расстоянии и показаны закрашенными кружками. Противоположные стороны квадрата (за исключением вершин) лежат на конечном расстоянии и показаны пунктиром. Их необходимо отождествить, т.к. они соответствуют одним и тем же точкам гиперболоида. Все четыре вершины квадрата также отождествляются. Центральная точка стереографической проекции $t = x = 0$, $y = -a$ (четыре вершины квадрата на рис. 25.3,*b*) отображается в точку $u = v = a$. Следуя стрелкам, легко проследить отображение областей (1,5,3,2,6,4), (2,5,4) и (1,6,3).

Из формул преобразования координат (25.38) вытекают формулы преобразования дифференциалов:

$$d\xi = -2a^2\frac{du}{(u-a)^2},$$
$$d\eta = -2a^2\frac{dv}{(v-a)^2}.$$

Якобиан преобразования координат сингулярен на линиях $u = a$ и $v = a$:

$$\frac{\partial(\xi,\eta)}{\partial(u,v)} = \frac{4a^2}{(u-a)^2(v-a)^2}.$$

В новых светоподобных координатах интервал принимает вид

$$ds^2 = \frac{4a^2}{(u+v)^2}dudv. \tag{25.39}$$

Теперь снова введем временну́ю T и пространственную X координаты с помощью соотношений:

$$u := T + X, \qquad v := T - X.$$

Тогда метрика однополостного гиперболоида примет вид

$$ds^2 = \frac{a^2}{T^2}(dT^2 - dX^2). \tag{25.40}$$

Она является вейлевски лоренцевой, а соответствующее пространство – однородным.

При решении уравнений движения в различных моделях гравитации мы получаем для метрики некоторое выражение, например, формулу (25.36) или (25.40). То есть мы имеем плоскость τ, σ или T, X с заданной на ней метрикой. Восстановить по этой метрике глобальную структуру пространства-времени (однополостный гиперболоид) довольно сложно, но можно. В разделе **??** будет описан общий подход к этой проблеме, основанный на анализе экстремалей. Если окажется, что экстремали для найденного решения уравнений движения неполны, то решение необходимо продолжить и восстановить глобальную структуру. Это можно сделать с помощью метода конформных блоков. Процедура восстановления поверхности по виду метрики, допускающей один вектор Киллинга, конструктивна, и в результате возникает диаграмма Картера–Пенроуза, представляющая глобальную структуру пространства-времени.

25.4 Уравнение Лиувилля

Рассмотрим евклидову плоскость $\mathbb{R}^2$ с декартовыми координатами x, y или, что эквивалентно, комплексную плоскость $\mathbb{C}$ с координатами $z := x + iy$ и $\bar{z} := x - iy$. Пусть задана вещественнозначная функция $\phi(z, \bar{z}) \in \mathcal{C}^\infty(\mathbb{C})$. Тогда *уравнением Лиувилля* для функции ϕ называется нелинейное уравнение второго порядка в частных производных

$$\triangle\phi + K\,\mathsf{e}^{2\phi} = 0, \qquad K = \mathsf{const}, \tag{25.41}$$

где $\triangle := 4\partial^2_{z\bar{z}}$ – оператор Лапласа.

Уравнение Лиувилля часто встречается в моделях математической физики и является одним из немногих нелинейных явно интегрируемых уравнений. Общее решение уравнения Лиувилля в односвязной области имеет вид [**?**]

$$\mathsf{e}^{2\phi} = \frac{4F'\bar{F}'}{(aF\bar{F} + bF + \bar{b}\bar{F} + d)^2}, \tag{25.42}$$

где $F = F(z)$ – произвольная голоморфная функция, введены обозначения:

$$F' := \partial_z F, \quad \bar{F} := \bar{F}(\bar{z}), \quad \bar{F}' := \partial_{\bar{z}}\bar{F},$$

и произвольные постоянные $a, d \in \mathbb{R}$ и $b \in \mathbb{C}$ удовлетворяют одному условию

$$ad - b\bar{b} = K. \tag{25.43}$$

При этом, если $K = 0$, то по крайней мере одна из постоянных a, b, или d должна быть отлична от нуля. Мы также предполагаем, что $F' \neq 0$.

Таким образом, произвольное решение уравнения Лиувилля параметризуется одной голоморфной функцией $F(z)$, двумя вещественными (a, d) и одной комплексной (b) постоянными, на которые наложено одно условие (25.42).

Если задана поверхность $\mathbb{M}$ с римановой метрикой $g \in \mathcal{C}^3(\mathbb{M})$, то согласно теореме 24.3.1 локально существует изотермическая система координат, в которой метрика является вейлевски евклидовой:

$$ds^2 = \mathsf{e}^{2\phi} dz d\bar{z}.$$

Если поверхность имеет постоянную кривизну, то ее гауссова кривизна также постоянна, $K = \mathsf{const}$. Это условие в изотермических координатах сводится к уравнению Лиувилля (25.41) (см. раздел 24.3.3). Отсюда вытекает следующая

Теорема 25.4.1. *Любая поверхность с римановой метрикой $g \in \mathcal{C}^3(\mathbb{M})$ постоянной кривизны локально изометрична сфере $\mathbb{S}^2$ $(K > 0)$, евклидовой плоскости $\mathbb{R}^2$ $(K = 0)$ или двуполостному гиперболоиду $\mathbb{H}^2$ $(K < 0)$.*

Доказательство. При умножением метрики на положительную постоянную k, $g_{\alpha\beta} \mapsto kg_{\alpha\beta}$, гауссова кривизна делится на ту же постоянную, $K \mapsto K/k$. Поэтому достаточно рассмотреть три случая $K = 1, 0, -1$. Выберем произвольную точку на поверхности. Тогда в некоторой окрестности этой точки существуют изотермические координаты. Следовательно, для конформного множителя возникнет уравнение Лиувилля. Рассмотрим три случая.

$K = 1$. Положим в общем решении уравнения Лиувилля

$$d = 1, \quad a = 1, \quad b = 0, \quad F = z.$$

Тогда решение (25.42) примет вид

$$ds^2 = \frac{4}{(1 + z\bar{z})^2} dz d\bar{z}.$$

Это – метрика сферы единичного радиуса (25.10).

$K = 0$. Положим

$$d = 1, \quad a = 0, \quad b = 0, \quad F = z/2.$$

Тогда получим евклидову метрику $ds^2 = dzd\bar{z}$.

$K = -1$. Положим

$$d = 1, \quad a = -1, \quad b = 0, \quad F = z.$$

Тогда метрика примет вид

$$ds^2 = \frac{4}{(1 - z\bar{z})^2} dzd\bar{z}.$$

Это – метрика (25.24) двуполостного гиперболоида единичного "радиуса" в комплексных координатах. □

Следствие. Риманова метрика поверхности постоянной кривизны бесконечно дифференцируема. □

Доказательство. Поскольку функция $F(z)$ голоморфна, то функция (25.42) бесконечно дифференцируема. □

Доказанная теорема локальна. Согласно теореме 12.3.3 каждая поверхность имеет единственную с точностью до диффеоморфизма универсальную накрывающую поверхность. Поскольку поверхности $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$ связны и односвязны, то они являются универсальными накрывающими поверхностями для всех поверхностей, соответственно, положительной, нулевой и отрицательной кривизны. Кроме того, они полны, т.е. любую экстремаль можно продолжить до бесконечного значения канонического параметра в обе стороны. Чтобы найти все полные поверхности постоянной кривизны, для поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$ необходимо найти все группы преобразований $\mathbb{G}$, действующие собственно разрывно и свободно. Тогда все полные поверхности постоянной кривизны будут являться фактор пространствами $\widetilde{\mathbb{M}}/\mathbb{G}$, где $\widetilde{\mathbb{M}} = \mathbb{S}^2, \mathbb{R}^2, \mathbb{H}^2$ (см. главу 12).

Теперь рассмотрим уравнение Лиувилля на плоскости Минковского $\mathbb{R}^{1,1}$. Обозначим декартовы координаты через t, x. Введем также конусные координаты

$$u := t + x, \qquad v := t - x.$$

Тогда аналогом уравнения Лиувилля (25.41) будет являться нелинейное уравнение в частных производных

$$\Box\phi + Ke^{2\phi} = 0, \qquad K = \mathsf{const}. \tag{25.44}$$

где $\Box := 4\partial^2_{uv}$ – оператор Даламбера на плоскости Минковского (2.112). Это уравнение также будем называть *уравнением Лиувилля*.

Уравнение Лиувилля на плоскости Минковского также явно интегрируется. Его общее решение имеет вид

$$e^{2\phi} = \frac{4F'G'}{(aFG + bF + cG + d)^2}, \tag{25.45}$$

где $F = F(u)$ и $G = G(v)$ – произвольные дифференцируемые монотонные функции от светоподобных координат u и v, соответственно, $F'G' \neq 0$. Четыре вещественные постоянные a, b, c и d, должны удовлетворять одному условию

$$ad - bc = K. \tag{25.46}$$

Если $K = 0$, то хотя бы одна из постоянных a, b, c или d должна быть отлична от нуля.

Рассмотрим псевдориманову поверхность. Пусть метрика дважды непрерывно дифференцируема. Тогда согласно теореме 24.3.2 в некоторой окрестности произвольной точки поверхности существует система координат, в которой метрика является вейлевски лоренцевой (24.49). Если поверхность имеет постоянную кривизну, то для конформного множителя возникает уравнение Лиувилля (24.63).

Теорема 25.4.2. *Любая псевдориманова поверхность с метрикой $g \in \mathcal{C}^2(\mathbb{M})$ постоянной кривизны локально изометрична однополостному гиперболоиду $\mathbb{L}^2$ $(K \neq 0)$ или плоскости Минковского $\mathbb{R}^{1,1}$ $(K = 0)$.*

Доказательство. Для лоренцевых поверхностей изменение знака метрики меняет знак гауссовой кривизны. С другой стороны, это эквивалентно перестановке местами времени и пространства. Поэтому можно не различать псевдоримановы поверхности положительной и отрицательной кривизны. Так же, как и для римановых поверхностей метрику можно умножить на положительную постоянную. Следовательно, достаточно рассмотреть два случая $K = 1$ и $K = 0$.

$K = 1$. Положим в решении уравнения Лиувилля

$$d = 1, \quad a = 1, \quad b = c = 0, \quad F = u, \quad G = v.$$

Тогда метрика примет вид

$$ds^2 = \frac{4}{(1+uv)^2} du dv.$$

Это – метрика однополостного гиперболоида (25.37) единичного "радиуса" (с точностью до замены общего знака метрики).

$K = 0$. Положим

$$d = 1, \quad a = 0, \quad b = c = 0, \quad F = u/2, \quad G = v/2.$$

Тогда получим метрику Лоренца $ds^2 = du dv$. □

Доказанная теорема локальна. Как и в случае римановых поверхностей любая полная псевдориманова поверхность является фактор пространством $\widetilde{\mathbb{M}}/\mathbb{G}$, где $\widetilde{\mathbb{M}} = \mathbb{L}^2, \mathbb{R}^{1,1}$ для поверхностей, соответственно, ненулевой и нулевой гауссовой кривизны, а $\mathbb{G}$ – группа преобразований, действующая собственно разрывно и свободно.

Произвольные функции $F(u)$ и $G(v)$, входящие в решение (25.45) уравнения Лиувилля, соответствуют конформным преобразованиям светоподобных координат (24.55), в которых метрика является вейлевски лоренцевой. Поскольку $F' \neq 0$ и $G' \neq 0$, то всегда можно совершить конформное преобразование координат $u, v \mapsto F, G$, в которых метрика примет вид

$$ds^2 = e^{2\phi} du dv = \frac{4 dF dG}{(aFG + bF + cG + d)^2}. \tag{25.47}$$

Область определения решения уравнения Лиувилля (25.45) зависит от выбора функций F и G, а также от значения постоянных a, b, c и d.

При $K > 0$, положим в решении (25.45)

$$F = u, \quad G = v, \quad a = d = 0, \quad b = -c = \sqrt{K}.$$

Тогда метрика примет вид

$$ds^2 = \frac{4 du dv}{K(u-v)^2} = \frac{1}{K} \frac{dt^2 - dx^2}{x^2}. \tag{25.48}$$

Эта метрика определена либо в правой, $x > 0$, либо в левой, $x < 0$, полуплоскости.

Аналогично, при $K < 0$ положим

$$F = u, \quad G = v, \quad a = d = 0, \quad b = c = \sqrt{|K|}.$$

Тогда метрика примет вид

$$ds^2 = \frac{4dudv}{|K|(u+v)^2} = \frac{1}{|K|}\frac{dt^2 - dx^2}{t^2}. \tag{25.49}$$

В таком виде метрика определена в верхней, $t > 0$, или нижней, $t < 0$, полуплоскости. Она уже была получена ранее (25.40).

Два последних выражения для метрики связаны между собой перестановкой координат $t \leftrightarrow x$ и изменением знака метрики $g_{\alpha\beta} \to -g_{\alpha\beta}$. Это означает, что поверхности постоянной положительной и отрицательной кривизны, как уже отмечалось, в случае псевдоримановой метрики эквивалентны.

Замечание. Модели математической физики, как правило, помимо метрики включают также другие поля. Поэтому может случиться так, что суммарное действие для всех полей не будет инвариантно относительно изменения знака метрики. В этом случае знак гауссовой кривизны является существенным. □

При любом значении гауссовой кривизны K мы можем положить

$$F = \frac{u}{2}, \quad G = \frac{v}{2}, \quad a = K, \quad d = 1, \quad b = c = 0.$$

Тогда метрика примет вид

$$ds^2 = \frac{dudv}{(1 + \frac{K}{4}uv)^2}.$$

Эта метрика определена на всей плоскости за исключением двух ветвей гиперболы $uv = -4/K$.

Положим $F = u$ и $G = v$ в общем решении уравнения Лиувилля (25.45). Тогда метрика примет вид

$$ds^2 = \frac{4dudv}{(auv + bu + cv + d)^2}. \tag{25.50}$$

Предложение 25.4.1. *При дробно линейном преобразовании светоподобных координат*

$$u \mapsto u' := \frac{\alpha_1 u + \beta_1}{\gamma_1 u + \delta_1}, \qquad v \mapsto v' := \frac{\alpha_2 v + \beta_2}{\gamma_2 v + \delta_2}, \tag{25.51}$$

где $\alpha_1, \dots, \delta_1, \alpha_2, \dots, \delta_2$ – некоторые постоянные, такие что $\det_1 := \alpha_1\delta_1 - \beta_1\gamma_1 \neq 0$ и $\det_2 := \alpha_2\delta_2 - \beta_2\gamma_2 \neq 0$, метрика (25.50) принимает вид

$$ds^2 = \frac{4 \det_1 \det_2 \, du'dv'}{(a'u'v' + b'u' + c'v' + d')^2}, \tag{25.52}$$

где

$$\begin{aligned}
a' &= a\alpha_1\alpha_2 + b\alpha_1\gamma_2 + c\gamma_1\alpha_2 + d\gamma_1\gamma_2,\\
b' &= a\alpha_1\beta_2 + b\alpha_1\delta_2 + c\gamma_1\beta_2 + d\gamma_1\delta_2,\\
c' &= a\beta_1\alpha_2 + b\beta_1\gamma_2 + c\delta_1\alpha_2 + d\delta_1\gamma_2,\\
d' &= a\beta_1\beta_2 + b\beta_1\delta_2 + c\delta_1\beta_2 + d\delta_1\delta_2.
\end{aligned}$$

Доказательство. Прямые вычисления. □

Дробно линейные преобразования (25.51) образуют подгруппу группы конформных преобразований координат (24.55). Если $\det_1 = \det_2 = 1$, то соответствующие дробно линейные преобразования образуют подгруппу, которая сохраняет вид метрики (25.50). Эти преобразования можно использовать для фиксирования постоянных a, b, c и d.

Глава 26

Поверхности с одним вектором Киллинга

(Псевдо-) римановы поверхности с одним вектором Киллинга часто встречаются при построении глобальных решений в различных моделях гравитации. Дадим

Определение. Пара $(\mathbb{M}, g)$, где $\mathbb{M}$ – многообразие и g – заданная на нем метрика, называется *глобальным решением* в данной модели гравитации, если метрика удовлетворяет соответствующим уравнениям движения на $\mathbb{M}$, а само многообразие $\mathbb{M}$ является максимально продолженным. □

Напомним, что максимальное продолжение означает, что любую экстремаль можно либо продолжить в обе стороны до бесконечного значения канонического параметра, либо при конечном значении канонического параметра она попадает в сингулярную точку, где один из геометрических инвариантов становится сингулярным. Это определение инвариантно и не зависит от выбора системы отсчета, т.к. канонический параметр инвариантен и определен с точностью до линейного преобразования.

Пример 26.0.1. Нетривиальным глобальным решением в общей теории относительности является расширение Крускала–Секереша [**?**, **?**] решения Шварцшильда [**?**] (независимо найденного также Дросте [**?**]). В этом случае четырехмерное пространство-время представляет собой топологическое произведение сферы $\mathbb{S}^2$ на псевдориманову поверхность с одним вектором Киллинга, которая обычно изображается в виде диаграммы Картера–Пенроуза, которая будет описана в настоящем разделе. □

В настоящей главе развит конструктивный метод конформных блоков построения глобальных решений $(\mathbb{M}, g)$ для двумерных метрик лоренцевой и евклидовой сигнатуры специального вида, обладающих одним вектором Киллинга.

Необходимость построения глобальных решений, т.е. одновременного построения и самих многообразий и заданных на них метрик, связана с инвариантностью моделей гравитации относительно общих преобразований координат. Действительно, для решения уравнений движения, как правило, фиксируется некоторая система координат, что уменьшает число искомых функций. Но, с другой стороны, фиксирование системы отсчета означает, что, за исключением тривиальных случаев, найденное решение является только локальным и представляет собой часть некоторого бóльшего пространства-времени. Отметим также, что только глобальная структура многообразия позволяет дать физическую интерпретацию решений уравнений движения для метрики.

Построение глобальных решений в гравитации является трудной задачей, поскольку, помимо решения уравнений движения, что само по себе сложно, предполагает решение уравнений для экстремалей, анализ их полноты и продолжения многообразия, если оно оказалось неполным. В общей теории относительности из-за сложности уравнений движения известно лишь небольшое число глобальных решений (см. главу 27). В двумерных моделях гравитации ситуация проще и удалось построить все глобальные решения [?] в двумерной гравитации с кручением [?], а также в широком классе моделей дилатонной гравитации [?, ?, ?]. Конструктивный метод конформных блоков построения глобальных решений для псевдоримановых поверхностей был предложен в [?] для двумерной гравитации с кручением в конформной калибровке. Для поверхностей с римановой метрикой метод конформных блоков был предложен в [?].

Замечание. В настоящей главе мы рассматриваем продолжение решений только вдоль экстремалей. В аффинной геометрии с кручением и неметричностью геодезические и экстремали в общем случае различны. Поэтому имеет смысл рассматривать максимальное продолжение многообразий как вдоль экстремалей, так и вдоль геодезических. Если геодезические и экстремали совпадают как, например, в общей теории относительности, то продолжение решения вдоль экстремалей автоматически влечет за собой продолжение вдоль геодезических. □

26.1 Локальный вид лоренцевой метрики

Рассмотрим плоскость $\mathbb{R}^2$ с декартовыми координатами $\{\zeta^\alpha\} = \{\tau, \sigma\}$, $\alpha = 0, 1$. Пусть в некоторой области на плоскости задана метрика лоренцевой сигнатуры, которая имеет вейлевски плоский вид

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2). \tag{26.1}$$

Будем считать, что конформный множитель $\Phi(q) \in \mathcal{C}^l$, $l \geq 2$, является l раз непрерывно дифференцируемой функцией одного аргумента $q \in \mathbb{R}$ за исключением конечного числа особенностей. Этого достаточно для того, чтобы компоненты метрики (26.1) могли удовлетворять некоторой системе уравнений второго порядка вне особых точек. Пусть аргумент q зависит только от одной из координат, τ или σ, и эта связь задается обыкновенным дифференциальным уравнением

$$\left|\frac{dq}{d\zeta}\right| = \pm\Phi(q), \tag{26.2}$$

где выполняется следующее правило знаков:

$$\begin{array}{ll} \Phi > 0: & \zeta = \sigma, \quad \text{знак } + \text{ (статическое локальное решение)}, \\ \Phi < 0: & \zeta = \tau, \quad \text{знак } - \text{ (однородное локальное решение)}. \end{array} \tag{26.3}$$

На первый взгляд двумерная метрика (26.1) имеет очень специфический вид. Однако, это – метрика достаточно общего вида. В дальнейшем будет показано, что некоторые модели гравитации при решении полной системы уравнений движения приводят к двумерной метрике вида (26.1). Например, при отыскании сферически симметричных вакуумных решений уравнений Эйнштейна в четырехмерном пространстве-времени (см. следующую главу), метрика (26.1) возникает на двумерных поверхностях, точки которых параметризуются временем и радиусом. Метрика вида (26.1)

возникает также при решении уравнений движения двумерной гравитации. По этой причине в настоящей главе мы будем, для краткости, называть метрику вида (26.1) просто локальным решением.

Мы допускаем, что конформный множитель метрики (26.1) может обращаться в нуль и иметь особенности в конечном числе точек, которые мы пронумеруем в порядке возрастания: q_i, $i = 1, \ldots, k$. В эту последовательность включены также бесконечно удаленные точки $q_1 = -\infty$ и $q_k = \infty$. Будем считать, что вблизи каждой точки q_i конформный множитель ведет себя степенным образом:

$$|q_i| < \infty : \qquad \Phi(q) \sim (q - q_i)^m, \tag{26.4}$$

$$|q_i| = \infty : \qquad \Phi(q) \sim q^m. \tag{26.5}$$

При конечных значениях q_i показатель степени не равен нулю, $m \neq 0$, т.к. конформный множитель в этой точке, по-предположению, либо равен нулю, либо сингулярен.

Строго говоря, показатель m различен для различных точек, и мы должны были бы писать m_i, а не просто m. Однако эта упрощенная запись не приводит к недоразумениям, поскольку при анализе полноты экстремалей будет рассматриваться одна конкретная точка q_i. Показатель степени m может быть любым вещественным числом, для которого правые части (26.4), (26.5) определены.

В промежутках между нулями и особенностями, где функция Φ или положительна, или отрицательна, локальное решение, соответственно, статично с вектором Киллинга $K = \partial_\tau$ или однородно с вектором Киллинга $K = \partial_\sigma$. Квадрат длины вектора Киллинга в обоих случаях равен конформному множителю, $K^2 = \Phi$. На горизонтах, которые, как будет показано ниже, соответствуют нулям конформного множителя, $\Phi(q) = 0$, вектор Киллинга является светоподобным. При этом он отличен от нуля, что будет продемонстрировано в координатах Эддингтона–Финкельстейна (см. раздел 26.6).

При конформных преобразованиях вид метрики (26.1), (26.2) меняется, т.к. переменная q становится зависящей от обеих координат на плоскости. При растяжке координат $\zeta \mapsto \tilde{\zeta} = k\zeta$, $k = \mathsf{const} \neq 0$, которые образуют подгруппу группы конформных преобразований, метрика сохраняет свой вид с $\tilde{q} = k^{-1}q$, $\tilde{\Phi} = k^{-2}\Phi$.

По сути дела, формулы (26.1), (26.2) задают четыре различные метрики: из-за наличия знака модуля у производной в (26.2) есть две области со статической метрикой и две области с однородной метрикой, отличающиеся знаком производной $dq/d\zeta$. Будем обозначать эти области римскими цифрами:

$$\begin{array}{rlll} \mathrm{I}: & \Phi > 0, & dq/d\sigma > 0, \\ \mathrm{II}: & \Phi < 0, & dq/d\tau < 0, \\ \mathrm{III}: & \Phi > 0, & dq/d\sigma < 0, \\ \mathrm{IV}: & \Phi < 0, & dq/d\tau > 0. \end{array} \tag{26.6}$$

Порядок следования областей будет ясен из дальнейшего построения.

Статическое локальное решение в области III получается из локального решения в области I пространственным отражением $\sigma \to -\sigma$, а однородное локальное решение в области IV связано с локальным решением в области II обращением времени $\tau \to -\tau$. Поскольку изменение знака конформного множителя у метрики можно компенсировать перестановкой пространственной и временно́й координаты $\tau \leftrightarrow \sigma$, то однородные локальные решения в областях II и IV можно получить из стационарных локальных решений для конформного множителя $-\Phi$ в областях III и I с последующим поворотом плоскости τ, σ на угол $\pi/2$ по часовой стрелке.

Пример 26.1.1. Рассмотрим t, r компоненты метрики Шварцшильда вне горизонта при $r > 2M > 0$

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \frac{dr^2}{1 - \dfrac{2M}{r}}.$$

Эту двумерную метрику можно записать в вейлевски плоском виде

$$ds^2 = \left(1 - \frac{2M}{r}\right) (dt^2 - d\sigma^2),$$

преобразовав радиальную координату $r = r(\sigma)$, где функция $r(\sigma)$ удовлетворяет уравнению

$$\left|\frac{dr}{d\sigma}\right| = 1 - \frac{2M}{r}.$$

Это – статические области I,III.

Если координата r лежит в интервале $0 < r < 2M$, то ее следует назвать временем, т.е. произвести замену $r \leftrightarrow t$. Следовательно, при $0 < t < 2M$ (под горизонтом) метрика Шварцшильда имеет вид

$$ds^2 = -\frac{dt^2}{1 - \dfrac{2M}{t}} + \left(1 - \frac{2M}{t}\right) dr^2.$$

Ее также можно записать в вейлевски плоском виде

$$ds^2 = \left|1 - \frac{2M}{t}\right| (d\tau^2 - dr^2),$$

где

$$\left|\frac{dt}{d\tau}\right| = -\left(1 - \frac{2M}{t}\right).$$

Таким образом мы получили однородную метрику в областях II, IV.

Как видим, двумерная t, r часть метрики Шварцшильда в точности имеет вид (26.1) с конформным множителем

$$\Phi(q) = 1 - \frac{2M}{q}, \tag{26.7}$$

где $q = r, t$. Конформный множитель $\Phi(q)$ для метрики Шварцшильда имеет простой полюс при $q = 0$ и простой нуль (горизонт) в точке $q = 2M$. □

Для метрики (26.1) в областях I и III можно перейти к координатам τ, q, в которых метрика примет вид

$$ds^2 = \Phi(q) d\tau^2 - \frac{dq^2}{\Phi(q)}.$$

В областях II и IV можно перейти к координатам q, σ:

$$ds^2 = -\frac{dq^2}{\Phi(q)} + \Phi(q) d\sigma^2.$$

Координаты τ, q в областях I, III и q, σ в областях II, IV называются *координатами Шварцшильда*. Метрика в этих координатах имеет простой вид, причем все компоненты метрики заданы явно. Недостатком координат Шварцшильда является то,

что они не различают между собой области I и III, а также II и IV. Это существенно, потому что для построения глобальных решений необходимо использовать все области.

Перейдем к вычислению геометрических объектов для метрики (26.1). Символы Кристоффеля имеют различный вид в различных областях. Ниже приведены явные выражения только для ненулевых компонент:

$$\text{I}: \ \Gamma_{\tau\tau}{}^{\sigma} = \Gamma_{\tau\sigma}{}^{\tau} = \Gamma_{\sigma\tau}{}^{\tau} = \Gamma_{\sigma\sigma}{}^{\sigma} = \frac{1}{2}\Phi', \tag{26.8}$$

$$\text{II}: \ \Gamma_{\tau\tau}{}^{\tau} = \Gamma_{\tau\sigma}{}^{\sigma} = \Gamma_{\sigma\tau}{}^{\sigma} = \Gamma_{\sigma\sigma}{}^{\tau} = \frac{1}{2}\Phi', \tag{26.9}$$

$$\text{III}: \ \Gamma_{\tau\tau}{}^{\sigma} = \Gamma_{\tau\sigma}{}^{\tau} = \Gamma_{\sigma\tau}{}^{\tau} = \Gamma_{\sigma\sigma}{}^{\sigma} = -\frac{1}{2}\Phi', \tag{26.10}$$

$$\text{IV}: \ \Gamma_{\tau\tau}{}^{\tau} = \Gamma_{\tau\sigma}{}^{\sigma} = \Gamma_{\sigma\tau}{}^{\sigma} = \Gamma_{\sigma\sigma}{}^{\tau} = -\frac{1}{2}\Phi', \tag{26.11}$$

где штрих обозначает производную по q. В статичных областях отличны от нуля символы Кристоффеля с нечетным числом пространственных индексов, а в однородных областях – с нечетным числом временны́х индексов. В статичных и однородных областях символы Кристоффеля отличаются знаком. Это является следствием того, что однотипные области связаны преобразованием $\zeta \to -\zeta$, а символы Кристоффеля линейны по производным. У тензора кривизны в каждой области отличны от нуля только четыре компоненты:

$$\text{I, III}: \ R_{\tau\sigma\tau}{}^{\sigma} = -R_{\sigma\tau\tau}{}^{\sigma} = R_{\tau\sigma\sigma}{}^{\tau} = -R_{\sigma\tau\sigma}{}^{\tau} = -\frac{1}{2}\Phi''\Phi, \tag{26.12}$$

$$\text{II, IV}: \ R_{\tau\sigma\tau}{}^{\sigma} = -R_{\sigma\tau\tau}{}^{\sigma} = R_{\tau\sigma\sigma}{}^{\tau} = -R_{\sigma\tau\sigma}{}^{\tau} = \frac{1}{2}\Phi''\Phi. \tag{26.13}$$

Тензор Риччи диагонален:

$$\text{I, III}: \ R_{\tau\tau} = -R_{\sigma\sigma} = -\frac{1}{2}\Phi''\Phi, \tag{26.14}$$

$$\text{II, IV}: \ R_{\tau\tau} = -R_{\sigma\sigma} = \frac{1}{2}\Phi''\Phi. \tag{26.15}$$

При этом во всех четырех областях скалярная кривизна одинакова,

$$\text{I, II, III, IV}: \quad R = -\Phi''. \tag{26.16}$$

Замечание. При заданном конформном множителе Φ это алгебраическое уравнение связывает q с R. Отсюда вытекает, что функция $q = q(\tau, \sigma)$, также как и скалярная кривизна, может рассматриваться, как функция на плоскости. В той области, где уравнение (26.16) разрешимо относительно q в качестве одной из координат можно выбрать скалярную кривизну R. □

Максимальное продолжение многообразия означает, что, если у поверхности (пространства-времени) с метрикой (26.1) есть край, лежащий на конечном расстоянии, т.е. соответствующий конечному значению канонического параметра для экстремалей, то он может быть только сингулярным. В противном случае многообразие можно и необходимо продолжить.

Поскольку в двумерном случае все компоненты тензора кривизны определяются скалярной кривизной, то проанализируем ее поведение подробнее.

Из уравнения (26.16) вытекает, что для поверхностей постоянной кривизны, которые рассмотрены в разделе 26.5.5, конформный множитель представляет собой квадратичный полином:

$$\Phi(q) = -(aq^2 + bq + c), \qquad a, b, c = \mathsf{const}. \tag{26.17}$$

При этом $R = 2a$.

Из уравнения (26.16) также следует, что скалярная кривизна сингулярна вблизи точки q_i при следующих показателях степени асимптотического поведения (26.4):

$$|q_i| < \infty : \qquad m < 0, \quad 0 < m < 1, \quad 1 < m < 2, \tag{26.18}$$

$$|q_i| = \infty : \qquad m > 2. \tag{26.19}$$

Отметим, что при конечных значениях q_i и при $m = 1$ скалярная кривизна может иметь отличные от нуля значения за счет поправок второго порядка. При бесконечных значениях параметра, $q \to \pm\infty$, скалярная кривизна стремится к отличной от нуля постоянной при $m = 2$ и к нулю при $m < 2$. Отметим, что ненулевое значение кривизны в конечной точке $|q_i| < \infty$ может возникнуть и при $m = 1$ за счет поправок следующего порядка в разложении (26.4).

26.2 Конформные блоки

При построении глобальных решений будет использовано понятие конформного блока, который ставится в соответствие каждому интервалу (q_i, q_{i+1}). Внутри каждого интервала конформный множитель Φ либо строго положителен, либо строго отрицателен.

Сначала найдем область определения соответствующей метрики (26.1) на плоскости τ, σ. Для определенности, рассмотрим статическое локальное решение I на интервале (q_i, q_{i+1}). Тогда временна́я координата пробегает всю числовую ось, $\tau \in \mathbb{R}$. Уравнение (26.2) для пространственной координаты σ принимает вид

$$\frac{dq}{d\sigma} = \Phi(q). \tag{26.20}$$

Постоянная интегрирования этого уравнения соответствует сдвигу пространственной координаты, $\sigma \to \sigma + \mathsf{const}$, т.е. выбору начала отсчета σ. Из уравнения (26.20) следует, что область изменения координаты σ зависит от сходимости интеграла

$$\sigma_{i,i+1} \sim \int^{q_i, q_{i+1}} \frac{dq}{\Phi(q)} \tag{26.21}$$

в граничных точках. Интеграл (26.21) сходится или расходится в зависимости от показателя степени m:

$$\begin{array}{ll} |q_i| < \infty : & \left\{ \begin{array}{llll} m < 1 & - \text{сходится}, & \text{прямая}, \\ m \geq 1 & - \text{расходится}, & \text{угол}, \end{array} \right. \\ |q_i| = \infty : & \left\{ \begin{array}{llll} m \leq 1 & - \text{расходится}, & \text{угол}, \\ m > 1 & - \text{сходится}, & \text{прямая}. \end{array} \right. \end{array} \tag{26.22}$$

В этой таблице справа указана форма границы соответствующих конформных блоков, которые будут введены ниже. Если на обоих концах интервала (q_i, q_{i+1}) интеграл

расходится, то $\sigma \in (-\infty, \infty)$, и метрика определена на всей плоскости τ, σ. Если на одном из концов q_{i+1} или q_i интеграл сходится, то метрика определена на полуплоскости $\sigma \in (-\infty, \sigma_{i+1})$ или $\sigma \in (\sigma_i, \infty)$, соответственно. При этом выбор граничных точек σ_{i+1} и σ_i произволен, и, не ограничивая общности, можно положить $\sigma_{i,i+1} = 0$. Если на обоих концах интеграл сходится, то локальное решение определено в полосе $\sigma \in (\sigma_i, \sigma_{i+1})$, при этом совместить с началом координат можно только один из концов интервала.

Для наглядного изображения максимально продолженных решений введем понятие конформного блока. С этой целью отобразим плоскость τ, σ на квадрат вдоль светоподобных направлений:

$$\xi = \tau + \sigma, \qquad \eta = \tau - \sigma. \tag{26.23}$$

Для этого произведем конформную замену переменных:

$$u = u(\xi), \quad v = v(\eta), \qquad u, v \in \mathcal{C}^{l+1}, \tag{26.24}$$

где функции u и v ограничены на всей оси, а их первые производные положительны. Например, $u = \operatorname{arctg} \xi$, $v = \operatorname{arctg} \eta$. Предположение о классе гладкости функций перехода сохраняет класс гладкости метрики после преобразования координат. Тогда статическому локальному решению, определенному на всей плоскости τ, σ, будет поставлен в соответствие квадратный *конформный блок*, изображенный на рис. 26.1, *a*.

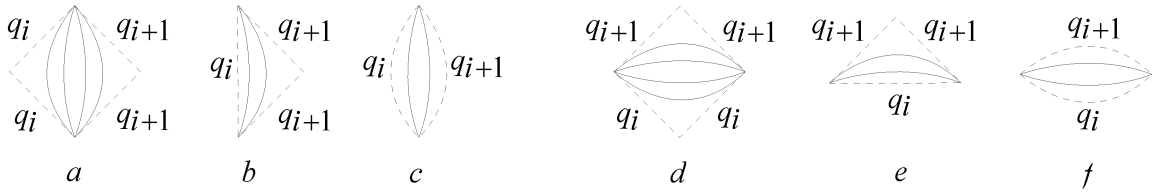


Рис. 26.1: Конформные блоки для статических (*a,b,c*) и однородных (*d,e,f*) локальных решений. Тонкие сплошные линии обозначают траектории Киллинга.

В общем случае под конформным блоком понимается конечная часть плоскости u, v и заданная на ней метрика (26.1) в координатах (26.23), (26.24). Если потребовать, чтобы экстремали подходили к границе конформного блока под всеми возможными углами, то это однозначно фиксирует асимптотическое поведение функций перехода (26.24) [?]. В остальном функции перехода остаются произвольными. Мы не будем останавливаться на этом вопросе подробно, т.к. гладкость метрики на горизонтах будет доказана путем перехода к координатам Эддингтона–Финкельстейна в разделе 26.6.

Значение переменной q и, значит, скалярной кривизны постоянно вдоль времениподобных траекторий Киллинга, которые внутри блока показаны тонкими сплошными линиями. Переменная q непрерывно меняется слева направо: монотонно возрастает в области I и монотонно убывает в области III. При этом на двух левых сторонах квадрата она принимает значение q_i, а на двух правых – значение q_{i+1}. Нижняя и верхняя вершины квадрата являются существенно особыми точками, т.к. в этих точках предел q зависит от пути, по которому произвольно выбранная последовательность подходит к вершине. Будем говорить, что у статического конформного блока есть две границы: левая и правая, на которых переменная q принимает значения q_i и q_{i+1}, соответственно.

Если решение уравнения (26.20) определено на полуинтервале $(\sigma_i = 0, \sigma_{i+1} = \infty)$, то статическому локальному решению ставится в соответствие треугольный конформный блок, пример которого показан на рис. 26.1,*b*. При этом конечной граничной точке $\sigma_i = 0$ ставится в соответствие вертикальная прямая. Этого всегда можно добиться путем выбора функций $u(\xi)$ и $v(\eta)$.

В случае, когда решение уравнения (26.20) определено на конечном интервале (σ_i, σ_{i+1}), конформный блок изображается в виде линзы, изображенной на рис. 26.1,*c*. За счет выбора функций $u(\xi)$ и $v(\eta)$ левую или правую границу этого конформного блока можно сделать вертикальной, но не обе границы одновременно.

Напомним, что каждому статическому интервалу (q_i, q_{i+1}) ставится в соответствие два конформных блока, т.к. благодаря знаку модуля, уравнение (26.2) инвариантно относительно инверсии пространственной координаты $\sigma \leftrightarrow -\sigma$.

Аналогично каждому однородному локальному решению ставится в соответствие один из конформных блоков, изображенных на рис. 26.1,*d,e,f*, или перевернутый блок, полученный обращением времени $\tau \leftrightarrow -\tau$. Для этих конформных блоков переменная q постоянна вдоль пространственноподобных траекторий Киллинга и монотонно возрастает или убывает от значения q_i на нижней границе, до значения q_{i+1} на верхней границе. Левая и правая вершины однородных конформных блоков являются существенно особыми точками.

Конформные блоки для статических и однородных локальных решений будем называть, соответственно, статическими и однородными.

Итак, мы показали, что лоренцева поверхность, на которой задана конформно плоская метрика (26.1) на каждом интервале (q_i, q_{i+1}) диффеоморфна одному из конформных блоков. Для построения максимально продолженных решений необходимо найти и проанализировать экстремали на этих поверхностях, при необходимости продолжить многообразие и метрику, а также доказать дифференцируемость метрики при склеивании конформных блоков. Использование конформных блоков для построения глобальных решений удобно и наглядно, т.к. все светоподобные экстремали изображаются в виде двух семейств прямых линий, проходящих под углом $\pm\pi/4$ к оси времени. Поэтому при продолжении поверхности конформные блоки можно склеивать вдоль ребер, сохраняя гладкость светоподобных экстремалей.

26.3 Экстремали

26.3.1 Форма экстремалей

Для того, чтобы понять, как устроено глобальное решение для метрики (26.1), необходимо проанализировать поведение экстремалей $\{\zeta^\alpha(t)\} = \{\tau(t), \sigma(t)\}$, $t \in \mathbb{R}$, которые подчиняются системе уравнений

$$\ddot{\zeta}^\alpha = -\Gamma_{\beta\gamma}{}^{\alpha}\dot{\zeta}^\beta\dot{\zeta}^\gamma, \tag{26.25}$$

где точка обозначает дифференцирование по каноническому параметру t и $\Gamma_{\beta\gamma}{}^{\alpha}$ – символы Кристоффеля.

Исследуем подробно поведение экстремалей для статического локального решения в области

$$\mathrm{I}: \quad ds^2 = \Phi(q)(d\tau^2 - d\sigma^2), \qquad \frac{dq}{d\sigma} = \Phi(q) > 0. \tag{26.26}$$

Используя выражение для символов Кристоффеля (26.8), получаем уравнения для экстремалей

$$\ddot{\tau} = -\Phi'\dot{\tau}\dot{\sigma}, \tag{26.27}$$

$$\ddot{\sigma} = -\frac{1}{2}\Phi'(\dot{\tau}^2 + \dot{\sigma}^2). \tag{26.28}$$

Эта система уравнений имеет два первых интеграла (см. раздел 17.3):

$$\Phi(\dot{\tau}^2 - \dot{\sigma}^2) = C_0 = \mathsf{const}, \tag{26.29}$$

$$\Phi\dot{\tau} = C_1 = \mathsf{const}. \tag{26.30}$$

Интеграл движения (26.29) существует для любой метрики и позволяет выбрать длину времениподобной экстремали в качестве канонического параметра. Значение постоянной C_0 определяет тип экстремали:

$C_0 > 0$ – времениподобные экстремали,
$C_0 = 0$ – светоподобные экстремали,
$C_0 < 0$ – пространственноподобные экстремали.

Отметим, что тип экстремали не может меняться от точки к точке. Второй из интегралов (26.30) связан с наличием вектора Киллинга, т.е. с симметрией задачи, и для произвольной метрики он отсутствует.

В системе уравнений для экстремалей (26.27), (26.28) конформный множитель Φ рассматривается, как сложная функция от σ: $\Phi = \Phi\Big(q(\sigma)\Big)$, где $q(\sigma)$ – решение уравнения (26.26). Если форма экстремали $\sigma(\tau)$ найдена, то конформный множитель можно рассматривать также как сложную функцию от τ вдоль каждой экстремали: $\Phi = \Phi\Big(\sigma(\tau)\Big)$. В дальнейшем анализе это будет подразумеваться.

Теорема 26.3.1. *Любая экстремаль в статическом пространстве-времени типа* I, III *принадлежит одному из следующих четырех классов.*
1) Светоподобные экстремали

$$\tau = \pm\sigma + \mathsf{const}, \tag{26.31}$$

для которых канонический параметр определен уравнением

$$\dot{\tau} = \frac{1}{\Phi}. \tag{26.32}$$

2) Экстремали общего вида, форма которых определена уравнением

$$\frac{d\tau}{d\sigma} = \pm\frac{1}{\sqrt{1 - C_2\Phi}}, \tag{26.33}$$

где C_2 – произвольная отличная от нуля постоянная, причем значения $C_2 < 0$ и $C_2 > 0$ описывают соответственно пространственно- и времениподобные экстремали. Канонический параметр определяется любым из двух уравнений:

$$\dot{\tau} = \frac{1}{\Phi}, \tag{26.34}$$

$$\dot{\sigma} = \pm\frac{\sqrt{1 - C_2\Phi}}{\Phi}. \tag{26.35}$$

При этом в уравнениях (26.33) и (26.35) знаки плюс или минус выбираются одновременно.

3) Прямые пространственноподобные экстремали, параллельные оси σ и проходящие через каждую точку $\tau = \mathsf{const}$. Канонический параметр определен уравнением

$$\dot{\sigma} = \frac{1}{\sqrt{\Phi}}. \tag{26.36}$$

4) Прямые вырожденные времениподобные экстремали, параллельные оси τ и проходящие через критические точки $\sigma_0 = \mathsf{const}$, в которых

$$\Phi'(\sigma_0) = 0. \tag{26.37}$$

Канонический параметр для них совпадает с временно́й координатой,

$$t = \tau. \tag{26.38}$$

Доказательство. Доказательство теоремы сводится к интегрированию системы уравнений (26.27), (26.28). При получении интеграла движения (26.30) необходимо делить на $\dot{\tau}$ и $\dot{\sigma}$, поэтому эти вырожденные случаи рассматриваются отдельно.

Начнем с вырожденных случаев. Для $\tau = \mathsf{const}$ уравнение (26.27) выполняется автоматически, а уравнение (26.28) после интегрирования приводит к условию

$$-\Phi\dot{\sigma}^2 = C_0 < 0.$$

Это уравнение сводится к (26.36) растяжкой канонического параметра.

Второй вырожденный случай соответствует $\dot{\sigma} = 0$. При этом уравнение (26.28) выполняется только в тех точках $\sigma = \sigma_0$, в которых $\Phi'(\sigma_0) = 0$, а уравнение (26.27) сводится к уравнению $\ddot{\tau} = 0$. Отсюда вытекает, что координату τ можно выбрать в качестве канонического параметра. Здесь можно лишь заметить справедливость равенства

$$\frac{d\Phi}{d\sigma} = \Phi\frac{d\Phi}{dq} = 0,$$

и, поскольку $\Phi > 0$, то в уравнении (26.37) дифференцирование по q можно заменить дифференцированием по σ. Таким образом исчерпываются оба специальных случая.

Остальные экстремали можно получить из анализа интегралов (26.29), (26.30), причем константа C_0 может быть произвольной, а $C_1 \neq 0$, т.к. случай $C_1 = 0$ относится к прямым экстремалям третьего класса. Уравнение (26.29) с учетом (26.30) приводит к равенству

$$\dot{\sigma} = \pm\frac{C_1}{\Phi}\sqrt{1 - C_2\Phi}, \tag{26.39}$$

где

$$C_2 = \frac{C_0}{C_1^2}. \tag{26.40}$$

Уравнение для формы экстремалей общего вида (26.33) является следствием (26.30) и (26.39), причем значение постоянной C_2 определяет тип экстремали. Уравнения для канонического параметра получаются из (26.30) и (26.39) после растяжки. В частном случае $C_2 = 0$ получаем светоподобные экстремали (26.31). □

Замечание. Из уравнения (26.33) следует, что постоянная C_2 параметризует угол, под которым экстремаль общего вида проходит через заданную точку. Можно проверить, что через произвольную точку в каждом направлении проходит одна и только одна экстремаль. Это подтверждает тот факт, что найдены все экстремали. □

Доказанная теорема позволяет качественно понять поведение экстремалей для произвольного конформного множителя.

На рис. 26.2 в первом ряду показано типичное поведение конформного множителя $\Phi(q)$ с одним и тремя локальными экстремумами между двумя нулями и с двумя локальными экстремумами между нулем и особенностью для статических конформных блоков. Во втором ряду показана зависимость этих конформных множителей $\Phi(\sigma)$ от пространственной координаты. В третьем ряду показано качественное поведение времени- и пространственноподобных экстремалей. Чтобы не загромождать рисунок, мы опустили светоподобные экстремали, проходящие через каждую точку пространства-времени, а также экстремали, параллельные оси σ. Все экстремали можно сдвигать вдоль времениподобной координаты τ.

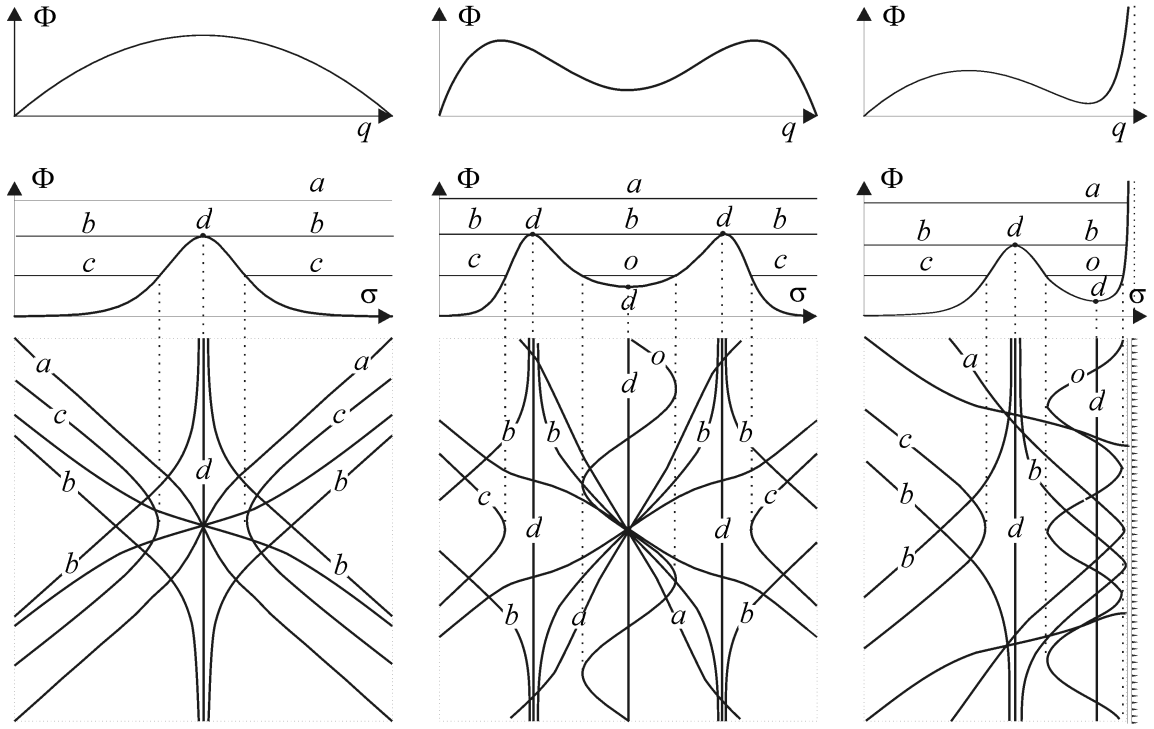


Рис. 26.2: Типичное поведение конформного множителя $\Phi(q)$ между двумя нулями и нулем и сингулярностью для статического локального решения (верхний ряд). Зависимость конформного множителя $\Phi(\sigma)$ от пространственной координаты (средний ряд). Типичные времени- и пространственно подобные экстремали (нижний ряд). Буквами отмечены времениподобные экстремали, имеющие качественно различное поведение при разных значениях постоянной C_2. Через каждый локальный экстремум проходит вырожденная экстремаль d. Во втором и третьем случае имеются осциллирующие экстремали o.

Если обе граничные точки интервала q_i и q_{i+1} являются нулями, то из непрерывности конформного множителя Φ следует, что у него есть по крайней мере один максимум, через который проходит вырожденная экстремаль d. В общем случае через каждый локальный экстремум функции $\Phi(\sigma)$ проходит вертикальная экстремаль. На рисунках они помечены буквой d.

Пространственноподобные экстремали общего вида соответствуют отрицательным значениям постоянной C_2 в уравнении (26.33), и, поскольку $\Phi > 0$, подынтегральное выражение всюду определено. Они начинаются на левой границе кон-

формного блока и заканчиваются на правой. Пространственноподобные экстремали общего вида показаны на рисунке без каких либо пометок.

Времениподобные экстремали общего вида имеют качественно различное поведение при различных значениях постоянной $C_2 > 0$. Они помечены буквами *a,b,c* и *o*.

Если постоянная C_2 достаточно мала,

$$C_2 < \frac{1}{\max \Phi(\sigma)},$$

то времениподобная экстремаль общего вида определена при всех значениях $\sigma \in (\sigma_i, \sigma_{i+1})$. Такие экстремали соединяют левую и правую границу конформного блока, пересекая все вырожденные экстремали, и помечены на рисунке буквой *a*.

Во втором ряду рис.26.2 горизонтальными линиями показаны области определения координаты σ,

$$C_2\Phi(\sigma) \leq 1, \tag{26.41}$$

при различных значениях постоянной C_2.

Если у функции Φ имеется локальный минимум, которому также соответствует вырожденная экстремаль, то среди экстремалей общего вида существуют времениподобные экстремали, осциллирующие вблизи локального минимума, которые помечены буквой *o*. Это следует из уравнения (26.33), т.к. при достаточно больших положительных значениях C_2 область изменения координаты σ определяется неравенством (26.41). Критические точки σ_*, в которых $\Phi(\sigma_*) = 1/C_2$, являются точками поворота при конечном значении $\tau = \tau_*$ тогда и только тогда, когда интеграл

$$\int^{\tau_*} d\tau = \pm \int^{\sigma_*} \frac{d\sigma}{\sqrt{1 - C_2\Phi}} \tag{26.42}$$

сходится. Этот интеграл сходится, если точка σ_* не совпадает с локальным максимумом. То есть поворот экстремали происходит при конечном значении τ_* и канонического параметра, что вытекает из уравнения (26.35).

При том же значении постоянной C_2 существуют времениподобные неосциллирующие экстремали общего вида, которые приходят и уходят на одну и ту же (левую или правую) границу. У них есть точка поворота при конечных значениях τ. Такие экстремали на рисунке помечены буквой *c*.

Если критическая точка σ_* совпадает с локальным максимумом, то

$$\Phi \simeq \frac{1}{C_2} - C_3(\sigma - \sigma_*)^2 + \ldots, \qquad C_3 = \mathsf{const} > 0,$$

и интеграл (26.42) расходится. Это означает, что критическая точка достигается при бесконечном значении τ_* и канонического параметра. То есть соответствующая экстремаль в этом направлении полна. На рисунке такие экстремали помечены буквой *b*.

Приведенная теорема описывает все экстремали для статических локальных решений. Поведение экстремалей для однородных локальных решений аналогично. Они получаются заменой $\Phi \mapsto -\Phi$, перестановкой временно́й и пространственной координат $\tau \leftrightarrow \sigma$ и поворотом на угол $\pi/2$ всей плоскости.

26.3.2 Асимптотика экстремалей

Для решения вопроса о том является ли метрика, определенная на τ, σ плоскости, полной по экстремалям или локальное решение следует продолжить, необходимо исследовать асимптотику и полноту экстремалей вблизи границы конформного блока. Забегая вперед, заметим, что все экстремали, за исключением осциллирующих, имеют асимптотику при приближении к границе конформного блока.

Рассмотрим квадратный конформный блок. В верхнюю и нижнюю вершины конформного блока попадают все экстремали, для которых

$$\lim_{\tau\to\pm\infty}\left|\frac{d\tau}{d\sigma}\right| > 1,$$

если предел существует. В левую и правую вершины конформного блока попадают экстремали, для которых

$$\lim_{\sigma\to\pm\infty}\left|\frac{d\tau}{d\sigma}\right| < 1. \tag{26.43}$$

На стороны квадратного конформного блока могут попасть только экстремали, имеющие светоподобную асимптотику:

$$\lim_{\tau,\sigma\to\pm\infty}\left|\frac{d\tau}{d\sigma}\right| = 1. \tag{26.44}$$

В принципе, в вершину конформного блока могут попасть также экстремали, имеющие светоподобную асимптотику. Однако, как будет показано ниже, этого не происходит.

Если одна из границ q_i, например, статического конформного блока прямая, то на нее попадают экстремали со светоподобной (26.44) и пространственноподобной (26.43) асимптотикой.

Осциллирующие экстремали всегда зажаты прямыми вырожденными экстремалями. Поэтому они попадают в ту же вершину конформного блока, что и прямые вырожденные экстремали 4) из теоремы 26.3.1.

Рассмотрим асимптотику экстремалей вблизи границы статических конформных блоков. Любая светоподобная экстремаль на квадратном конформном блоке начинается и заканчивается на противоположных сторонах. В случае треугольного конформного блока или блока типа линзы светоподобные экстремали начинаются или заканчиваются на времениподобных границах.

Для статических локальных решений типа I, III прямые экстремали, параллельные пространственной оси σ начинаются в левой и заканчиваются в правой вершине квадратного конформного блока. В случае треугольного блока и блока типа линзы эти экстремали начинаются или (и) заканчиваются на времениподобной границе. Вырожденные экстремали начинаются в нижней и заканчиваются в верхней вершине статического конформного блока. Осциллирующие экстремали зажаты вырожденными экстремалями и поэтому также начинаются и заканчиваются в нижней и верхней вершине.

Таким образом, мы поняли в какие граничные точки конформного блока попадают все прямые экстремали, а также осциллирующие осциллирующие экстремали общего вида.

Осталось рассмотреть поведение неосциллирующих экстремалей общего вида вблизи границы. Сначала определим точки границы конформного блока, в которые они

попадают. Допустим, что граница q_i конформного блока является нулем, $\Phi(q_i) = 0$, $|q_i| < \infty$. Этому соответствует положительный показатель $m > 0$ (26.4). Соответствующая граница конформного блока является углом при $m \geq 1$ и прямой при $0 < m < 1$ (26.107). В этом случае из уравнения для экстремалей (26.33) следует, что векторы, касательные к ним, имеют светоподобную асимптотику

$$\lim_{q \to q_i} \left| \frac{d\tau}{d\sigma} \right| = 1.$$

Если граница конформного блока является углом, то экстремали общего вида попадают на стороны угла, а не в вершину. Действительно, из уравнений (26.34), (26.35) для экстремалей, идущих в правую верхнюю сторону квадратного конформного блока, следует равенство

$$\frac{d\eta}{d\xi} = \frac{1 - \sqrt{1 - C_2\Phi}}{1 + \sqrt{1 + C_2\Phi}}.$$

Для малых $\Phi \ll 1$ разложим правую часть этого равенства в ряд. В первом порядке по Φ получим равенство

$$\frac{d\eta}{d\xi} \simeq \frac{1}{2}\left[1 - \left(1 - \frac{C_2\Phi}{2}\right)\right] = \frac{C_2}{4}\Phi.$$

Откуда вытекает, что интеграл

$$\int d\eta \sim \int d\xi\Phi \sim \int d\sigma\Phi \sim \int^{q_i} dq = q_i$$

сходится. Это означает, что точка $\xi \to \infty$ достигается при конечных значениях η, и это соответствует некоторой точке на стороне угла, а не вершина. Если граница q_i конформного блока является времениподобной, то экстремали общего вида попадают на нее.

Пусть теперь на границе $\Phi(q_i) = \infty$, $|q_i| < \infty$. В этом случае $m < 0$, граница времениподобна и достигается при конечных σ_i. Из равенства (26.33) следует, что касательный вектор к пространственноподобной экстремали при $\Phi \to \infty$ имеет асимптотику

$$\frac{d\tau}{d\sigma} \sim \frac{1}{\sqrt{\Phi}} \to 0.$$

То есть, они попадают на границу $|\sigma_i| < \infty$ под прямым углом.

Времениподобные экстремали общего вида не могут попасть на времениподобную границу, потому что правая часть в (26.33) при $C_2 > 0$ вблизи особенности $\Phi(q_i) = \infty$ становится отрицательной. Эти экстремали вблизи границы $|\sigma_i| < \infty$ имеют точку поворота при конечных значениях τ, по той же причине, что и осциллирующие экстремали.

26.3.3 Полнота экстремалей

Теорема 26.3.1 позволяет проанализировать полноту экстремалей при подходе к границе конформного блока. Для определенности рассмотрим стационарное локальное решение. Прежде всего заметим, что вырожденные экстремали всегда полны, т.к. канонический параметр для них совпадает с временем (26.38). Осциллирующие экстремали также полны поскольку делают бесконечное число осцилляций, каждая из

которых соответствует конечному изменению канонического параметра. Если же экстремаль общего вида при $\tau \to \pm\infty$ приближается к вырожденной экстремали, то она полна, поскольку из уравнения для канонического параметра (26.34) следует, что

$$\lim_{\sigma\to\sigma_0} t \to \int^{\infty} d\tau \Phi(\sigma_0) \to \infty.$$

Для стационарного локального решения в нижнюю и верхнюю вершины попадают только вырожденные и осциллирующие экстремали, и поэтому они всегда полны. Возможна ситуация, когда эти экстремали отсутствуют. В этом случае будем считать эти вершины полными, потому что любая времениподобная кривая имеет бесконечную длину при $\tau \to \pm\infty$. Это означает, что нижняя и верхняя вершины стационарного блока всегда полны, т.е. представляет собой временну́ю бесконечность прошлого и будущего, соответственно.

Полнота светоподобных экстремалей определяется уравнением (26.32), которое позволяет вычислить предел

$$\lim_{\sigma\to\pm\infty} t \to \int^{\sigma_i} d\sigma \Phi = \int^{q_i} dq = q_i. \tag{26.45}$$

Это значит, что при приближении к границе конформного блока они неполны при конечных q_i и полны при $|q_i| = \infty$.

Полнота неосциллирующих экстремалей общего вида, попадающих на границу конформного блока, соответствующую значению q_i, вытекает из уравнения (26.35)

$$\lim_{q\to q_i} t \to \int^{\sigma_i} d\sigma \frac{\Phi}{\sqrt{1 - C_2\Phi}} \sim \int^{q_i} \frac{dq}{\sqrt{1 - C_2\Phi}}. \tag{26.46}$$

Вблизи нулей $\Phi \to 0$ поведение экстремалей общего вида такое же, как и у светоподобных экстремалей (26.45). Особенность $\Phi \to \infty$ достигается только пространственноподобными экстремалями, $C_2 < 0$, полнота которых определяется интегралом

$$\lim_{q\to q_i} t \to \int^{q_i} \frac{dq}{\sqrt{\Phi}}. \tag{26.47}$$

Эти экстремали неполны в конечных точках $|q_i| < \infty$ при $m < 0$. В бесконечно удаленных точках $|q_i| = \infty$ пространственноподобные экстремали полны при $0 < m \leq 2$ и неполны при $m > 2$.

Полнота прямых экстремалей, параллельных оси σ, определяется уравнением (26.36) и при приближении к границе q_i дается интегралом (26.47). Это значит, что их полнота вблизи особенностей такая же, как и у экстремалей общего вида. Вблизи нулей они всегда полны за исключением случая простого нуля в конечной точке, где они неполны.

Таким образом проанализирована полнота всех элементов границы. Итог анализа приведен на рис. 26.3. Здесь в зависимости от показателя степени m (26.4) для конечных и бесконечных значений q_i показаны соответствующие элементы границы конформных блоков. Для определенности, показана правая граница статических конформных блоков типа I. На рисунке времениподобная граница показана вертикальной линией, а светоподобная – углом. Если на границе скалярная кривизна сингулярна (26.18), (26.19), то на ней присутствуют выступы. Неполные и полные границы показаны соответственно пунктиром и жирной сплошной линией. Исключение составляет полуполная граница, соответствующая $|q_i| = \infty$ при $m > 2$, рис. 26.3.

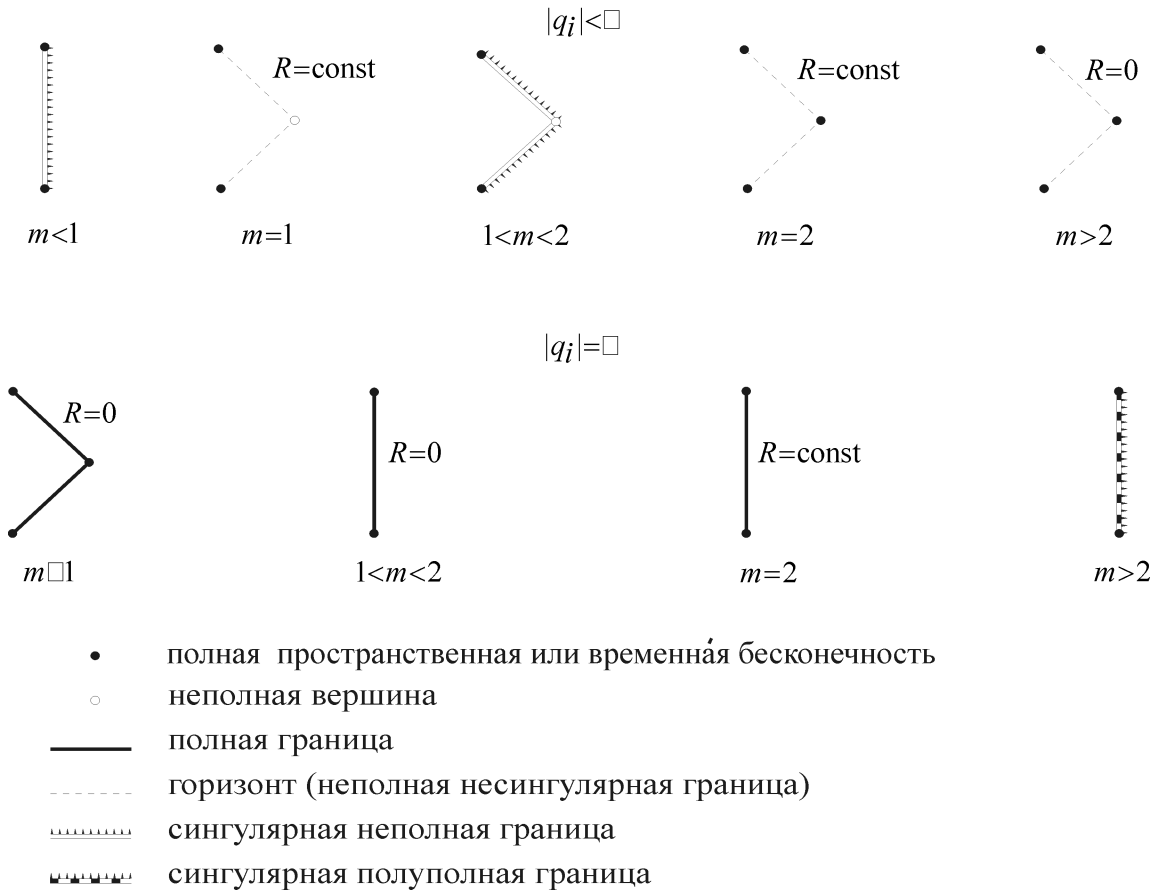


Рис. 26.3: Форма правой границы статических конформных блоков в зависимости от показателя степени m. В верхнем и нижнем рядах показаны границы для конечных и бесконечных значений q_i, соответственно.

Светоподобные экстремали, попадающие на эту сингулярную границу полны, а пространственноподобные экстремали – неполны. Нижние и верхние вершины всех статических конформных блоков (временоподобные бесконечности прошлого и будущего) являются существенно особыми точками и всегда полны, что отмечено закрашенными окружностями. Полнота или неполнота правых вершин границы в форме угла отмечена закрашенной или незакрашенной окружностью, соответственно. Левые границы статических конформных блоков имеют такое же строение, только углы нужно отразить относительно вертикальной прямой.

Качественное поведение всех экстремалей в однородном пространстве-времени типа II, IV определяется нулями и особенностями конформного фактора так же, как и для стационарного локального решения. При этом соответствующие конформные блоки должны быть повернуты на угол $\pi/2$.

26.4 Построение глобальных решений

В разделе 26.2 каждому локальному решению типа I-IV был сопоставлен конформный блок, определенный на интервале (q_i, q_{i+1}). Затем в разделе 26.3.3 из анализа экстремалей была доказана полнота или неполнота границ этих блоков. В результате оказалось, что нули конформного множителя, которые соответствуют *горизонтам* пространства-времени, и только они определяют неполную по экстремалям границу

конформного блока, на которой кривизна конечна. Во всех остальных случаях граница либо полна, либо на ней кривизна сингулярна. Поэтому решения вида (26.1) необходимо продолжить только через нули функции $\Phi(q)$. Напомним, что нулям конформного множителя (горизонтам) соответствуют границы в виде угла (см. рис. 26.3).

Сформулируем правила, по которым осуществляется максимальное продолжение поверхности с метрикой (26.1) при заданной функции $\Phi(q)$. Дифференцируемость максимально продолженной по экстремалям псевдоримановой поверхности, построенной по этим правилам дается приведенной ниже теоремой 26.10.1.

1. Каждое глобальное решение для метрики (26.1) соответствует некоторому интервалу значений переменной $q \in (q_-, q_+)$, где $q_\pm$ либо бесконечно удаленные точки $q_\pm = \pm\infty$, либо сингулярности кривизны, определяемые условием (26.18). Внутри интервала сингулярности должны отсутствовать.

2. Если внутри интервала (q_-, q_+) конформный блок не имеет нулей, то соответствующий конформный блок представляет собой глобальное решение.

3. Если внутри интервала (q_-, q_+) нули существуют, то пронумеруем их в порядке возрастания, $\Phi(q_j) = 0$, $j = 1, \ldots, n$, и поставим в соответствие каждому интервалу $(q_-, q_1), \ldots, (q_n, q_+)$ пару статических или однородных конформных блоков для $\Phi > 0$ и $\Phi < 0$, соответственно.

4. Склеим конформные блоки вдоль границ, соответствующих нулям q_j, причем склеиваются только блоки, соответствующие соседним интервалам (q_{j-1}, q_j) и (q_j, q_{j+1}).

5. Диаграмма Картера–Пенроуза получается склеиванием всех соседних конформных блоков. Она представляет собой связную фундаментальную область, если внутри интервала (q_-, q_+) конформный множитель меняет знак. Если $\Phi \geq 0$ или $\Phi \leq 0$ всюду внутри интервала (q_-, q_+), то получим две фундаментальные области, которые получаются друг из друга отражением пространства или времени.

6. При наличии в интервале (q_-, q_+) только одного нуля нечетного порядка граница фундаментальной области состоит из границ конформных блоков, соответствующих точкам q_- и q_+, и диаграмма Картера–Пенроуза представляет глобальное решение.

7. При наличии одного нуля четного порядка или двух и более нулей произвольного порядка граница фундаментальной области включает нули конформного множителя, и ее необходимо либо продолжить периодически в пространстве и (или) во времени, либо отождествить противоположные стороны вдоль горизонтов.

8. Если у диаграммы Картера–Пенроуза фундаментальная группа тривиальна, то она представляет собой универсальную накрывающую глобального решения.

9. Если у диаграммы Картера–Пенроуза фундаментальная группа нетривиальна, то построить соответствующую универсальную накрывающую.

Определение. *Диаграммой Картера–Пенроуза* называется образ на плоскости максимально продолженной поверхности, соответствующей интервалу (q_-, q_+) и полученный в результате склеивания всех конформных блоков для интервала (q_-, q_+) по правилам 1)–9). □

Диаграммы Картера–Пенроуза являются наглядным изображением максимально продолженной поверхности, т.к. оба класса непересекающихся между собой светоподобных экстремалей изображаются двумя классами перпендикулярных прямых, как на плоскости Минковского.

Поясним приведенные правила. Утверждение 1) является следствием того, что продолжение решения через точки $q_\pm$ невозможно, т.к. эти точки либо полны, либо сингулярны. Решения продолжаются только через горизонты $|q_j| < \infty$ с $m = 1$ или $m \geq 2$. В этих случаях граница конформного блока является углом. Если точка q_j – нуль нечетного порядка, то все четыре конформных блока, соответствующих интервалам (q_{j-1}, q_j) и (q_j, q_{j+1}) склеиваются согласно правилам 3) и 4) вокруг вершины q_j, причем эта точка является седловой для конформного множителя $\Phi(q)$, рассматриваемого как функция на диаграмме Картера–Пенроуза. Если q_j – нуль четного порядка, то вдоль соответствующего горизонта склеиваются между собой только области одного типа, например, I-I или III-III. Если внутри интервала (q_-, q_+) функция Φ не меняет знака, то возникают две несвязные между собой фундаментальные области, каждую из которых можно периодически продолжить. При изменении знака Φ возникает седловая точка, и области различных типов образуют связную фундаментальную область, которую можно либо периодически продолжить, либо нет, в зависимости от значений q на границе. Это составляет содержание правил 5) и 6).

Правила 5), 6) и 7) являются утверждениями, которые будут доказаны дальнейшим построением.

Сформулируем основную теорему, оправдывающую приведенные выше правила построения глобальных решений.

Теорема 26.4.1. *Универсальное накрывающее пространство, построенное по правилам 1)–9), представляет собой максимально продолженную вдоль экстремалей дифференцируемую псевдориманову поверхность класса $\mathcal{C}^{l+1}$ с метрикой класса $\mathcal{C}^l$, $l \geq 2$, такую, что внутренность каждого конформного блока изометрична поверхности с метрикой (26.1).*

Доказательство. По построению, внутренность каждого конформного блока покрывается одной картой и является многообразием класса $\mathcal{C}^{l+1}$, что следует из (26.24). Поэтому внутри конформного блока метрика принадлежит тому же классу гладкости, что и конформный множитель. При этом функции перехода (26.24) изометрично отображают плоскость τ, σ или ее часть, на которой определена метрика, на внутренность конформного блока. Поэтому необходимо доказать дифференцируемость многообразия и метрики только на горизонтах и в седловых точках. Это сделано путем перехода к новым системам координат в разделах 26.6 и 26.7, соответственно. □

Перед завершением доказательства теоремы рассмотрим несколько примеров, чтобы пояснить, как пользоваться правилами построения глобальных решений в конкретных случаях, и какие глобальные решения могут возникнуть.

26.5 Примеры

Если в результате решения уравнений движения какой либо модели гравитации возникает двумерная метрика вида (26.1), то, следуя правилам 1)–9) из предыдущего раздела можно построить глобальное решение, не заботясь о переходе к новым координатным системам, покрывающим бóльшие области. Достоинством рассматриваемого метода является его конструктивность, т.к. для построения глобальных решений достаточно элементарного анализа конформного множителя $\Phi(q)$. Начнем с трех известных примеров из общей теории относительности.

26.5.1 Решение Шварцшильда

Конформный множитель для решения Шварцшильда имеет вид (26.7). Он имеет простой полюс в точке $q = 0$ $(m = -1)$, который соответствует сингулярности кривизны (26.18). Точка $q_1 = 2M$ $(m = 1)$ является простым нулем и соответствует горизонту. При $M > 0$ нуль лежит на положительной полуоси. Значения $q = \pm\infty$ $(m = 0)$ соответствуют асимптотически плоской пространственной бесконечности $(R = 0)$. Поведение конформного множителя показано на рис. 26.4 слева.

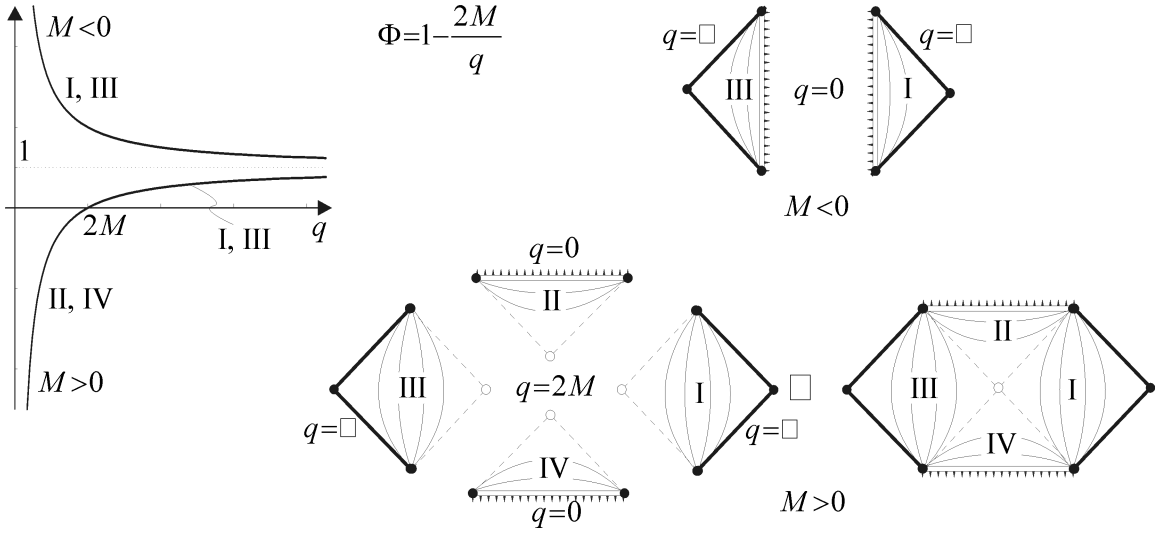


Рис. 26.4: Поведение конформного множителя для решения Шварцшильда для положительной и отрицательной массы. При $M < 0$ существует два глобальных решения, представленных треугольными диаграммами Картера–Пенроуза. При $M > 0$ четыре конформных блока, соответствующих двум интервалам $(0, 2M)$ и $(2M, \infty)$, склеиваются в одно глобальное решение.

Бесконечному интервалу $q \in (-\infty, \infty)$ соответствует два глобальных решения для положительных и отрицательных q. Рассмотрим случай $M > 0$. Для положительных значений $q \in (0, \infty)$ имеем $q_- = 0$, $q_1 = 2M$ и $q_+ = \infty$. Каждому интервалу $(q_-, q_1) = (0, 2M)$ и $(q_1, q_+) = (2M, \infty)$ соответствует по два однородных (II, IV) и статических (I, III) конформных блока, показанных на рис. 26.4 внизу в центре. При этом элементы границы определяются по рис. 26.3. Из этих четырех конформных блоков по правилу 4) единственным образом склеивается глобальное решение, показанное на рис. 26.4. Эта диаграмма Картера–Пенроуза представляет собой расширение Крускала–Секереша решения Шварцшильда [**?**, **?**]. Напомним, что расширение

Крускала–Секереша представляет собой запись метрики Шварцшильда в такой системе координат, которая покрывает сразу все области I–IV. Картер впервые изобразил это расширение в виде конечной области на плоскости [?].

Поскольку фундаментальная группа для полученной диаграммы тривиальна, то построенное глобальное решение является универсальной накрывающей для других глобальных решений. Оба статических конформных блока I или III диффеоморфны внешнему решению Шварцшильда при $r > 2M$. Теорема 26.10.1 обеспечивает, в данном случае, гладкость соответствующего глобального решения, и нет необходимости в явном построении глобальной системы координат.

Аналогичные глобальные решения будут возникать для широкого класса метрик, у которых конформный множитель ведет себя качественно так же, как и нижняя ветвь на рис. 26.4. То есть определен на полуинтервале (q_-, ∞), имеет особенность в q_-, один нуль и стремится к постоянной на бесконечности.

Отметим, что поверхность, представленная диаграммой Картера–Пенроуза имеет непостоянную скалярную кривизну,

$$R = \frac{4M}{q^3}. \tag{26.48}$$

Это – двумерная скалярная кривизна на лоренцевой поверхности. Скалярная кривизна, соответствующая четырехмерной метрике Шварцшильда с учетом угловой зависимости тождественно равна нулю в силу уравнений Эйнштейна. Заметим, что двумерная скалярная кривизна (26.48) совпадает с инвариантным собственным значением четырехмерного тензора Вейля [?]. Центр диаграммы Картера–Пенроуза является седловой точкой для переменной q и, следовательно, для скалярной кривизны (26.48). Эта точка неполна и отмечена незакрашенной окружностью.

Физическая интерпретация диаграммы Картера–Пенроуза для решения Шварцшильда состоит в следующем. В общей теории относительности пространство-время четырехмерно. Поэтому максимально продолженное сферически симметричное решение вакуумных уравнений Эйнштейна представляет собой топологическое произведение сферы $\mathbb{S}^2$ на диаграмму Картера–Пенроуза, изображенную на рис. 26.4 и представляющую собой ту часть пространства-времени, которая описывается координатами t и r. Нижняя и верхняя сингулярные границы диаграммы Картера–Пенроуза, умноженные на сферу, называются, соответственно, *белой* и *черной дырой*. Сфера $r = 2M$ называется *горизонтом* черной дыры. Конформный блок II, лежащий под горизонтом черной дыры, соответствует внутренности черной дыры. Если выбрать произвольную точку в области II, то любая времениподобная или светоподобная кривая неизбежно попадет на сингулярную пространственноподобную границу (черную дыру) за конечное собственное время. При этом нет никакой возможности пересечь горизонт. Внешние области I и III соответствуют двум разным вселенным для одной и той же черной дыры. Эти вселенные причинно не связаны между собой, поскольку их точки можно соединить только пространственноподобными кривыми. Если наблюдатель находится, например, во вселенной I, то у него есть две возможности: либо за конечное время упасть на черную дыру, либо жить бесконечно долго, поскольку бесконечность будущего (верхняя вершина конформного блока I) полна.

Для отрицательных значений q горизонты отсутствуют. Поэтому треугольные конформные блоки, показанные на рис. 26.4 вверху справа, представляют собой максимально продолженные решения. В этом пространстве-времени сингулярность кривизны располагается вдоль времениподобной границы, находящейся на конечном расстоянии и не окруженной горизонтом. Такие сингулярности в общей теории от-

носительности называются *голыми*. Заметим, однако, что на голые сингулярности попадают только свето- и пространственноподобные экстремали при конечном значении канонического параметра. Времениподобные экстремали отражаются вблизи сингулярности (см. рис. 26.2).

Глобальное решение для отрицательных q можно рассматривать как решение для положительных q (что соответствует интерпретации координаты q, как радиуса), но с отрицательной массой, $M < 0$. Эти решения рассматриваются в общей теории относительности, как нефизические.

26.5.2 Решение Рейснера–Нордстрема

Конформный множитель для решения Рейснера–Нордстрема [?, ?], описывающего заряженную черную дыру, имеет вид

$$\Phi = 1 - \frac{2M}{q} + \frac{Q^2}{q^2}, \qquad 0 < Q < M, \tag{26.49}$$

где M и Q – масса и заряд черной дыры. Функцию Φ при заданных соотношениях между константами можно записать в виде

$$\Phi = \frac{(q - q_1)(q - q_2)}{q^2}, \tag{26.50}$$

где

$$q_{1,2} = M \pm \sqrt{M^2 - Q^2}.$$

Отсюда следует, что конформный множитель имеет полюс второго порядка при $q = 0$ и два положительных простых нуля (горизонта) $q_{1,2}$. Его поведение для положительных q показано на рис. 26.5 слева.

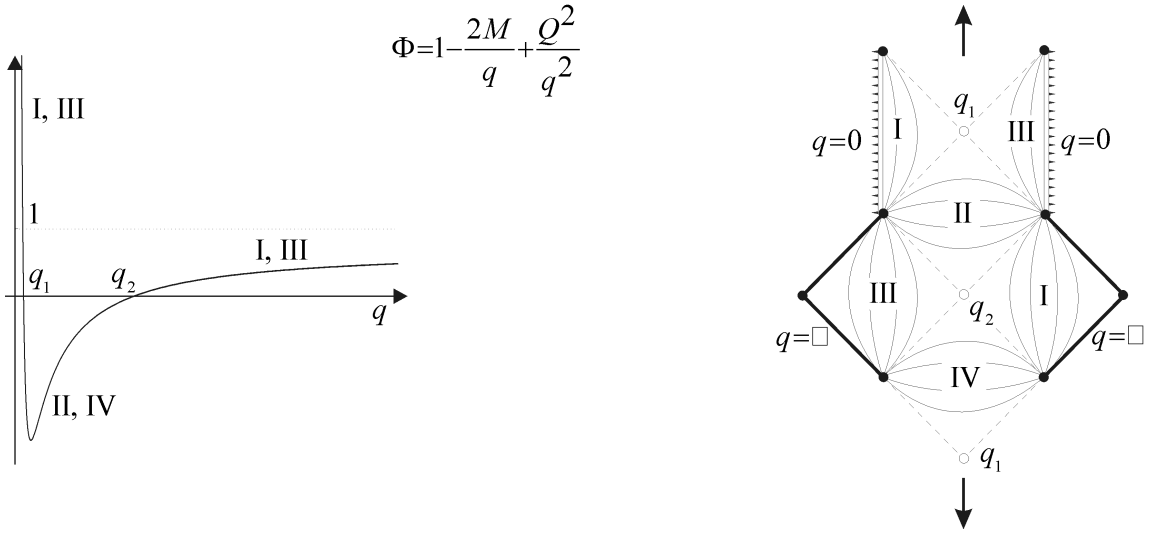


Рис. 26.5: Конформный множитель и фундаментальная область для решения Рейснера–Нордстрема. Стрелки показывают возможное продолжение или отождествление решения вдоль горизонтов.

Глобальная структура решения для положительных q, по сравнению с решением Шварцшильда, меняется качественно из-за наличия двух горизонтов. Каждому из

интервалов $(0, q_1)$, (q_1, q_2) и (q_2, ∞) соответствует по два конформных блока. Поскольку внутри интервала $(0, \infty)$ конформный множитель меняет знак, то по правилу 6) диаграмма Картера–Пенроуза, склеенная из шести конформных блоков, рис. 26.5, представляет собой фундаментальную область для решения Рейснера–Нордстрема. Ее граница состоит не только из сингулярных и бесконечно удаленных точек, но также включает отрезки, соответствующие горизонтам. В соответствии с правилом 7) ее можно периодически продолжить, склеивая одинаковые фундаментальные области вдоль горизонтов в направлениях, показанных стрелками. В этом случае мы получим универсальную накрывающую для решения Рейснера–Нордстрема. Другая возможность заключается в том, что после склеивания произвольного конечного числа фундаментальных областей, граничные точки на горизонтах снизу и сверху можно отождествить. Тогда глобальное решение с топологической точки зрения будет представлять собой цилиндр.

Сингулярности в решении Рейснера–Нордстрема являются голыми.

Скалярная кривизна (26.16) для решения Рейснера–Нордстрема непостоянна:

$$R = \frac{4M}{q^3} - \frac{6Q^2}{q^4}. \tag{26.51}$$

Это – двумерная скалярная кривизна поверхности, а не четырехмерного пространства-времени.

Левая ветвь конформного множителя $-\infty < q < 0$ для решения Рейснера–Нордстрема соответствует голой сингулярности и рассматриваться не будет, т.к. изменение порядка полюса не влияет на структуру глобального решения.

26.5.3 Экстремальная черная дыра

Экстремальная черная дыра возникает из решения Рейснера–Нордстрема (26.49) в том случае, когда заряд равен массе, $Q = M$. Соответствующий конформный множитель,

$$\Phi = \frac{(q - M)^2}{q^2}, \tag{26.52}$$

имеет полюс второго порядка при $q = 0$ и положительный нуль $q = M$ также второго порядка. Его поведение показано на рис. 26.6 слева. Левая ветвь $-\infty < q < 0$, как и в предыдущих случаях, описывает голую сингулярность.

Положительные значения q разбиваются на два интервала $(q_-, q_1) = (0, M)$ и $(q_1, q_+) = (M, \infty)$. Этим интервалам ставятся в соответствие два конформных блока и их пространственное отражение. Поскольку конформный множитель не меняет знак, то в соответствии с правилом 6) имеются две несвязные фундаментальные области, показанные на рис. 26.6. Вторая фундаментальная область, построенная из областей типа III, получается пространственным отражением из области типа I. Как и в случае решения Рейснера–Нордстрема границы фундаментальных областей включают горизонты и их можно либо склеивать до бесконечности, либо отождествить, что приводит к универсальной накрывающей и цилиндрам, соответственно.

26.5.4 Плоскость Минковского

Предыдущие примеры демонстрируют правила построения глобальных лоренцевых поверхностей непостоянной кривизны с одним вектором Киллинга. Следующие два

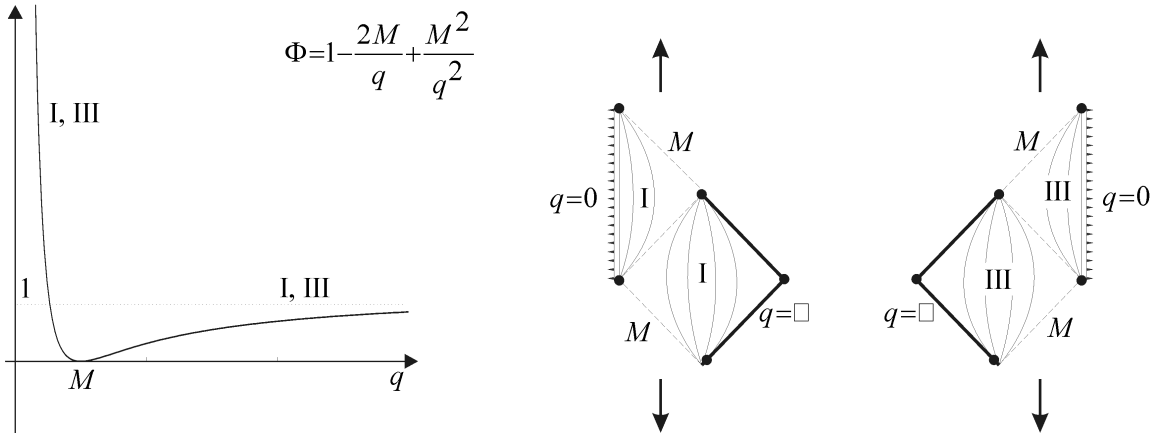


Рис. 26.6: Конформный множитель и две фундаментальные области для экстремальной черной дыры. Стрелками показаны направления, вдоль которых решения периодически продолжаются.

примера показывают, как эти правила работают в случае поверхностей постоянной кривизны, которые имеют по три вектора Киллинга. Пример плоскости Минковского важен также потому, что будет использован в дальнейшем для доказательства дифференцируемости глобальных решений в седловых точках.

Для плоскости Минковского скалярная кривизна равна нулю, а конформный множитель является линейной функцией

$$\Phi = bq + c, \qquad b, c = \mathsf{const}. \tag{26.53}$$

При этом возможны два качественно отличных случая: $b = 0$ и $b \neq 0$.

При $b = 0$ метрика (26.1) является метрикой Минковского

$$ds^2 = c(d\tau^2 - d\sigma^2), \tag{26.54}$$

где, для определенности, будем считать, что $c > 0$. Конформный множитель $\Phi = c$ не имеет ни особенностей, ни нулей. При этом уравнение (26.2) имеет общее решение

$$q = \pm c\sigma,$$

где мы отбросили несущественную постоянную интегрирования, соответствующую сдвигу σ. Знак $\pm$ соответствует различной ориентации переменной q относительно пространственной координаты. Поскольку эта переменная не входит в метрику, то ее можно исключить из рассмотрения. Поэтому интервалу $q \in (-\infty, \infty)$ ставится в соответствие один квадратный конформный блок, показанный на рис. 26.7 слева.

Теперь рассмотрим случай $b \neq 0$. Не ограничивая общности, положим $b > 0$. Тогда конформный множитель (26.53) имеет один простой нуль в точке $q_1 = -c/b$. Интервалам $(-\infty, q_1)$ и (q_1, ∞) соответствует по два квадратных конформных блока, которые склеиваются вдоль горизонтов, как показано на рис. 26.7 справа.

Покажем, как связано это глобальное решение с предыдущим представлением плоскости Минковского в виде одного конформного блока. Не ограничивая общности, положим $c = 0$, т.е.

$$ds^2 = b|q|(d\tau^2 - d\sigma^2), \qquad b > 0. \tag{26.55}$$

Этого всегда можно добиться путем сдвига $q \to q + \mathsf{const}$, который не влияет на глобальную структуру решения. Рассмотрим по очереди все четыре области.

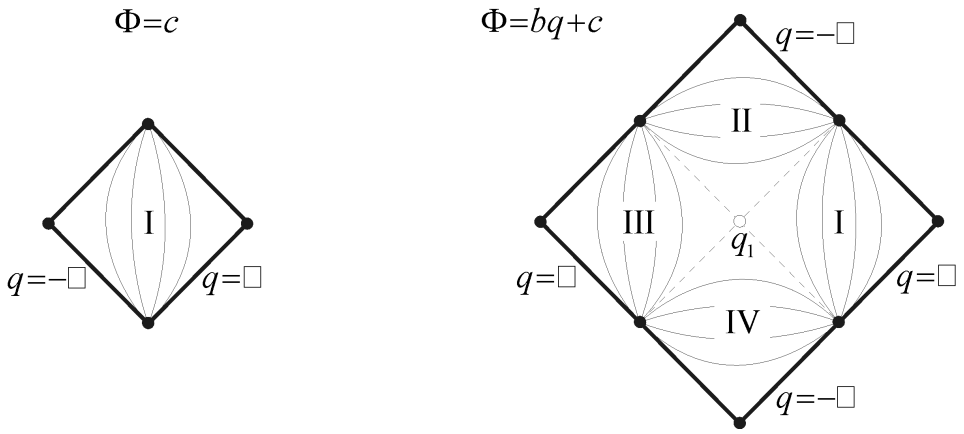


Рис. 26.7: Два представления плоскости Минковского при отсутствии и наличии одного горизонта. В последнем случае плоскость Минковского склеивается из четырех конформных блоков.

Область I. Уравнение (26.2) для $\Phi = bq$ дает

$$q = \mathsf{e}^{b\sigma} \tag{26.56}$$

с точностью до сдвига σ. Соответствующая метрика статична,

$$ds^2 = b\,\mathsf{e}^{b\sigma}(d\tau^2 - d\sigma^2), \qquad \tau, \sigma \in \mathbb{R}^2, \tag{26.57}$$

и определена на всей плоскости. Переход к конусным координатам (26.23) и конформное преобразование

$$u = \frac{2}{b}\,\mathsf{e}^{\frac{b\xi}{2}} > 0, \qquad v = -\frac{2}{b}\,\mathsf{e}^{-\frac{b\eta}{2}} < 0 \tag{26.58}$$

приводит к метрике Минковского

$$ds^2 = b\,du dv, \tag{26.59}$$

определенной в первом квадранте.

Область III. Для переменной q и метрики справедливы равенства:

$$q = \mathsf{e}^{-b\sigma}, \tag{26.60}$$

$$ds^2 = b\,\mathsf{e}^{-b\sigma}(d\tau^2 - d\sigma^2), \qquad \tau, \sigma \in \mathbb{R}^2. \tag{26.61}$$

Конформное преобразование, приводящее к метрике Минковского (26.59), имеет вид

$$u = -\frac{2}{b}\,\mathsf{e}^{-\frac{b\xi}{2}} < 0, \qquad v = \frac{2}{b}\,\mathsf{e}^{\frac{b\eta}{2}} > 0, \tag{26.62}$$

что соответствует третьему квадранту плоскости Минковского.

Область II. Эта область однородна:

$$q = -\,\mathsf{e}^{b\tau}, \tag{26.63}$$

$$ds^2 = b\,\mathsf{e}^{b\tau}(d\tau^2 - d\sigma^2), \qquad \tau, \sigma \in \mathbb{R}^2. \tag{26.64}$$

Конформное преобразование, приводящее к метрике Минковского (26.59), имеет вид

$$u = \frac{2}{b}\,\mathsf{e}^{\frac{b\xi}{2}} > 0, \qquad v = \frac{2}{b}\,\mathsf{e}^{\frac{b\eta}{2}} > 0, \tag{26.65}$$

что соответствует второму квадранту.

Область IV. Аналогично, для переменной q и метрики справедливы равенства

$$q = -\,\mathsf{e}^{-b\tau}, \tag{26.66}$$

$$ds^2 = b\,\mathsf{e}^{-b\tau}(d\tau^2 - d\sigma^2), \qquad \tau, \sigma \in \mathbb{R}^2. \tag{26.67}$$

Конформное преобразование, приводящее к метрике Минковского имеет вид

$$u = -\frac{2}{b}\,\mathsf{e}^{-\frac{b\xi}{2}} < 0, \qquad v = -\frac{2}{b}\,\mathsf{e}^{-\frac{b\eta}{2}} < 0, \tag{26.68}$$

что соответствует четвертому квадранту.

Для всех четырех областей горизонт $q_1 = 0$ соответствует координатным прямым $u = 0$ и $v = 0$. Простые вычисления показывают, что во всех областях

$$q = -\frac{b^2}{4}(t^2 - x^2), \tag{26.69}$$

где

$$u = t + x, \qquad v = t - x.$$

Эти координаты дают простейший пример координат Крускала–Секереша для метрики (26.55). То есть с точностью до знака переменная q представляет собой квадрат радиуса гиперболической полярной системы координат (2.101) на плоскости Минковского. Для гиперболического полярного угла в статичных областях

$$\varphi = \operatorname{arcth}\left(\frac{t}{x}\right),$$

справедливы равенства:

$$\mathrm{I}: \quad \varphi = \frac{b}{2}\tau, \qquad \mathrm{III}: \quad \varphi = -\frac{b}{2}\tau.$$

В однородных областях

$$\varphi = \operatorname{arcth}\left(-\frac{x}{t}\right),$$

и

$$\mathrm{II}: \quad \varphi = -\frac{b}{2}\sigma, \qquad \mathrm{IV}: \quad \varphi = \frac{b}{2}\sigma.$$

Поскольку конформные преобразования (26.58), (26.62), (26.65) и (26.68) во всех четырех областях дают одну и ту же метрику Минковского (26.59), но в разных квадрантах, то это доказывает гладкость метрики на горизонтах и в седловой точке для переменной q (26.69).

Таким образом, линейный конформный множитель (26.53) при $b \neq 0$ описывает плоскость Минковского, рис. 26.7, b слева, в гиперболической полярной системе координат.

26.5.5 Поверхности постоянной кривизны

Рассмотрим поверхности постоянной ненулевой кривизны, для которых конформный множитель является квадратичным полиномом (26.17). Для определенности, рассмотрим поверхности положительной скалярной (отрицательной гауссовой) кривизны, $a > 0$. Поверхности отрицательной скалярной кривизны получаются из поверхностей положительной кривизны преобразованием $\tau \leftrightarrow \sigma$, т.е. поворотом всех диаграмм на угол $\pi/2$.

В зависимости от значений постоянных a, b, c уравнение $\Phi(q) = 0$ может не иметь решений или иметь один или два корня. Рассмотрим последовательно все три случая, которые приводят к разным метрикам и различным диаграммам Картера–Пенроуза для поверхностей постоянной кривизны. Не ограничивая общности, положим $b = 0$, чего всегда можно добиться сдвигом переменной q.

Отсутствие горизонта. Рассмотрим конформный множитель вида

$$\Phi = -(aq^2 + c), \qquad a > 0,\ c > 0. \tag{26.70}$$

Поскольку $\Phi < 0$, а сингулярности и нули отсутствуют, то глобальное решение соответствует бесконечному интервалу $q \in (-\infty, \infty)$. Оно представляется однородным конформным блоком типа линзы II или IV, показанными на рис. 26.8 слева. В обла-

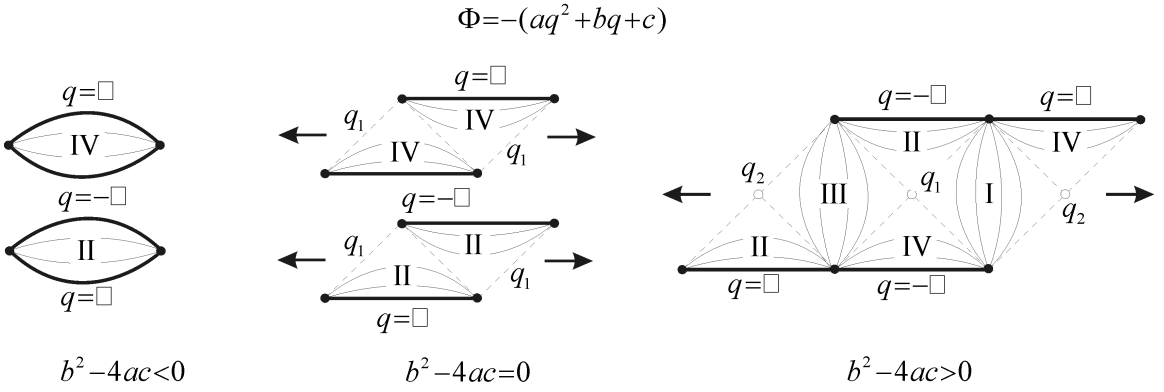


Рис. 26.8: Диаграммы Картера–Пенроуза для поверхностей постоянной положительной скалярной кривизны. Возможно три случая: отсутствие, один или два горизонта.

сти IV уравнение (26.2) для конформного множителя (26.70) приводит к равенству

$$q = \sqrt{\frac{c}{a}}\ \mathrm{tg}\,(\sqrt{ac}\,\tau) \tag{26.71}$$

с точностью до сдвига τ. Временна́я координата меняется в конечном интервале

$$-\frac{\pi}{2} < \sqrt{ac}\,\tau < \frac{\pi}{2},$$

что соответствует диаграмме типа линзы. Метрика в плоскости τ, σ имеет вид

$$ds^2 = \frac{c}{\cos^2(\sqrt{ac}\,\tau)}(d\tau^2 - d\sigma^2). \tag{26.72}$$

Эта метрика является глобальной и определена на всей универсальной накрывающей постоянной кривизны.

В области II метрика имеет такой же вид, меняется только ориентация переменной q по отношению к τ.

Один горизонт. Пусть конформный множитель имеет вид

$$\Phi = -aq^2, \qquad a > 0. \tag{26.73}$$

В точке $q_1 = 0$ он имеет нуль второго порядка, соответствующий горизонту. В области IV для обоих интервалов $q \in (-\infty, 0)$ и $q \in (0, \infty)$ получаем следующую метрику

$$q = -\frac{1}{a\tau}, \tag{26.74}$$

$$ds^2 = \frac{1}{a\tau^2}(d\tau^2 - d\sigma^2). \tag{26.75}$$

Фундаментальная область в этом случае состоит из двух треугольных конформных блоков, показанных на рис. 26.8 в центре. Чтобы получить универсальную накрывающую поверхность, диффеоморфную диаграмме типа линзы из предыдущего случая, фундаментальную область необходимо продолжить в обе стороны. Если отождествить левую и правую границы фундаментальной области, то получим поверхность, диффеоморфную однополостному гиперболоиду, рассмотренному в разделе 25.3.

Два горизонта. Пусть конформный множитель имеет вид

$$\Phi = -(aq^2 + c), \qquad a > 0,\ c < 0. \tag{26.76}$$

Он имеет два простых нуля в точках $q_{1,2} = \mp\sqrt{-c/a}$, которые соответствуют горизонтам. Поэтому трем интервалам $(-\infty, q_1)$, (q_1, q_2) и (q_2, ∞) соответствует шесть конформных блоков. Из симметрии конформного множителя $q \mapsto -q$ следует, что однородные конформные блоки II, IV и стационарные конформные блоки I, III связаны преобразованием $q \mapsto -q$. Для однородной области типа IV получаем метрику в виде

$$q = -\sqrt{-\frac{c}{a}}\ \mathsf{cth}\,(\sqrt{-ac}\,\tau), \tag{26.77}$$

$$ds^2 = \frac{-c}{\mathsf{sh}^{\,2}(\sqrt{-ac}\,\tau)}(d\tau^2 - d\sigma^2). \tag{26.78}$$

Для стационарной области I метрика имеет вид

$$q = \sqrt{-\frac{c}{a}}\ \mathsf{th}\,(\sqrt{-ac}\,\sigma), \tag{26.79}$$

$$ds^2 = \frac{-c}{\mathsf{ch}^{\,2}(\sqrt{-ac}\,\sigma)}(d\tau^2 - d\sigma^2). \tag{26.80}$$

На рис. 26.8 справа изображены конформные блоки и их склейка (диаграмма Картера–Пенроуза) для поверхностей постоянной кривизны.

Различные виды метрики при отсутствии или наличии горизонтов обусловлены различным выбором систем координат. Выпишем формулы преобразования координат, которые связывают между собой метрики в отсутствие горизонта, а также с одним и двумя горизонтами. Поскольку все метрики являются вейлевски плоскими, то связь между ними осуществляется конформным преобразованием координат. Запишем метрику с одним горизонтом (26.75) в конусных координатах (26.23) и совершим конформное преобразование $\xi \mapsto u(\xi)$, $\eta \mapsto v(\eta)$, тогда метрика примет вид

$$ds^2 = \frac{1}{a\tau^2}(d\tau^2 - d\sigma^2) = \frac{4}{a(\xi+\eta)^2}\frac{d\xi}{du}\frac{d\eta}{dv}dudv. \tag{26.81}$$

Преобразование конусных координат

$$\xi = \frac{2}{\sqrt{ac}}\,\mathsf{tg}\left(\frac{\sqrt{ac}}{2}u\right), \qquad \eta = -\frac{2}{\sqrt{ac}}\,\mathsf{ctg}\left(\frac{\sqrt{ac}}{2}v\right) \tag{26.82}$$

дает метрику (26.72), которая соответствует отсутствию горизонта. Преобразования конусных координат

$$\xi = \frac{2}{\sqrt{-ac}}\,\mathsf{th}\left(\frac{\sqrt{-ac}}{2}u\right), \qquad \eta = \frac{2}{\sqrt{-ac}}\,\mathsf{th}\left(\frac{\sqrt{-ac}}{2}v\right) \tag{26.83}$$

и

$$\xi = \frac{2}{\sqrt{-ac}}\,\mathsf{th}\left(\frac{\sqrt{-ac}}{2}u\right), \qquad \eta = -\frac{2}{\sqrt{-ac}}\,\mathsf{cth}\left(\frac{\sqrt{-ac}}{2}v\right) \tag{26.84}$$

приводят к однородной (26.78) и статической (26.80) метрикам при наличии двух горизонтов.

Рассмотренные примеры показывают, что одной и той же поверхности могут соответствовать различные вейлевски плоские метрики, связанные конформным преобразованиям координат. Однако независимо от выбора исходной метрики метод конформных блоков позволяет восстановить всю поверхность единственным образом.

26.6 Координаты Эддингтона–Финкельстейна

В теореме 26.10.1 был описан конструктивный метод конформных блоков построения максимально продолженных поверхностей с одним вектором Киллинга. Дифференцируемость метрики внутри каждого конформного блока следует из построения. Конформные блоки склеиваются вдоль горизонтов q_j (нулей конформного множителя $\Phi(q_j) = 0$) с показателями асимптотики $m = 1$ (простой нуль) и $m \geq 2$. В настоящем разделе мы докажем дифференцируемость метрики на горизонтах после склейки, записав ее в новой системе координат, которая покрывает горизонты.

Пусть внутри интервала (q_-, q_+) есть нуль q_j нечетного порядка. Для определенности, будем считать, что в соседних интервалах (q_{j-1}, q_j) и (q_j, q_{j+1}) конформный множитель положителен и отрицателен соответственно. В соответствии с правилом 4) стационарный конформный блок типа I должен быть склеен с двумя однородными блоками типа II и IV. Соответствующие горизонты можно накрыть *координатами Эддингтона–Финкельстейна* [**?**, **?**], которые вводятся следующим образом.

Склейка I–II. Для доказательства дифференцируемости метрики в склеенных областях I и II, включая горизонт, введем координаты Эддингтона–Финкельстейна $\tau, \sigma \mapsto \xi, q$. Эти координаты полностью покрывают указанные области, при этом функции перехода внутри областей I и II принадлежат классу $\mathcal{C}^{l+1}$, а метрика в новых координатах – классу $\mathcal{C}^l$.

В области I метрика для $q \in (q_{j-1}, q_j)$ имеет вид

$$\mathrm{I}: \qquad ds^2 = \Phi(q)(d\tau^2 - d\sigma^2), \qquad \frac{dq}{d\sigma} = \Phi(q) > 0. \tag{26.85}$$

перейдем от декартовых координат τ, σ к координатам Эддингтона–Финкельстейна ξ, q:

$$\tau := \xi - \int^q \frac{dr}{\Phi(r)}, \qquad \sigma := \int^q \frac{dr}{\Phi(r)}. \tag{26.86}$$

Интегралы в этих формулах на горизонте $q \to q_j$ расходятся, однако это не важно, т.к. преобразование координат (26.86) рассматривается только во внутренних точках области I. Координата $\xi = \tau + \sigma$ – это обычная светоподобная координата, а в качестве второй координаты выбрана сама переменная q. Очевидно, что во внутренних точках функции перехода принадлежат классу $\mathcal{C}^{l+1}$. Подставляя выражения для дифференциалов

$$d\tau = d\xi - \frac{dq}{\Phi}, \qquad d\sigma = \frac{dq}{\Phi}, \tag{26.87}$$

в выражение для метрики (26.85), получим квадратичную форму

$$ds^2 = \Phi d\xi^2 - 2dqd\xi. \tag{26.88}$$

Определитель этой метрики равен $\det g_{\alpha\beta} = -1$, и поэтому метрика (26.88) определена для всех $\xi \in (-\infty, \infty)$ и, что важно, при всех $q \in (q_-, q_+)$, а не только при $q \in (q_{j-1}, q_j)$. То есть метрика (26.88) определена в большей области, чем исходная метрика (26.1). Эта область – диагональная цепочка конформных блоков, идущих сверху слева вниз направо и склеенных вдоль светоподобной координаты ξ (см. рис. 26.9 справа). При этом гладкость метрики (26.88) совпадает с гладкостью конформного множителя всюду, включая горизонты.

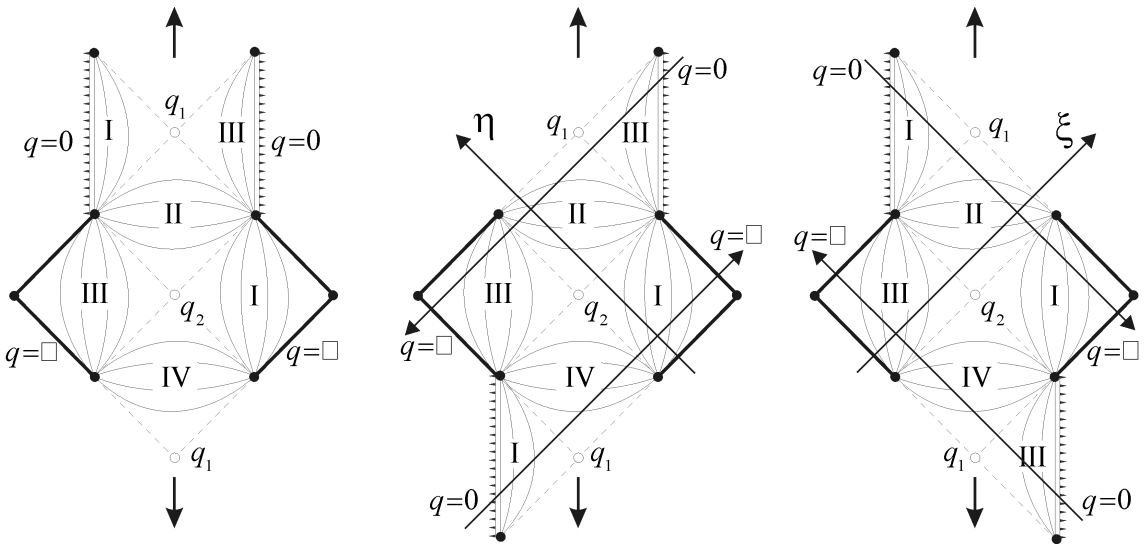


Рис. 26.9: Эквивалентные диаграммы Картера–Пенроуза для решения Рейснера–Нордстрема. Координаты Эддингтона–Финкельстейна покрывают диагональные цепочки конформных блоков, идущие снизу слева наверх направо (в центре) или снизу справа наверх налево (справа).

В однородной области для $q \in (q_j, q_{j+1})$ типа

$$\text{II}: \quad ds^2 = -\Phi(q)(d\tau^2 - d\sigma^2), \qquad \frac{dq}{d\tau} = \Phi(q) < 0, \tag{26.89}$$

перейдем к координатам Эддингтона–Финкельстейна ξ, q:

$$\tau := \int^q \frac{dr}{\Phi(r)}, \qquad \sigma := \xi - \int^q \frac{dr}{\Phi(r)}. \tag{26.90}$$

Преобразование координат в области II отличается от преобразования (26.86). Тем не менее метрика (26.89) снова принимает вид (26.88). Это означает, что координаты Эддингтона–Финкельстейна и заданная в них метрика (26.88) изометрично покрывают области I и II и, кроме того, определена на их объединении и горизонте. Это доказывает, что на горизонте дифференцируемость метрики на продолженной поверхности совпадает с дифференцируемостью конформного множителя.

Можно также проверить, что горизонт q_j в координатах Эддингтона–Финкельстейна сам является светоподобной экстремалью.

Метрика (26.88) в координатах Эддингтона–Финкельстейна имеет вектор Киллинга ∂_ξ. Квадрат его длины равен конформному множителю $(\partial_\xi, \partial_\xi) = \Phi(q)$. Поэтому на горизонте длина вектора Киллинга равна нулю.

Склейка III–IV. Области III–IV склеиваются аналогично областям I и II. Приведем явные формулы перехода к координатам Эддингтона–Финкельстейна:

$$\text{III}: \qquad \tau := \xi + \int^q \frac{dr}{\Phi(r)}, \qquad \sigma := -\int^q \frac{dr}{\Phi(r)}. \tag{26.91}$$

и

$$\text{IV}: \qquad \tau := -\int^q \frac{dr}{\Phi(r)}, \qquad \sigma := \xi + \int^q \frac{dr}{\Phi(r)}. \tag{26.92}$$

При этом в обеих областях метрика имеет одинаковый вид

$$ds^2 = \Phi d\xi^2 + 2dqd\xi. \tag{26.93}$$

Эта метрика определена на всей диагональной цепочке конформных блоков $q \in (q_-, q_+)$, идущей снизу справа наверх налево (см. рис. 26.9 справа).

Между собой метрики (26.88) и (26.93) связаны преобразованием $\xi \mapsto -\xi$.

Склейка I–IV. В области I перейдем к координатам Эддингтона–Финкельстейна η, q:

$$\text{I}: \qquad \tau := \eta + \int^q \frac{dr}{\Phi(r)}, \qquad \sigma := \int^q \frac{dr}{\Phi(r)}, \tag{26.94}$$

где $\eta = \tau - \sigma$ – светоподобная координата. В новых координатах метрика принимает вид

$$ds^2 = \Phi d\eta^2 + 2dqd\eta. \tag{26.95}$$

Эта метрика определена на всей диагональной цепочке конформных блоков ($q \in (q_-, q_+)$, идущей снизу слева наверх направо (см. рис. 26.9 в центре).

К этой же метрике в области IV приводят координаты

$$\text{IV}: \qquad \tau := -\int^q \frac{dr}{\Phi(r)}, \qquad \sigma := -\eta - \int^q \frac{dr}{\Phi(r)}. \tag{26.96}$$

Аналогичные координаты вводятся при склейке областей II и III, при этом соответствующая метрика связана с метрикой (26.95) преобразованием $\eta \mapsto -\eta$.

Если горизонт q_j имеет четный порядок, и в обоих интервалах (q_{j-1}, q_j) и (q_j, q_{j+1}) конформный множитель, например, положителен, то, согласно правилу 4), склеиваются только однотипные блоки. Для доказательства дифференцируемости склейки на блоках типа I и III вводятся координаты (26.86) и (26.91), соответственно.

Таким образом доказывается дифференцируемость метрики и, следовательно, символов Кристоффеля и тензора кривизны на всех горизонтах. Метрика в координатах Эддингтона–Финкельстейна (26.88) при $q \in (q_-, q_+)$ покрывает всю диагональную цепочку конформных блоков типа I и II для фундаментальной области. Для этой цепочки блоков координата q убывает с возрастанием времени τ. Цепочка блоков представляет собой многообразие класса $\mathcal{C}^{l+1}$, т.к. функции перехода (26.86), (26.90) принадлежат этому классу. Метрика (26.93) покрывает параллельную цепочку блоков типа III и IV. Метрика в виде (26.95) покрывает перпендикулярную диагональную цепочку конформных блоков типа I и IV. Внутри каждой цепочки метрика в координатах Эддингтона–Финкельстейна невырождена и класса $\mathcal{C}^l$.

Если внутри интервала (q_-, q_+), который соответствует глобальному решению, имеется n нулей, то для получения фундаментальной области необходимо склеить

$2(n+1)$ конформных блоков (удвоенное число интервалов (q_j, q_{j+1})). Их можно склеить в две диагональных цепочки блоков, например, идущих снизу слева наверх направо, как на рис. 26.9 в центре. В каждой цепочке склеиваются между собой либо только блоки I, IV, либо только блоки II, III. По-построению, каждая цепочка является связным многообразием. Если внутри интервала есть хотя бы один нуль нечетного порядка, то вокруг него происходит склейка цепочек, как для решения Шварцшильда. При этом достаточно одного нуля нечетного порядка, чтобы превратить фундаментальную область в связное многообразие. Если же конформный множитель внутри интервала (q_-, q_+) знак не меняет, то диагональные цепочки блоков представляет собой две компоненты связности фундаментальной области. Это доказывает утверждение в правиле 5).

Нетрудно проверить, что переход от декартовых координат τ, σ к координатам Эддингтона–Финкельстейна вырожден на всех горизонтах.

Координаты Эддингтона–Финкельстейна являются естественными в следующем смысле. Прежде всего отметим, что продолжать локальное решение надо по переменной q, т.к. именно она определяет полноту многообразия и сингулярности кривизны. Поскольку для стационарного решения переменная q зависит от σ, то вместо σ можно выбрать само q (26.86), (26.91), (26.94), при этом связь координат определяется уравнением (26.2) взаимно однозначно для каждого типа области с точностью до несущественного сдвига σ. После этого вместо временно́й координаты вводится светоподобная координата (26.23). При этом существует две возможности, и обе они реализуются, задавая координаты на взаимно перпендикулярных диагональных цепочках конформных блоков. Аналогично вводятся координаты Эддингтона–Финкельстейна на однородных блоках.

26.7 Дифференцируемость метрики в седловой точке

Координаты Эддингтона–Финкельстейна не покрывают седловых точек, расположенных в местах пересечения горизонтов. Седловые точки соответствуют нулям конформного множителя $\Phi(q_j) = 0$, расположенным в конечных точках, $|q_j| < \infty$ с показателем степени $m = 1$ или $m \geq 2$ (см. рис. 26.3). Если точка q_j является нулем кратности два или больше, то седловая точка полна, и доказывать ничего не надо. Если точка q_j – простой нуль, то можно ввести координаты, покрывающие окрестность седловой точки. Поскольку поведение конформного множителя вблизи простого нуля такое же, как и на плоскости Минковского, то перейдем к светоподобным *координатам Крускала–Секереша* $\tau, \sigma \mapsto u, v$. Явные формулы преобразования в областях I–IV имеют вид (26.58), (26.65), (26.62) и (26.68), соответственно. При этом мы положим $b := \Phi'(q_j)$ (см. раздел 26.5.4).

Рассмотрим область I более подробно. Метрика имеет вид

$$ds^2 = \Phi(d\tau^2 - d\sigma^2) = \Phi d\xi d\eta,$$

где ξ, η – светоподобные координаты (26.23). Преобразуем светоподобные координаты $\xi, \eta \mapsto u, v$ по формулам (26.58). Тогда метрика примет вид

$$ds^2 = \Phi\, \mathsf{e}^{-b\sigma} du dv = \Phi\, \mathsf{exp}\left(-b \int^q \frac{dp}{\Phi(p)}\right) du dv,$$

где мы учли связь σ с q, которая дается уравнением

$$\frac{dq}{d\sigma} = \Phi.$$

Проведем аналогичное преобразование координат во всех четырех квадрантах и положим $b := \Phi'(q_j)$. Тогда во всех четырех квадрантах метрика (26.1) примет одинаковый вид

$$ds^2 = |\Phi| \exp\left(-\int^q dp \frac{\Phi'(q_j)}{\Phi(p)}\right) dudv = \exp\left(\int^q dp \frac{\Phi'(p) - \Phi'(q_j)}{\Phi(p)}\right) dudv. \qquad (26.97)$$

Как и в случае плоскости Минковского координаты u, v покрывают окрестность седловой точки, состоящую из областей всех четырех типов, склеенных вдоль горизонтов.

Вблизи седловой точки при $m = 1$ справедливо разложение

$$\Phi(p) = \Phi_1(p - q_j) + \frac{1}{2}\Phi_2(p - q_j)^2 + \mathsf{o}\Big((p - q_j)^2\Big), \qquad \Phi_{1,2} = \mathsf{const}, \quad \Phi_1 \neq 0.$$

Поэтому

$$\lim_{p \to q_j} \frac{\Phi'(p) - \Phi'(q_j)}{\Phi(p)} = \frac{\Phi_2}{\Phi_1}.$$

Поскольку в седловой точке предел подынтегрального выражения конечен, то интеграл является достаточно гладкой функцией. Это доказывает невырожденность и дифференцируемость метрики в седловой точке, которая совпадает с дифференцируемостью конформного множителя.

Очевидно, что вблизи каждой седловой точки первого порядка можно ввести координаты Крускала–Секереша. Поскольку функции перехода к координатам Крускала–Секереша класса $\mathcal{C}^\infty$, то класс гладкости универсальной накрывающей поверхности определяется гладкостью функций перехода к координатам Эддингтона–Финкельстейна, $\mathcal{C}^{l+1}$. Тем самым доказательство основной теоремы 26.10.1 завершено.

Таким образом, мы показали, что локальное решение вида (26.1) имеет дифференцируемое максимальное продолжение. Описанная процедура является однозначной. Однако, это не доказывает единственность продолжения. Дело в том, что преобразование координат к координатам Эддингтона–Финкельстейна и Крускала–Секереша вырождено на горизонтах. Поэтому нельзя утверждать, что указанное продолжение решения вдоль экстремалей является единственным.

Метод построения глобальных решений для двумерных метрик лоренцевой сигнатуры, рассмотренный в настоящей главе, является простым и конструктивным. Если уравнения теории гравитации приводят к метрике (26.1) с некоторым конформным множителем, то для построения глобального решения достаточно проанализировать его нули и особенности. При этом восстанавливается универсальная накрывающая поверхность. Остальные глобальные решения являются факторпространствами универсальной накрывающей по группе преобразований, действующей собственно разрывно и свободно. Выше мы отметили возможность построения решений с нетривиальной фундаментальной группой путем отождествления элементов границы (горизонтов) фундаментальной области. В общем случае нахождение групп преобразований, действующих собственно разрывно и свободно, зависит от конкретного вида конформного множителя и является предметом самостоятельного исследования.

26.8 Локальный вид римановой метрики

Чтобы преобразовать лоренцеву метрику (26.1) к метрике евклидовой сигнатуры, совершим поворот в комплексной плоскости той из координат τ или σ, от которой конформный множитель не зависит. Поскольку метрика вида (26.1), по-предположению, возникает в результате решения системы уравнений движения для некоторой модели гравитации, то соответствующая риманова метрика также будет являться локальным решением той же системы уравнений после поворота в комплексной плоскости.

В явном виде комплексный поворот задается следующим образом. Введем новую координату $\tau := i\rho$ в статичных областях I, III и $\sigma := i\rho$ – в однородных областях II, IV. В результате получим метрику

$$ds^2 = -\Phi(q)(d\sigma^2 + d\rho^2), \tag{26.98}$$

где в областях II и IV мы переобозначили координату $\tau \mapsto \sigma$. Знак конформного множителя $\Phi(q)$ не фиксирован, т.е. рассматриваются как положительно, так и отрицательно определенные метрики. После комплексного поворота аргумент конформного множителя q зависит только от σ, и эта связь задается обыкновенным дифференциальным уравнением

$$\left|\frac{dq}{d\sigma}\right| = |\Phi(q)|\,. \tag{26.99}$$

Знаки модуля в этом уравнении раскрываются следующим образом:

$$\begin{array}{rlll} \text{I}: & \Phi > 0, & dq/d\sigma > 0, & \mathsf{sign}\, g_{\alpha\beta} = (--), \\ \text{II}: & \Phi < 0, & dq/d\sigma < 0, & \mathsf{sign}\, g_{\alpha\beta} = (++), \\ \text{III}: & \Phi > 0, & dq/d\sigma < 0, & \mathsf{sign}\, g_{\alpha\beta} = (--), \\ \text{IV}: & \Phi < 0, & dq/d\sigma > 0, & \mathsf{sign}\, g_{\alpha\beta} = (++), \end{array} \tag{26.100}$$

где в последней колонке показана сигнатура метрики (26.98).

Как и случае метрики лоренцевой сигнатуры, у нас есть четыре области: две области с положительно определенной и две области с отрицательно определенной метрикой. Метрики в областях I и III, а также II и IV по существу одинаковы, т.к. связаны между собой простым преобразованием $\sigma \mapsto -\sigma$.

Метрика вида (26.98) обладает по-крайней мере одним вектором Киллинга $K = \partial_\rho$, квадрат длины которого равен $(\partial_\rho, \partial_\rho) = -\Phi(q)$.

Риманова метрика, заданная уравнениями (26.98), (26.99), и будет предметом исследования в настоящем разделе. Мы допускаем, что конформный множитель может обращаться в нуль и иметь особенности в конечном числе точек q_i, $i = 1, \ldots, k$. В эту последовательность включены также бесконечно удаленные точки $q_1 = -\infty$ и $q_k = \infty$. Таким образом, вещественная ось q разбивается точками q_i на интервалы, внутри которых конформный множитель либо строго больше, либо строго меньше нуля. Будем считать, что вблизи граничных точек q_i конформный множитель ведет себя, как и в лоренцевом случае, степенным образом (26.4), (26.5). При конечных значениях q_i показатель степени не равен нулю, $m \neq 0$, так как конформный множитель в этой точке, по-предположению, либо равен нулю, либо сингулярен.

Символы Кристоффеля имеют следующие нетривиальные компоненты:

$$\text{I, II}: \qquad \Gamma_{\sigma\sigma}{}^{\sigma} = \Gamma_{\sigma\rho}{}^{\rho} = \Gamma_{\rho\sigma}{}^{\rho} = -\Gamma_{\rho\rho}{}^{\sigma} = \frac{1}{2}\Phi', \tag{26.101}$$

$$\text{III, IV}: \qquad \Gamma_{\sigma\sigma}{}^{\sigma} = \Gamma_{\sigma\rho}{}^{\rho} = \Gamma_{\rho\sigma}{}^{\rho} = -\Gamma_{\rho\rho}{}^{\sigma} = -\frac{1}{2}\Phi', \tag{26.102}$$

где штрих обозначает дифференцирование по аргументу, $\Phi' := d\Phi/dq$, и никакого суммирования по индексам ρ и σ не проводится. Тензор кривизны одинаков для всех областей и имеет только одну независимую компоненту

$$R_{\sigma\rho\sigma}{}^{\rho} = \frac{1}{2}\Phi\Phi'' \tag{26.103}$$

Тензор Риччи диагонален. Его ненулевые компоненты и скалярная кривизна равны

$$R_{\sigma\sigma} = R_{\rho\rho} = \frac{1}{2}\Phi\Phi'', \tag{26.104}$$

$$R = -\Phi''. \tag{26.105}$$

Отсюда следует, что для поверхностей постоянной кривизны, $R = \mathsf{const}$, как и в лоренцевом случае, конформный множитель Φ является квадратичным полиномом.

Из уравнения (26.105) следует, что скалярная кривизна сингулярна вблизи граничной точки q_i при тех же показателях (26.18), (26.19), что и в лоренцевом случае. В граничных точках $q_i = \pm\infty$ скалярная кривизна стремится к отличной от нуля постоянной при $m = 2$ и к нулю при $m < 2$. Отметим, что ненулевое значение кривизны в конечной точке $|q_i| < \infty$ может возникнуть и при $m = 1$ за счет поправок следующего порядка в разложении (26.4).

Значение переменной q и, значит, скалярной кривизны постоянно вдоль траекторий Киллинга $\sigma = \mathsf{const}$. В соответствии с определением областей (26.100) q монотонно возрастает по σ в областях I, IV и монотонно убывает в областях II, III.

Область определения метрики (26.98) на плоскости σ, ρ зависит от вида конформного множителя. Координата $\rho \in \mathbb{R}$ пробегает всю вещественную прямую, поскольку от нее ничего не зависит, а область изменения координаты σ определяется уравнением (26.99). Каждому интервалу переменной $q \in (q_i, q_{i+1})$ соответствует конечный, полубесконечный или бесконечный интервал координаты σ в зависимости от сходимости или расходимости интеграла

$$\sigma_{i,i+1} \sim \int^{q_i, q_{i+1}} \frac{dq}{\Phi(q)} \tag{26.106}$$

в граничных точках. Сходимость интеграла определяется показателем степени m:

$$\begin{aligned} |q_i| < \infty: &\quad \begin{cases} m < 1, & \text{сходится}, \\ m \geq 1, & \text{расходится}, \end{cases} \\ |q_i| = \infty: &\quad \begin{cases} m \leq 1, & \text{расходится}, \\ m > 1, & \text{сходится}, \end{cases} \end{aligned} \tag{26.107}$$

Если на обоих концах интервала (q_i, q_{i+1}) интеграл расходится, то координата σ пробегает всю вещественную ось, $\sigma \in (-\infty, \infty)$, и метрика определена на всей плоскости $\sigma, \rho \in \mathbb{R}^2$. Если на одном из концов q_{i+1} или q_i интеграл сходится, то метрика определена на полуплоскости $\sigma \in (-\infty, \sigma_{i+1})$ или $\sigma \in (\sigma_i, \infty)$, соответственно. При этом выбор граничных точек σ_{i+1} и σ_i произволен, и, не ограничивая общности, можно положить $\sigma_{i,i+1} = 0$. Если на обоих концах интервала интеграл сходится, то локальное решение определено в полосе $\sigma \in (\sigma_i, \sigma_{i+1})$, при этом можно приравнять нулю только один из концов интервала.

В лоренцевом случае для построения максимально продолженных поверхностей каждому интервалу (q_i, q_{i+1}) ставится в соответствие конформный блок определенной формы, которые затем склеиваются между собой. При евклидовой сигнатуре метрики в этом нет необходимости, поскольку "светоподобные" экстремали с асимптотикой $\rho = \pm\sigma + \mathsf{const}$, как будет показано ниже, в общем случае отсутствуют.

26.9 Экстремали

Для того чтобы описать максимально продолженную поверхность для римановой метрики (26.98) с одним вектором Киллинга, необходимо проанализировать поведение экстремалей $\{\zeta^\alpha(t)\} = \{\sigma(t), \rho(t)\}$, $t \in \mathbb{R}$, которые задаются системой нелинейных обыкновенных дифференциальных уравнений (26.25), где точка обозначает дифференцирование по каноническому параметру t, который определен с точностью до линейного преобразования.

Для определенности, рассмотрим область I. Используя выражение для символов Кристоффеля (26.101), получаем систему уравнений для экстремалей:

$$\ddot{\sigma} = \frac{1}{2}\Phi'(\dot{\rho}^2 - \dot{\sigma}^2), \tag{26.108}$$

$$\ddot{\rho} = -\Phi'\dot{\sigma}\dot{\rho}. \tag{26.109}$$

Эта система уравнений имеет два первых интеграла:

$$-\Phi(\dot{\sigma}^2 + \dot{\rho}^2) = C_0 = \mathsf{const}, \tag{26.110}$$

$$-\Phi\dot{\rho} = C_1 = \mathsf{const}. \tag{26.111}$$

Существование интеграла движения (26.110) позволяет выбрать длину экстремали в качестве канонического параметра. Второй интеграл (26.111) связан с наличием вектора Киллинга. На этом этапе проявляется важность наличия у метрики одного вектора Киллинга, т.к. существование двух интегралов позволяет проанализировать экстремали в общем случае, не конкретизируя вид конформного множителя $\Phi(q)$.

В системе уравнений для экстремалей (26.108), (26.109) конформный множитель Φ рассматривается, как сложная функция от σ: $\Phi = \Phi\Big(q(\sigma)\Big)$, где $q(\sigma)$ – решение уравнения (26.99). Если форма экстремали $\sigma(\rho)$ найдена, то конформный множитель можно рассматривать также как сложную функцию от ρ вдоль каждой экстремали: $\Phi = \Phi\Big(\sigma(\rho)\Big)$. В дальнейшем анализе это будет подразумеваться.

Теорема 26.9.1. *Любая экстремаль в области I принадлежит одному из следующих четырех классов.*
1. Прямые экстремали вида (аналог светоподобных экстремалей)

$$\rho = \pm\sigma + \mathsf{const}, \tag{26.112}$$

существуют только для евклидовой метрики $\Phi = \mathsf{const}$*, при этом канонический параметр можно выбрать в виде* $\sigma = t$*.*
2. Экстремали общего вида, форма которых определена уравнением

$$\frac{d\rho}{d\sigma} = \pm\frac{1}{\sqrt{-1 - C_2\Phi}}, \tag{26.113}$$

где $C_2 < 0$ *– отрицательная постоянная. Соответствующий канонический параметр определяется любым из двух уравнений:*

$$\dot{\sigma} = \pm\frac{\sqrt{-1 - C_2\Phi}}{\Phi}, \tag{26.114}$$

$$\dot{\rho} = \frac{1}{\Phi}. \tag{26.115}$$

При этом в уравнениях (26.113) и (26.114) знаки плюс или минус выбираются одновременно.

3. Прямые экстремали, параллельные оси σ и проходящие через каждую точку $\rho = \mathsf{const}$. Канонический параметр определен уравнением

$$\dot{\sigma} = \frac{1}{\sqrt{\Phi}}. \tag{26.116}$$

4. Прямые вырожденные экстремали, параллельные оси ρ и проходящие через критические точки $\sigma_0 = \mathsf{const}$, в которых

$$\Phi'(\sigma_0) = 0. \tag{26.117}$$

Канонический параметр для них можно выбрать в виде

$$t = \rho. \tag{26.118}$$

Доказательство. Почти дословное повторение доказательства теоремы 26.3.1 в лоренцевом случае. □

Постоянная C_2, как и в лоренцевом случае, определяется интегралами (26.110) и (26.111):

$$C_2 := \frac{C_0}{C_1^2}. \tag{26.119}$$

Для области I $\Phi > 0$ и, следовательно, $C_0 < 0$ и $C_2 < 0$. Из уравнения (26.113) следует, что постоянная C_2 параметризует угол, под которым экстремаль общего вида проходит через заданную точку.

Подчеркнем качественное отличие в поведении экстремалей для метрик евклидовой и лоренцевой сигнатуры. Во-первых, для римановой метрики в общем случае отсутствует аналог светоподобных экстремалей. Это важно, так как в лоренцевом случае светоподобные экстремали были неполны на горизонтах и их необходимо было продолжить. Для римановых метрик эта проблема снимается. Во-вторых, уравнения для экстремалей общего вида (26.113) и (26.114) отличаются от соответствующих уравнений в лоренцевом случае знаком перед единицей, стоящей под знаком корня. Это небольшое на первый взгляд отличие приводит к тому, что у экстремалей общего вида уже нет светоподобной асимптотики при $q \to q_i$ вблизи нулей конформного множителя $\Phi(q_i) = 0$, которые определяют горизонты.

Качественное поведение экстремалей общего вида анализируется в общем случае, для чего достаточно знать поведение конформного множителя $\Phi(q)$ вблизи граничных точек q_i. Поскольку конформный множитель в области I положителен, то экстремали общего вида существуют только при отрицательных значениях постоянной C_2, т.к. в противном случае правая часть уравнения (26.115) становится мнимой. При этом должно быть выполнено неравенство

$$\Phi(q) \geq -1/C_2.$$

При достаточно больших значениях модуля C_2 это неравенство определяет интервал изменения переменной $q \in (q', q'')$, где граничные точки q' и q'' определяются уравнением $\Phi(q) = -1/C_2$. Точкам q', q'' соответствуют некоторые точки σ' и σ''. Экстремаль общего вида не может выйти за пределы полосы $\sigma \in (\sigma', \sigma'')$, $\rho \in (-\infty, \infty)$. Несложный анализ уравнения (26.115) показывает, что экстремали общего вида осциллируют

между значениями σ' и σ'', как показано на рис.26.10 a_3, b_3. Осциллирующие экстремали всегда полны, поскольку правая часть уравнения (26.115) ограничена сверху значением $1/\Phi(q')$ или $1/\Phi(q'')$ и снизу $1/\max\Phi(q)$.

Если конформный множитель $\Phi(q)$ равен бесконечности в точке q_{i+1}, как показано на рис.26.10 b, то экстремаль общего вида может начинаться и заканчиваться на границе, где $\Phi(q_{i+1}) = \infty$. Эта граница соответствует конечному значению σ_{i+1} при $|q_{i+1}| < \infty$, $m < 1$ и $|q_{i+1}| = \infty$, $m > 1$. При этом все экстремали подходят к границе $\Phi = \infty$ под прямым углом, т.к. правая часть уравнения (26.113) стремится к нулю. Полнота экстремалей общего вида, подходящих к границе $\Phi = \infty$ определяется интегралом

$$\lim_{q\to q_{i+1}} t \to \int^{q_{i+1}} \frac{dq}{\sqrt{\Phi}}, \tag{26.120}$$

что следует из уравнений (26.114) и (26.99). Тем самым полнота экстремалей общего вида, подходящих к сингулярной границе, такая же, как и прямых экстремалей, параллельных оси σ (26.116). Они полны только при $|q_{i+1}| = \infty$ и $1 < m \leq 2$.

В лоренцевом случае в области I к границе $\Phi = \infty$ подходят только прямые и пространственноподобные экстремали (см. раздел 26.3). Их полнота связана с показателем степени m так же, как и для римановой метрики.

На рис.26.10 в верхнем ряду (a_1, b_1) показано типичное поведение конформного множителя с одним экстремумом между двумя нулями и с двумя локальными экстремумами между нулем и особенностью. В последнем случае мы предполагаем, что $|q_{i+1}| < \infty$ и $m < 0$, что соответствует сингулярной кривизне. В среднем ряду (a_2, b_2) представлена зависимость этих конформных множителей от координаты σ. На рис.26.10, b_2 значение координаты σ_{i+1} конечно, что следует из сходимости интеграла (26.106). В нижнем ряду (a_3, b_3) показано качественное поведение экстремалей на σ, ρ плоскости. Через каждый локальный экстремум конформного множителя проходит вырожденная экстремаль (d). Экстремали общего вида (o) осциллируют вблизи локального максимума $\Phi(\sigma_0)$ между значениями σ' и σ'', которые определяются постоянной C_2. Экстремали общего вида (r) могут также начинаться и заканчиваться на границе $\Phi(\sigma_{i+1}) = \infty$, имея при этом конечную длину. Существуют также экстремали общего вида (s), которые асимптотически приближаются при $\rho \to \pm\infty$ к вырожденной экстремали, проходящей через локальный минимум σ'_0. Часть из этих экстремалей заканчивается на границе $\Phi(\sigma_{i+1}) = \infty$ при конечном значении канонического параметра.

Все экстремали можно произвольно сдвигать вдоль оси ρ. Через каждую точку ρ проходит также экстремаль (p), параллельная оси σ.

Если обе граничные точки интервала q_i и q_{i+1} являются нулями, то из непрерывности конформного множителя Φ следует, что у него есть по крайней мере один экстремум, через который проходит вырожденная экстремаль (d). Вырожденные экстремали всегда полны, т.е. имеют бесконечную длину, т.к. канонический параметр совпадает с координатой ρ (26.118).

Проведенный анализ экстремалей показывает, что неполнота экстремалей в полосе $\sigma \in (\sigma_i, \sigma_{i+1})$, $\rho \in (-\infty, \infty)$ полностью определяется поведением экстремалей, параллельных оси σ, при подходе к граничным точкам σ_i, σ_{i+1}. Что, в свою очередь, определяется сходимостью интеграла (26.120). Эти экстремали неполны в конечных точках $|q_i| < \infty$ при $m < 2$. В бесконечно удаленных точках $|q_i| = \infty$ они неполны при $m > 2$ и полны во всех остальных случаях. Учитывая, что продолжать поверхность необходимо только при конечных значениях кривизны, приходим к выводу, что продолжение поверхности необходимо только в случае простого нуля, $m = 1$, у

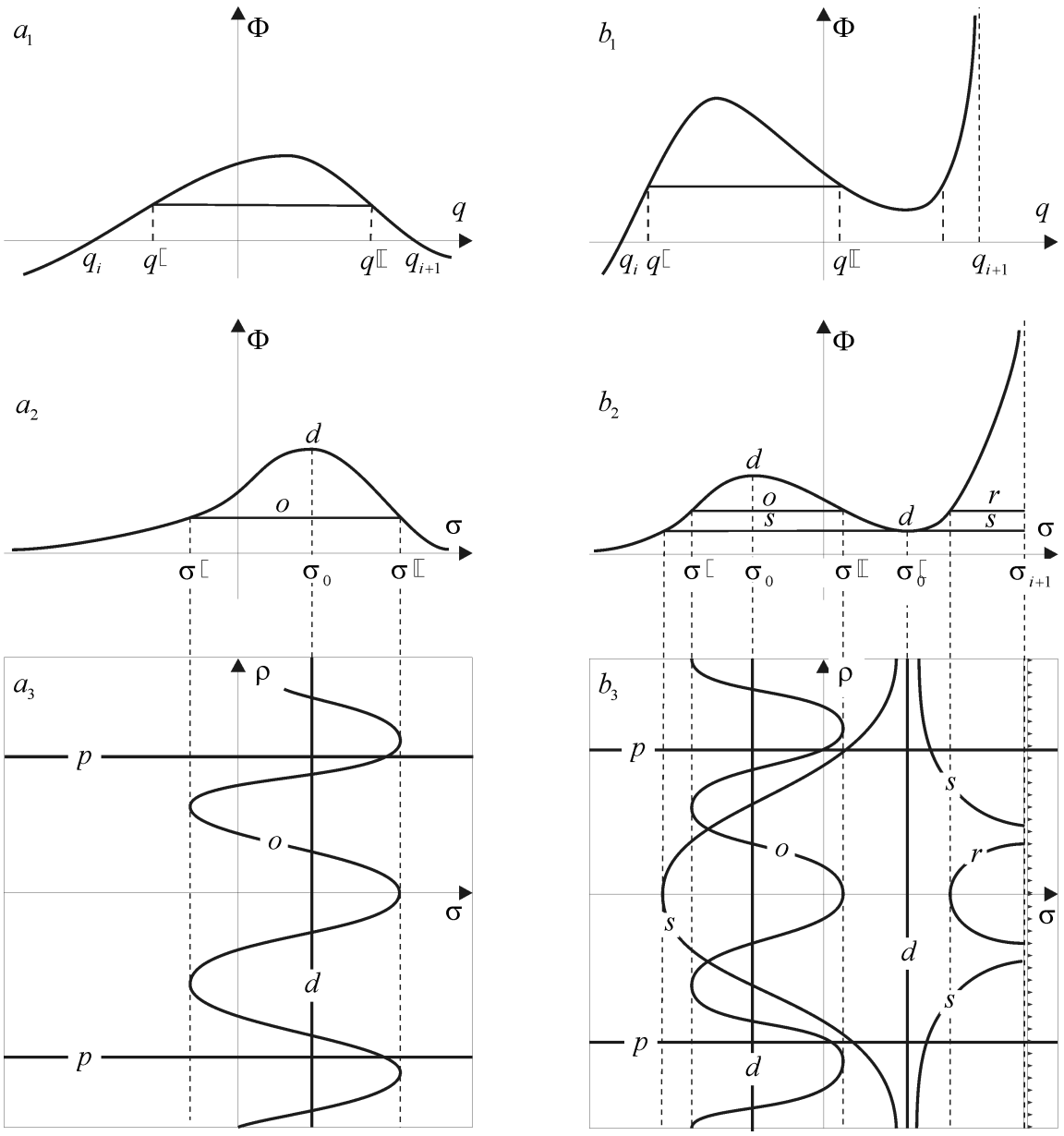


Рис. 26.10: Верхний ряд (a_1, b_1): типичное поведение конформного множителя $\Phi(q)$ между двумя нулями и нулем и сингулярностью. Средний ряд (a_2, b_2): зависимость конформного множителя $\Phi(\sigma)$ от координаты σ. Точка σ_{i+1} может находиться как на конечном, так и на бесконечном расстоянии от начала координат. Нижний ряд (a_3, b_3): типичные экстремали общего вида при разных значениях постоянной C_2. Вблизи локального максимума экстремали (o) осциллируют между σ' и σ''. Через каждый локальный экстремум проходит вырожденная экстремаль (d). Существуют экстремали (s), которые асимптотически приближаются к вырожденной экстремали в локальном минимуме σ'_0 и могут заканчиваться на сингулярной границе σ_{i+1}. Экстремали (r) конечной длины начинаются и заканчиваются на сингулярной границе. Все экстремали можно произвольно сдвигать вдоль оси ρ. Через каждую точку ρ проходит также экстремаль (p), параллельная оси σ.

$|q_i| < \infty$

	$m<0$	$0<m<1$	$m=1$	$1<m<2$	$m=2$	$m>2$
R	∞	∞	const	∞	const	0
σ_i	const	const	∞	∞	∞	∞
P^2	∞	0	0	0	0	0
Полнота	$-$	$-$	$-$	$-$	$+$	$+$

$|q_i| = \infty$

	$m<0$	$m=0$	$0<m\leq 1$	$1<m<2$	$m=2$	$m>2$
R	0	0	0	0	const	∞
σ_i	∞	∞	∞	const	const	const
P^2	0	const	∞	∞	∞	∞
Полнота	$+$	$+$	$+$	$+$	$+$	$-$

Таблица 26.1: Свойства граничных точек в зависимости от показателя степени m. Символ const в строках для скалярной кривизны и длины направляющей окружности обозначает отличную от нуля константу.

конформного множителя в конечной точке $|q_i| < \infty$. Напомним, что простой нуль конформного множителя в лоренцевом случае соответствует горизонту.

26.10 Построение глобальных решений

Чтобы построить максимальное продолжение поверхностей с римановой метрикой (26.98), имеющей по крайней мере один вектор Киллинга, произведем следующее построение, которое будет также полезно для наглядного изображения поверхностей. Поскольку конформный множитель не зависит от ρ, то можно отождествить прямые ρ и $\rho + L$, где $L > 0$ – произвольная положительная постоянная. При таком отождествлении плоскость σ, ρ сворачивается в цилиндр. Длина направляющей окружности P определяется конформным множителем и на границе $q \to \pm\infty$ равна либо постоянной, либо бесконечности:

$$P^2 = -L^2\Phi(q) \overset{q\to\pm\infty}{\longrightarrow} \begin{cases} 0, & \Phi(\pm\infty) = 0, \\ \text{const} \neq 0, & \Phi(\pm\infty) = \text{const} \neq 0, \\ \mp\infty & \Phi(\pm\infty) = \pm\infty. \end{cases} \tag{26.121}$$

В конечных точках $q \to q_i \neq \pm\infty$ квадрат длины направляющей окружности может принимать только два значения:

$$P^2 = -L^2\Phi(q) \overset{q\to q_i}{\longrightarrow} \begin{cases} 0, & \Phi(q_i) = 0, \\ \mp\infty & \Phi(q_i) = \pm\infty. \end{cases} \tag{26.122}$$

При этом плоскость σ, ρ является универсальным накрывающим пространством цилиндра.

Подытожим в таблице 26.1 свойства граничных точек q_i. В зависимости от показателя степени m здесь указаны значения скалярной кривизны R на соответствующей границе, конечность значения координаты σ_i в точке q_i, квадраты длин направляющих окружностей цилиндров P^2 (с точностью до знака) и полнота экстремалей, параллельных оси σ.

На рис. 26.11 показан вид поверхностей вблизи граничной точки q_{i+1} после отождествления $\rho \sim \rho + L$. Поверхности вблизи точки q_i имеют такой же вид, но повернуты в другую сторону. Поверхность, соответствующая всему интервалу (q_i, q_{i+1}), получается после склейки двух таких поверхностей для граничных точек q_i и q_{i+1}. Из таблицы 26.1 следует, что точка q_i является неполной по экстремалям, и при этом кривизна несингулярна в единственном случае, когда $|q_i| < \infty$ при $m = 1$. Таким образом, мы видим, что необходимо продолжить экстремали только вблизи точки $|q_i| < \infty$ при $m = 1$. Поскольку эта точка при лоренцевой сигнатуре метрики соответствует горизонту, то и в евклидовом случае оставим это же название.

Продолжение поверхности через горизонт проводится следующим образом. Прежде всего отметим, что горизонт в евклидовом случае представляет собой точку, так как длина направляющей окружности стремится к нулю. Кроме того, эта "бесконечно удаленная" точка $\sigma = \infty$ в плоскости ρ, σ на самом деле находится на конечном расстоянии, т.к. все экстремали достигают этой точки при конечном значении канонического параметра. В окрестности точки q_i с точностью до членов более высокого порядка конформный множитель имеет вид

$$\Phi(q) = \Phi'_i(q - q_i), \tag{26.123}$$

где $\Phi'_i := \Phi'(q_i) = \mathsf{const} \neq 0$. В области I при $\Phi'_i > 0$ и $q > q_i$ уравнение (26.99) легко интегрируется

$$q - q_i = e^{\Phi'_i \sigma},$$

где отброшена несущественная постоянная интегрирования, соответствующая сдвигу σ. Таким образом, граничная точка $|q_i| < \infty$ достигается при $\sigma \to -\infty$. В координатах q, ρ, которые играют роль шварцшильдовских координат в лоренцевом случае, метрика (26.98) принимает вид

$$-ds^2 = \frac{dq^2}{\Phi'_i(q - q_i)} + \Phi'_i(q - q_i)d\rho^2. \tag{26.124}$$

В полярных координатах r, φ, определенных формулами

$$q - q_i := \frac{\Phi'_i}{4} r^2, \qquad \rho := \frac{2}{\Phi'_i}\varphi, \tag{26.125}$$

метрика становится евклидовой

$$-ds^2 = dr^2 + r^2 d\varphi^2.$$

При этом полярный угол, что важно, меняется в интервале $\varphi \in [0, L\Phi'_i/2]$ при $\rho \in [0, L]$, а радиус r определен в окрестности граничной точки q_i соотношением (26.125). Рассмотренное преобразование координат отображает "бесконечно удаленную" точку $\sigma_i = -\infty$ в начало координат евклидовой плоскости. Поскольку полярный угол меняется в интервале, который в общем случае отличен от интервала $[0, 2\pi]$, то в начале координат возможно наличие конической сингулярности. Соответствующий угол дефицита равен

$$2\pi\theta := \frac{L\Phi'_i}{2} - 2\pi. \tag{26.126}$$

Таким образом, мы получили евклидову метрику на плоскости $\mathbb{R}^2$ с конической сингулярностью. При $L = 4\pi/\Phi'_i$ угол дефицита равен нулю, коническая сингулярность отсутствует, и мы получаем плоскую евклидову метрику, гладкость которой в начале

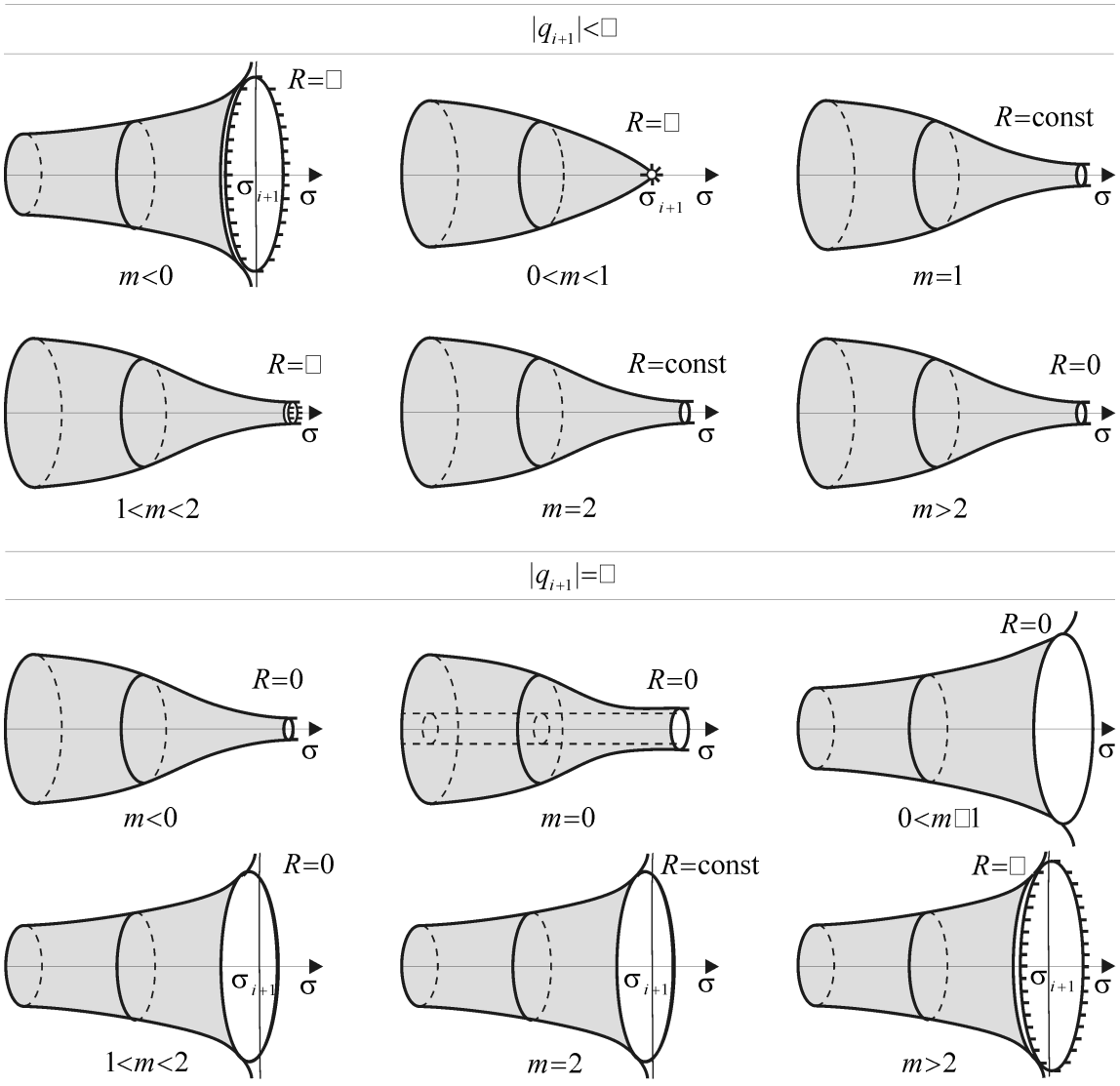


Рис. 26.11: Вид поверхности вблизи граничной точки q_{i+1} после отождествления $\rho \sim \rho + L$. Мы предполагаем, что координата σ возрастает слева направо и $\sigma_{i+1} > \sigma_i$. Поверхности вблизи точки q_i имеют такой же вид, но повернуты в другую сторону. Поверхность для интервала q_i, q_{i+1} получается после склейки двух таких поверхностей для точек q_i и q_{i+1}.

координат очевидна. В общем случае конформный множитель (26.123) вблизи граничной точки q_i содержит поправки более высокого порядка, и переход к полярным координатам при нулевом угле дефицита дает метрику того же класса гладкости, что и конформный множитель.

Таким образом, в общем случае продолжение решения через точку $|q_i| < \infty$ при $m = 1$ не имеет смысла, так как эта точка соответствует конической сингулярности. Мы считаем, что точка, в которой сосредоточена коническая сингулярность, как и носитель всякой другой сингулярности, не принадлежит многообразию. Поэтому универсальной накрывающей для такого решения будет плоскость σ, ρ или ее часть с метрикой (26.98). В исключительном случае отсутствия конической сингулярности при $L = 4\pi/\Phi_i'$ продолжение необходимо. При этом прямые экстремали, параллельные оси σ и проходящие через точки ρ и $\rho + L/2$, представляют собой две половины одной экстремали, как показано на рис.26.12. При отсутствии конической

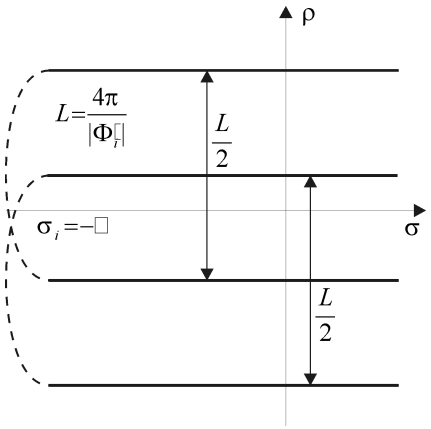


Рис. 26.12: Продолжение прямых экстремалей, проходящих через точки ρ и $\rho + L/2$ в отсутствие конической сингулярности $L = 4\pi/|\Phi_i'|$. Отождествление происходит в точке $\sigma_i = -\infty$.

сингулярности фундаментальная группа поверхности тривиальна, и поэтому соответствующая поверхность сама представляет собой универсальную накрывающую.

Если на обоих концах интервала (q_i, q_{i+1}) конформный множитель имеет асимптотики $\Phi \sim (q - q_i)$ и $\Phi \sim (q_{i+1} - q)$, то поверхность необходимо продолжить в обеих точках $\sigma = \pm\infty$. После отождествления $\rho \sim \rho + L$ в общем случае в граничных точках возникают конические сингулярности. Если $L\Phi_i' = 4\pi$ и $L\Phi_{i+1}' = 4\pi$, то конические сингулярности отсутствуют. На рис.26.13 показаны три возможных вида глобальных поверхностей, соответствующих наличию или отсутствию конических сингулярностей. Эти поверхности имеют топологию цилиндра, плоскости и сферы, соответственно.

Сформулируем правила построения максимально продолженных решений с метрикой вида (26.98)

1. После отождествления $\rho \sim \rho + L$ каждому интервалу (q_i, q_{i+1}) соответствует максимально продолженное решение, которое получается склеиванием двух поверхностей, изображенных на рис.26.11 и соответствующих граничным точкам q_i и q_{i+1}.

2. Во всех случаях, кроме отсутствия конической сингулярности $|q_i| < \infty$, $L\Phi_i' \neq 4\pi$ или $|q_{i+1}| < \infty$, $L|\Phi_{i+1}'| \neq 4\pi$, полоса $\sigma \in (\sigma_i, \sigma_{i+1})$, $\rho \in \mathbb{R}$ с метрикой (26.98)

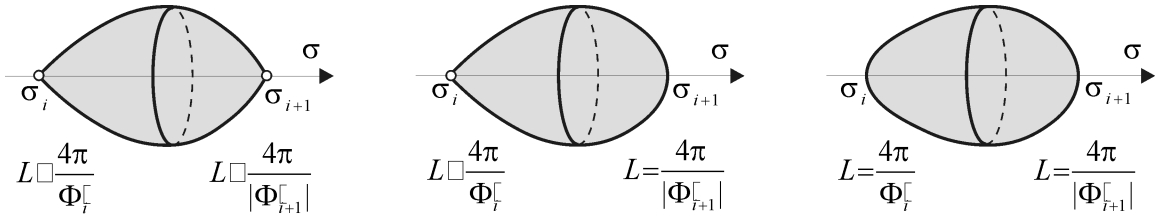


Рис. 26.13: Три возможных вида поверхностей, соответствующих интервалу (q_i, q_{i+1}), когда конформный множитель имеет асимптотику $\Phi \sim q - q_{i,i+1}$ в граничных точках. При отождествлении $\rho \sim \rho + L$ в общем случае возникают две конические сингулярности. При $L|\Phi'_{i,i+1}| = 4\pi$ конические сингулярности отсутствуют.

является универсальным накрывающим пространством для соответствующего максимально продолженного решения.

3. При отсутствии одной из конических сингулярностей, $|q_i| < \infty$, $L\Phi'_i = 4\pi$ или $|q_{i+1}| < \infty$, $L|\Phi'_{i+1}| = 4\pi$, поверхность, полученная из плоскости σ, ρ путем отождествления $\rho \sim \rho + L$, представляет собой максимально продолженное решение с тривиальной фундаментальной группой.

Отметим, что в общем случае при построении максимально продолженных решений нам не приходилось склеивать карты. Это значит, что соответствующие поверхности принадлежат классу $\mathcal{C}^\infty$. При отсутствии конической сингулярности переход к полярным координатам (26.125) не использует явный вид конформного множителя, и возникающая поверхность так же, как и евклидова плоскость является бесконечно дифференцируемым многообразием. Тем самым мы доказали теорему, оправдывающую приведенные выше правила построения глобальных решений.

Теорема 26.10.1. *Универсальное накрывающее пространство, построенное по правилам 1–3, является максимально продолженным гладким многообразием класса $\mathcal{C}^\infty$ с римановой метрикой класса $\mathcal{C}^l$, $l \geq 2$, таким, что область вне горизонта изометричная (части) плоскости ρ, σ с метрикой (26.98).*

Таким образом построены глобальные решения для широкого класса двумерных метрик евклидовой сигнатуры, которые имеют один вектор Киллинга. Доказана гладкость возникающих двумерных поверхностей и дифференцируемость метрики. Этот метод конструктивен и является аналогом метода конформных блоков для двумерных метрик лоренцевой сигнатуры. Почти всегда каждый конформный блок представляет собой универсальное накрывающее пространство для максимально продолженной римановой поверхности. Исключение составляют поверхности с горизонтом, который в евклидовом случае представляет собой точку.

26.11 Решение Шварцшильда

В качестве примера рассмотрим решение Шварцшильда. Для этого решения t, r часть метрики имеет вид

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2), \tag{26.127}$$

где

$$\Phi(q) = 1 - \frac{2M}{q}, \qquad q = \begin{cases} r, & \Phi(q) > 0 \\ t, & \Phi(q) < 0. \end{cases}$$

При этом связь переменной q с одной из координат τ или σ задается дифференциальным уравнением (26.2). Максимально продолженное решение Шварцшильда имеет вид топологического произведения двух поверхностей $\mathbb{U} \times \mathbb{S}^2$, где $\mathbb{S}^2$ – сфера и $\mathbb{U}$ – двумерная лоренцева поверхность, которая наглядно изображается в виде диаграммы Картера–Пенроуза, рис.26.14 слева.

Диаграмма Картера–Пенроуза после изменения сигнатуры метрики распадается на четыре несвязные между собой поверхности, которые показаны на рис.26.14 справа. Областям вне горизонта черной дыры соответствуют две римановы поверхности I и III с отрицательно определенной метрикой. Четырехмерная метрика является отрицательно определенной, $\mathsf{sign}\, g_{\mu\nu} = (----)$. Отрицательная определенность евклидовой метрики связана с выбором сигнатуры для метрики Шварцшильда (26.127) и может быть изменена. Это решение обычно рассматривается, как евклидова версия решения Шварцшильда. Области под горизонтами черной и белой дыр соответствуют поверхностям II и IV соответственно. При этом сигнатура полной метрики равна $\mathsf{sign}\, g_{\mu\nu} = (++--)$.

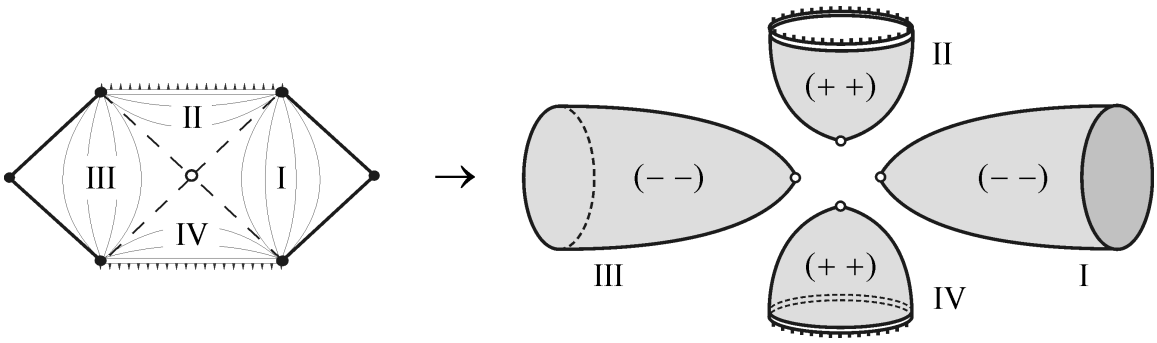


Рис. 26.14: Диаграмма Картера–Пенроуза для решения Шварцшильда при евклидовой сигнатуре метрики распадается на четыре несвязных между собой поверхности. Областям вне горизонта соответствуют две поверхности с отрицательно определенной метрикой I, III, а областям внутри горизонта – две поверхности с положительно определенной метрикой.

Ситуация с евклидовой версией решения Шварцшильда типична. Связная диаграмма Картера–Пенроуза в лоренцевом случае разрезается вдоль всех горизонтов и переходит в несвязные римановы поверхности с отрицательно или положительно определенной метрикой. При этом переход через горизонт с нечетным m соответствует изменению знака римановой метрики.

Глава 27

Сплетенные решения в общей теории относительности

В настоящей главе построены глобальные решения вакуумных уравнений общей теории относительности (21.4) с космологической постоянной в предположении, что четырехмерное пространство-время является *сплетенным произведением* двух поверхностей. При этом не делается никаких предположений о симметрии решений. Как следствие уравнений движения по крайней мере одна из двух поверхностей должна быть поверхностью постоянной кривизны. Отсюда вытекает, что метрика имеет по крайней мере три вектора Киллинга. Другими словами, свойства симметрии решений при таком подходе являются следствием самих уравнений движения. Построенные решения включают, в частности, сферически симметричные решения, которые соответствуют произведению некоторой лоренцевой поверхности на сферу. Многие глобальные решения имеют интересную физическую интерпретацию. В частности, построены решения, описывающие кротовые норы, доменные стенки сингулярностей кривизны, космические струны, комические струны, окруженные доменными стенками, решения с замкнутыми времениподобными кривыми и др.

27.1 Сплетенное произведение

Дадим

Определение. Пусть задано два многообразия $\mathbb{M}_1$ и $\mathbb{M}_2$ с метриками g и h соответственно. Касательное пространство в каждой точке топологического произведения $(x_1, x_2) \in \mathbb{M}_1 \times \mathbb{M}_2$ разлагается в прямую сумму:

$$\mathbb{T}_{(x_1,x_2)}(\mathbb{M}_1 \times \mathbb{M}_2) = \mathbb{T}_{(x_1,x_2)}(\mathbb{M}_1) \oplus \mathbb{T}_{(x_1,x_2)}(\mathbb{M}_2).$$

Сплетенным произведением (warped product) двух многообразий называется их топологическое произведение $\mathbb{M}_1 \times \mathbb{M}_2$ с метрикой $\hat{g}$, которая определена следующим соотношением

$$\hat{g}(X, Y) := k(x_2)g(X_1, Y_1) + m(x_1)h(X_2, Y_2), \tag{27.1}$$

где $k(x_2) \in \mathcal{C}^{n_2}(\mathbb{M}_2)$ и $m(x_1) \in \mathcal{C}^{n_1}(\mathbb{M}_1)$ – достаточно гладкие отличные от нуля функции на многообразиях $\mathbb{M}_2$ и $\mathbb{M}_1$, и

$$\mathbb{T}(\mathbb{M}_1 \times \mathbb{M}_2) \ni X = X_1 \oplus X_2 \in \mathbb{T}(\mathbb{M}_1) \oplus \mathbb{T}(\mathbb{M}_2),$$
$$\mathbb{T}(\mathbb{M}_1 \times \mathbb{M}_2) \ni Y = Y_1 \oplus Y_2 \in \mathbb{T}(\mathbb{M}_1) \oplus \mathbb{T}(\mathbb{M}_2),$$

– разложение векторных полей X, Y, касательных к $\mathbb{M}_1 \times \mathbb{M}_2$, в прямую сумму. □

Предположим, что четырехмерное пространство-время является сплетенным произведением двух поверхностей: $\mathbb{M} = \mathbb{U} \times \mathbb{V}$, где $\mathbb{U}$ – поверхность с лоренцевой метрикой g и $\mathbb{V}$ – поверхность с римановой метрикой h. Обозначим локальные координаты на $\mathbb{U}$ и $\mathbb{V}$ соответственно через x^α $(\alpha, \beta, \ldots = 0, 1)$ и y^μ $(\mu, \nu, \ldots = 2, 3)$. Тогда топологическому произведению $\mathbb{U} \times \mathbb{V}$ соответствуют координаты $\{x^i\} := \{x^\alpha, y^\mu\}$ $(i, j, \ldots = 0, 1, 2, 3)$. В этой системе координат четырехмерная метрика имеет блочно-диагональный вид:

$$\widehat{g}_{ij} = \begin{pmatrix} k(y) g_{\alpha\beta}(x) & 0 \\ 0 & m(x) h_{\mu\nu}(y) \end{pmatrix}, \tag{27.2}$$

где $k(y)$ и $m(x)$ – достаточно гладкие отличные от нуля функции на $\mathbb{V}$ и $\mathbb{U}$ соответственно.

В настоящей главе шляпка над символом означает, что соответствующий геометрический объект относится ко всему четырехмерному пространству-времени $\mathbb{M}$, а символы без шляпки относятся к двумерным поверхностям $\mathbb{U}$ или $\mathbb{V}$. Соответственно $g_{\alpha\beta}$ и $h_{\mu\nu}$ являются метриками на $\mathbb{U}$ и $\mathbb{V}$. Греческие буквы из начала $(\alpha, \beta, \ldots)$ и середины $(\mu, \nu, \ldots)$ алфавита всегда относятся к координатам на первой и второй поверхностям соответственно.

В физике функции $k(y)$ и $m(x)$ часто называют (*дилатонными*) полями на поверхностях $\mathbb{V}$ и $\mathbb{U}$.

Для определенности будем считать поверхность $\mathbb{U}$ псевдоримановым многообразием с метрикой лоренцевой сигнатуры, а поверхность $\mathbb{V}$ – римановым многообразием с положительно определенной метрикой. Тогда с точностью до перестановки первых двух координат сигнатура метрики на $\mathbb{M}$ будет либо $(+\,-\,-\,-)$, либо $(-\,+\,+\,+)$ в зависимости от знака m. Эти метрики связаны между собой инверсией $\widehat{g}_{ij} \to -\widehat{g}_{ij}$, относительно которой уравнения Эйнштейна при отсутствии полей материи и космологической постоянной инвариантны. Предположим также, что обе поверхности являются ориентируемыми.

Отметим, что относительно вида метрики (27.2) не делается никаких дополнительных предположений, связанных с симметрией. Однако в дальнейшем мы увидим, что уравнения Эйнштейна и требование полноты многообразий приводят к тому, что по крайней мере одна из поверхностей $\mathbb{U}$ или $\mathbb{V}$ должна быть поверхностью постоянной кривизны. То есть любое максимально продолженное решение уравнений Эйнштейна вида (27.2) допускает по крайней мере три вектора Киллинга. Следовательно, в рассматриваемом случае симметрия решений является следствием уравнений движения. В частном случае будут получены сферически симметричные решения, когда поверхность $\mathbb{V}$ является сферой $\mathbb{S}^2$.

Все решения, рассмотренные в настоящей главе относятся к решениям типа D по классификации Петрова [**?**]. Физическая интерпретация решений опирается на глобальную структуру пространства-времени.

27.2 Двумерная редукция

Вид метрики (27.2) позволяет решить явно четырехмерные вакуумные уравнения Эйнштейна с космологической постоянной Λ

$$\widehat{R}_{ij} = \Lambda \widehat{g}_{ij}, \tag{27.3}$$

и построить глобальные (максимально продолженные) решения.

Мы увидим, что уравнения Эйнштейна существенно ограничивают дилатонные поля: по крайней мере одно дилатонное поле должно быть постоянно. Поэтому все решения делятся на три основных класса: оба дилатонных поля постоянны (случай A), только $k = \mathsf{const}$ (случай B) или $m = \mathsf{const}$ (случай C). В первом случае решение уравнений Эйнштейна представляет собой топологическое произведение двух поверхностей постоянной кривизны. В случае B риманова поверхность $\mathbb{V}$ должна быть поверхностью постоянной кривизны. Сюда входят сферически симметричные решения, а также другие решения, когда поверхность $\mathbb{V}$ представляет собой евклидову плоскость или плоскость Лобачевского (двуполостный гиперболоид). Последние решения соответствуют кротовым норам. В случае C поверхность $\mathbb{U}$ должна быть поверхностью постоянной кривизны. Эти решения описывают космические струны и доменные стенки сингулярности кривизны.

Приступим к решению уравнений Эйнштейна (27.3). Метрика, обратная к (27.2), имеет вид

$$\widehat{g}^{ij} = \begin{pmatrix} \dfrac{1}{k} g^{\alpha\beta} & 0 \\ 0 & \dfrac{1}{m} h^{\mu\nu} \end{pmatrix}, \tag{27.4}$$

где $g^{\alpha\beta}$ и $h^{\mu\nu}$ матрицы, обратные соответственно к $g_{\alpha\beta}$ и $h_{\mu\nu}$. Символы Кристоффеля (7.22) равны

$$\begin{aligned}
\widehat{\Gamma}_{\alpha\beta}{}^{\gamma} &= \Gamma_{\alpha\beta}{}^{\gamma}, \\
\widehat{\Gamma}_{\alpha\beta}{}^{\mu} &= -\frac{1}{2} g_{\alpha\beta} \frac{h^{\mu\nu}\partial_\nu k}{m}, \\
\widehat{\Gamma}_{\alpha\mu}{}^{\gamma} &= \widehat{\Gamma}_{\mu\alpha}{}^{\gamma} = \frac{1}{2}\delta^{\gamma}_{\alpha} \frac{\partial_\mu k}{k}, \\
\widehat{\Gamma}_{\alpha\mu}{}^{\nu} &= \widehat{\Gamma}_{\mu\alpha}{}^{\nu} = \frac{1}{2}\delta^{\nu}_{\mu} \frac{\partial_\alpha m}{m}, \\
\widehat{\Gamma}_{\mu\nu}{}^{\alpha} &= -\frac{1}{2} h_{\mu\nu} \frac{g^{\alpha\beta}\partial_\beta m}{k}, \\
\widehat{\Gamma}_{\mu\nu}{}^{\rho} &= \Gamma_{\mu\nu}{}^{\rho}.
\end{aligned} \tag{27.5}$$

Компоненты тензора Риччи (7.77) принимают вид

$$\begin{aligned}
\widehat{R}_{\alpha\beta} &= R_{\alpha\beta} + \frac{\nabla_\alpha\nabla_\beta m}{m} - \frac{\nabla_\alpha m \nabla_\beta m}{2m^2} + \frac{g_{\alpha\beta}\nabla^2 k}{2m} \\
\widehat{R}_{\alpha\mu} &= \widehat{R}_{\mu\alpha} = -\frac{\nabla_\alpha m \nabla_\mu k}{2mk} \\
\widehat{R}_{\mu\nu} &= R_{\mu\nu} + \frac{\nabla_\mu\nabla_\nu k}{k} - \frac{\nabla_\mu k \nabla_\nu k}{2k^2} + \frac{h_{\mu\nu}\nabla^2 m}{2k},
\end{aligned} \tag{27.6}$$

где, для краткости, введены обозначения

$$\nabla^2 m := g^{\alpha\beta}\nabla_\alpha\nabla_\beta m, \qquad \nabla^2 k := h^{\mu\nu}\nabla_\mu\nabla_\nu k. \tag{27.7}$$

Здесь и далее в этой главе символ ∇ обозначает ковариантную производную с символами Кристоффеля. Соответствующая скалярная кривизна равна

$$\widehat{R} = \frac{1}{k}R^g + 2\frac{\nabla^2 m}{km} - \frac{(\nabla m)^2}{2km^2} + \frac{1}{m}R^h + 2\frac{\nabla^2 k}{km} - \frac{(\nabla k)^2}{2k^2 m}, \tag{27.8}$$

где введены обозначения

$$(\nabla m)^2 := g^{\alpha\beta}\partial_\alpha m\partial_\beta m, \qquad (\nabla k)^2 := h^{\mu\nu}\partial_\mu k\partial_\nu k. \tag{27.9}$$

Скалярные кривизны поверхностей $\mathbb{U}$ и $\mathbb{V}$ обозначены через R^g и R^h соответственно.

Таким образом уравнений Эйнштейна (27.3) для метрики (27.2) принимают вид

$$R_{\alpha\beta} + \frac{\nabla_\alpha\nabla_\beta m}{m} - \frac{\nabla_\alpha m\nabla_\beta m}{2m^2} + \frac{1}{2}g_{\alpha\beta}\left(\frac{\nabla^2 k}{m} - 2k\Lambda\right) = 0, \tag{27.10}$$

$$R_{\mu\nu} + \frac{\nabla_\mu\nabla_\nu k}{k} - \frac{\nabla_\mu k\nabla_\nu k}{2k^2} + \frac{1}{2}h_{\mu\nu}\left(\frac{\nabla^2 m}{m} - 2m\Lambda\right) = 0, \tag{27.11}$$

$$\frac{\nabla_\alpha m\nabla_\mu k}{mk} = 0. \tag{27.12}$$

Перепишем уравнения (27.10) и (27.11) в более удобном виде, выделив из них след, который определяет скалярные кривизны поверхностей:

$$R^g + \frac{\nabla^2 m}{m} - \frac{(\nabla m)^2}{2m^2} + \frac{\nabla^2 k}{m} - 2k\Lambda = 0, \tag{27.13}$$

$$R^h + \frac{\nabla^2 k}{k} - \frac{(\nabla k)^2}{2k^2} + \frac{\nabla^2 m}{k} - 2m\Lambda = 0. \tag{27.14}$$

Бесследовые части уравнений (27.10) и (27.11), умноженные на m и k, принимают простой вид

$$\nabla_\alpha\nabla_\beta m - \frac{\nabla_\alpha m\nabla_\beta m}{2m} - \frac{1}{2}g_{\alpha\beta}\left[\nabla^2 m - \frac{(\nabla m)^2}{2m}\right] = 0, \tag{27.15}$$

$$\nabla_\mu\nabla_\nu k - \frac{\nabla_\mu k\nabla_\nu k}{2k} - \frac{1}{2}h_{\mu\nu}\left[\nabla^2 k - \frac{(\nabla k)^2}{2k}\right] = 0. \tag{27.16}$$

Они не содержат слагаемые с кривизной вовсе, потому что в двух измерениях тензор Риччи полностью определяется скалярной кривизной (7.77) и не имеет бесследовой части.

Отметим, что отсутствие сингулярности у скалярной кривизны на поверхности означает ее отсутствие в полном тензоре кривизны.

Таким образом, четырехмерные уравнения Эйнштейна (27.3) для метрики вида (27.2) эквивалентны системе уравнений (27.12)–(27.16). Уравнения (27.15) и (27.16) содержат функции, зависящие только от координат x и y соответственно. В то же время координаты различных поверхностей в уравнениях (27.12), (27.13) и (27.14) перемешаны.

Уравнение (27.12) накладывает жесткие ограничения. Как следствие имеем, что либо поле дилатона k, либо поле дилатона m, либо k и m одновременно должны быть постоянны. Соответственно, возможны три случая:

$$\begin{array}{lll} \mathsf{A}: & k = \mathsf{const} \neq 0, & m = \mathsf{const} \neq 0, \\ \mathsf{B}: & k = \mathsf{const} \neq 0, & \nabla_\alpha m \neq 0, \\ \mathsf{C}: & \nabla_\mu k \neq 0, & m = \mathsf{const} \neq 0. \end{array} \tag{27.17}$$

Рассмотрим эти случаи подробно.

27.3 Произведение поверхностей постоянной кривизны

Наиболее симметричные решения вакуумных уравнений Эйнштейна (27.3) в виде топологического произведения двух поверхностей постоянной кривизны возникают, когда оба дилатонных поля k и m постоянны (случай A в (27.17)). Если k и m постоянны, то уравнения (27.12), (27.15), и (27.16) удовлетворяются, и скалярные кривизны обеих поверхностей $\mathbb{U}$ и $\mathbb{V}$ должны быть постоянны как следствие уравнений (27.13), (27.14), которые принимают следующий вид

$$R^g = 2k\Lambda, \qquad R^h = 2m\Lambda. \tag{27.18}$$

Если $\Lambda = 0$, то обе поверхности $\mathbb{U}$ и $\mathbb{V}$ имеют нулевую кривизну, и все пространство-время $\mathbb{M}$ представляет пространство Минковского или его фактор-пространство по группе преобразований, действующей свободно и собственно разрывно, (цилиндр или тор) с метрикой Лоренца

$$\widehat{g}_{ij} = \mathsf{diag}\,(+--\,-) \quad \text{или} \quad \widehat{g}_{ij} = \mathsf{diag}\,(-+\,+\,+). \tag{27.19}$$

При ненулевой космологической постоянной $\Lambda \neq 0$ обе поверхности $\mathbb{U}$ и $\mathbb{V}$ имеют постоянную ненулевую кривизну. Поверхности постоянной кривизны были рассмотрены в главе 25, и мы приведем лишь выражения для соответствующих метрик. Если $\mathbb{U}$ является полной псевдоримановой поверхностью ненулевой кривизны $R^g = -2K = \mathsf{const}$, то она представляет собой однополостный гиперболоид $\mathbb{L}^2$, вложенный в трехмерное пространство Минковского, с индуцированной метрикой или его универсальной накрывающей (раздел 25.3). Его группой симметрии является группа Лоренца $\mathbb{SO}(1,2)$. В стереографических координатах метрика однополостного гиперболоида $\mathbb{L}^2$ имеет хорошо известный вид (25.36)

$$ds^2_{\mathbb{L}} = g_{\alpha\beta}dx^\alpha dx^\beta = \frac{dt^2 - dx^2}{\left[1 + \frac{K}{4}(t^2 - x^2)\right]^2}, \tag{27.20}$$

где введены обозначения $t := x^0$ и $x := x^1$. В отличие от риманова случая псевдориманова поверхность постоянной кривизны одна и та же как для положительной, так и для отрицательной гауссовой кривизны K, при этом меняется только общий знак метрики (27.20), что соответствует перестановке координат $t \leftrightarrow x$.

При $K = 0$ метрика (27.20) совпадает с обычной двумерной метрикой Минковского, и соответствующая поверхность представляет собой плоскость Минковского $\mathbb{R}^{1,1}$ с группой Пуанкаре $\mathbb{IO}(1,1)$ в качестве группы симметрии.

Положительно определенная метрика на двумерной римановой поверхности постоянной кривизны $R^h = -2K \neq 0$ в стереографических координатах имеет вид (25.24)

$$ds^2_{\mathbb{H}} = h_{\mu\nu}dy^\mu dy^\nu = \frac{dy^2 + dz^2}{\left[1 + \frac{K}{4}(y^2 + z^2)\right]^2}, \tag{27.21}$$

где $y = y^1$ и $z = y^2$. Эта метрика отличается от (27.20) только знаками.

Для положительных $K > 0$ она соответствует сфере $\mathbb{S}^2$, рассмотренной в разделе 25.1. При $K = 0$ метрика (27.21) соответствует евклидовой плоскости $\mathbb{R}^2$, или цилиндру, или тору. При отрицательных $K < 0$ мы имеем плоскость Лобачевского (гиперболическую плоскость) $\mathbb{H}^2$, рассмотренную в разделе 25.2, или компактную

риманову поверхность рода два или выше. Группами симметрии сферы $\mathbb{S}^2$, евклидовой плоскости $\mathbb{R}^2$, и плоскости Лобачевского $\mathbb{H}^2$ являются, соответственно, группы $\mathbb{O}(3)$, $\mathbb{IO}(2)$ и $\mathbb{O}(1,2)$.

При ненулевой постоянной кривизне всегда можно произвести растяжку координат таким образом, чтобы $K = \pm 1$.

Если скалярные кривизны постоянны (27.18), то решение для ненулевой космологической постоянной $\Lambda \neq 0$ является топологическим произведением двух поверхностей постоянной кривизны с метрикой

$$ds^2 = k\frac{dt^2 - dx^2}{\left[1 - \frac{k\Lambda}{4}(t^2 - x^2)\right]^2} + m\frac{dy^2 + dz^2}{\left[1 - \frac{m\Lambda}{4}(y^2 + z^2)\right]^2}. \tag{27.22}$$

В данном случае можно не говорить о сплетенном произведении поверхностей, т.к. дилатонные поля постоянны. Растягивая координаты, всегда можно добиться выполнения равенств $k = \pm 1$, $m = \pm 1$. Выберем $k = 1$ и $m = -1$ с тем, чтобы метрика имела сигнатуру $(+---)$. Тогда возможны три качественно отличных случая, соответствующих положительной, нулевой и отрицательной космологической постоянной:

$$\begin{array}{lllll} \Lambda < 0: & K^g = +|\Lambda|, & K^h = -|\Lambda|, & \mathbb{M} = \mathbb{L}^2 \times \mathbb{H}^2, \\ \Lambda = 0: & K^g = 0, & K^h = 0, & \mathbb{M} = \mathbb{R}^{1,1} \times \mathbb{R}^2 = \mathbb{R}^{1,3}, \\ \Lambda > 0: & K^g = -|\Lambda|, & K^h = +|\Lambda|, & \mathbb{M} = \mathbb{L}^2 \times \mathbb{S}^2, \end{array} \tag{27.23}$$

где K^g и K^h – гауссовы кривизны соответственно поверхностей $\mathbb{U}$ и $\mathbb{V}$.

Напомним, что вакуумные уравнения Эйнштейна (27.3) допускают решение в виде пространства-времени постоянной кривизны, которое называется пространством-временем (анти-) де Ситтера. Это пространство-время имеет максимальное число – десять – векторов Киллинга. Хотя полная (четырехмерная) скалярная кривизна для решения в виде произведения двух поверхностей постоянной кривизны с метрикой (27.22), впрочем как и для всех других решений вакуумных уравнений Эйнштейна (27.3), постоянна, $\widehat{R} = 4\Lambda$, решения (27.23) при $\Lambda \neq 0$ не совпадают с решением (анти-) де Ситтера. Действительно, каждая из поверхностей $\mathbb{L}^2$, $\mathbb{H}^2$ и $\mathbb{S}^2$ имеет по три вектора Киллинга, и можно показать (см. например [**?**]), что четырехмерное пространство-время имеет всего шесть векторов Киллинга. Поэтому решения в виде произведения двух поверхностей не совпадают с решением (анти-) де Ситтера.

Все решения в виде произведения двух поверхностей постоянной кривизны относятся к типу D по классификации Петрова [**?**].

27.4 Пространственно симметричные решения

Во втором случае B (27.17) дилатонное поле k постоянно. Не ограничивая общности, положим $k = 1$. Тогда вся система уравнений Эйнштейна (27.12)–(27.16) сводится к следующей системе:

$$\nabla_\alpha \nabla_\beta m - \frac{\nabla_\alpha m \nabla_\beta m}{2m} - \frac{1}{2} g_{\alpha\beta}\left[\nabla^2 m - \frac{(\nabla m)^2}{2m}\right] = 0, \tag{27.24}$$

$$R^h + \nabla^2 m - 2m\Lambda = 0, \tag{27.25}$$

$$R^g + \frac{\nabla^2 m}{m} - \frac{(\nabla m)^2}{2m^2} - 2\Lambda = 0. \tag{27.26}$$

Уравнение (27.25) представляет собой сумму двух слагаемых, зависящих от координат на разных поверхностях, $x \in \mathbb{U}$ и $y \in \mathbb{V}$, которая должна быть равна нулю. Это значит, что каждое слагаемое равно некоторой постоянной. Зафиксируем эту постоянную следующим уравнением $R^h = -2K = \mathsf{const}$. Таким образом, в случае B поверхность $\mathbb{V}$ – это поверхность постоянной кривизны. При этом возможны три случая, когда гауссова кривизна римановой поверхности $\mathbb{V}$ положительна, $\mathbb{V} = \mathbb{S}^2$, равна нулю, $\mathbb{V} = \mathbb{R}^2$, или отрицательна, $\mathbb{V} = \mathbb{H}^2$. Тогда соответствующие решения вакуумных уравнений Эйнштейна инвариантны относительно групп преобразований $\mathbb{O}(3)$, $\mathbb{IO}(2)$ или $\mathbb{O}(1,2)$, которые являются группами изометрий поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ и $\mathbb{H}^2$. Соответствующее четырехмерное пространство-время представляет собой сплетенное произведение поверхности $\mathbb{U}$ с одной из поверхностей $\mathbb{S}^2$, $\mathbb{R}^2$ или $\mathbb{H}^2$, где $\mathbb{U}$ представляется диаграммой Картера–Пенроуза. В частном случае при $K = 1$, возникают сферически симметричные решения. Таким образом, в рассматриваемом случае группа симметрии пространства-времени возникает как следствие уравнений движения.

При $K = \mathsf{const}$ уравнение (27.25) принимает вид

$$\nabla^2 m - 2(m\Lambda + K) = 0. \tag{27.27}$$

Исключая случай A, рассмотренный в предыдущем разделе, двинемся дальше, считая, что $\nabla_\alpha m \neq 0$.

Предложение 27.4.1. *Уравнение (27.27) является первым интегралом уравнений (27.24) и (27.26).*

Доказательство. Продифференцируем уравнение (27.27), используем тождество

$$[\nabla_\alpha, \nabla_\beta]\, A_\gamma = -R^g_{\alpha\beta\gamma}{}^\delta A_\delta\,,$$

где A_α – компоненты произвольного ковекторного поля, для изменения порядка ковариантных производных и используем уравнение (27.24) три раза для исключения вторых производных от m. После небольших алгебраических выкладок мы получим уравнение (27.26). □

Из доказательства предложения следует, что достаточно решить только уравнения (27.24) и (27.27), при этом уравнение (27.26) будет удовлетворено автоматически.

Замечание. Исходное действие Гильберта–Эйнштейна инвариантно относительно общих преобразований координат, и, согласно второй теореме Нетер, между уравнениями движения существует линейная зависимость. Поэтому зависимость уравнений (27.24)–(27.26) не является чем то удивительным и связана с инвариантностью исходного действия. □

Для явного решения уравнений движения (27.24) и (27.27) зафиксируем конформную калибровку для метрики $g_{\alpha\beta}$ на лоренцевой поверхности $\mathbb{U}$:

$$g_{\alpha\beta}dx^\alpha dx^\beta = \Phi d\xi d\eta, \tag{27.28}$$

где $\Phi(\xi, \eta)$ – конформный множитель, который зависит от координат светового конуса ξ, η на $\mathbb{U}$. Соответствующая четырехмерная метрика примет вид

$$ds^2 = \Phi d\xi d\eta + m d\Omega, \tag{27.29}$$

где $d\Omega$ – метрика на римановой поверхности постоянной кривизны $\mathbb{V} = \mathbb{S}^2$, $\mathbb{R}^2$ или $\mathbb{H}^2$. Знак конформного множителя Φ пока не фиксируем.

При $\Phi > 0$ и $m < 0$ сигнатура метрики (27.29) равна $(+ - --)$, если пространственная и временна́я координаты определены так же, как и раньше (26.23). Если изменить знак $m > 0$, то сигнатура метрики станет $(+ - ++)$. Такое же преобразование сигнатуры можно получить, изменив общий знак метрики, $\hat{g}_{ij} \mapsto -\hat{g}_{ij}$, и переставив первые две координаты. Вакуумные уравнения Эйнштейна (27.3) инвариантны относительно одновременного изменения знаков метрики и космологической постоянной. Поскольку в дальнейшем мы построим глобальные решения для всех возможных значений космологической постоянной, $\Lambda \in \mathbb{R}$, то, не ограничивая общности, достаточно рассмотреть случай $m < 0$. При отрицательных m удобно ввести параметризацию

$$m = -q^2, \qquad q(\xi, \eta) > 0. \tag{27.30}$$

Символы Кристоффеля для метрики (27.28) в конформной калибровке имеют только две ненулевые компоненты:

$$\Gamma_{\xi\xi}{}^{\xi} = \frac{\partial_\xi \Phi}{\Phi}, \qquad \Gamma_{\eta\eta}{}^{\eta} = \frac{\partial_\eta \Phi}{\Phi}, \tag{27.31}$$

и уравнения (27.24), (27.27) принимают простой вид

$$-\partial^2_{\xi\xi} q + \frac{\partial_\xi \Phi \partial_\xi q}{\Phi} = 0, \tag{27.32}$$

$$-\partial^2_{\eta\eta} q + \frac{\partial_\eta \Phi \partial_\eta q}{\Phi} = 0, \tag{27.33}$$

$$-2\frac{\partial^2_{\xi\eta} q^2}{\Phi} - (K - \Lambda q^2) = 0. \tag{27.34}$$

Таким образом, полная система уравнений (27.24)–(27.26) в конформной калибровке (27.28) сводится к трем уравнениям на две неизвестные функции q и Φ. Первые два уравнения являются обыкновенными дифференциальными уравнениями, и определяют функции q и Φ с точностью до умножения на произвольную постоянную. Система уравнений (27.32)–(27.34) переопределена и может быть проинтегрирована явно.

Предложение 27.4.2. *Условия $\partial_\xi q = 0$ и $\partial_\eta q = 0$ эквивалентны.*

Доказательство. Если $\partial_\xi q = 0$, то из уравнения (27.34) следует $q^2 = K/\Lambda = \mathsf{const}$ и, следовательно, $\partial_\eta q = 0$. Обратное утверждение верно по той же причине. □

Поскольку $q = \mathsf{const}$ соответствует уже рассмотренному случаю A, то предположим, что $\partial_\xi q \neq 0$ и $\partial_\eta q \neq 0$. Тогда, разделив уравнения (27.32) и (27.33) соответственно на $\partial_\xi q$ и $\partial_\eta q$, они легко интегрируются:

$$-\ln|\partial_\xi q| + \ln|\Phi| = \tilde{G}(\eta), \tag{27.35}$$

$$-\ln|\partial_\eta q| + \ln|\Phi| = \tilde{F}(\xi). \tag{27.36}$$

При этом возникают две произвольные функции $\tilde{F}(\xi)$ и $\tilde{G}(\eta)$. Введем монотонные функции $F(\eta)$ и $G(\xi)$ при помощи дифференциальных уравнений

$$F' := \frac{dF}{d\xi} = C\,\mathsf{e}^{\tilde{F}} > 0, \qquad G' := \frac{dG}{d\eta} = C\,\mathsf{e}^{\tilde{G}} > 0,$$

где $C > 0$ – некоторая положительная постоянная, которую мы зафиксируем чуть позже. Тогда разность уравнений (27.35) и (27.36) примет вид

$$\frac{|\partial_\xi q|}{F'} = \frac{|\partial_\eta q|}{G'}. \tag{27.37}$$

Конформная калибровка для двумерной метрики (27.28) определена с точностью до конформных преобразований. Воспользуемся этой свободой и перейдем к новым координатам $\xi, \eta \mapsto F, G$. Это всегда можно сделать, т.к. якобиан преобразования координат отличен от нуля, $F'G' \neq 0$. При конформном преобразовании координат конформный множитель преобразуется по-правилу $\Phi \mapsto \Phi/(F'G')$, что следует из вида метрики в конформной калибровке (27.28).

Предложение 27.4.3. *Уравнения (27.32)–(27.34) ковариантны относительно конформных преобразований*

$$\xi, \eta \mapsto F, G, \qquad \Phi \mapsto \tilde{\Phi} = \frac{\Phi}{F'G'}. \tag{27.38}$$

Доказательство. Прямая проверка. □

Таким образом, произвольные функции $\tilde{F}(\xi)$ и $\tilde{G}(\eta)$, возникшие в первых интегралах (27.35), (27.36), соответствуют конформным преобразованиям.

Перейдем к новым координатам $\xi, \eta \mapsto F, G$.

Предложение 27.4.4. *Если $\partial_\xi q\, \partial_\eta q > 0$, то в новых координатах функция $q(\tau)$ зависит только от временно́й координаты $\tau := \frac{1}{2}(F + G)$. Если $\partial_\xi q\, \partial_\eta q < 0$, то функция $q(\sigma)$ зависит только от пространственной координаты $\sigma := \frac{1}{2}(F - G)$.*

Доказательство. Из-за знаков модулей в уравнении (27.37), возможны два случая.

Если $\partial_\xi q\, \partial_\eta q > 0$, то справедливо равенство

$$\frac{\partial q}{\partial(F-G)} = \partial_\xi q \frac{\partial \xi}{\partial(F-G)} + \partial_\eta q \frac{\partial \eta}{\partial(F-G)} = \frac{\partial_\xi q}{F'} - \frac{\partial_\eta q}{G'} = 0. \tag{27.39}$$

Последнее равенство вытекает из уравнения (27.37). Поэтому, переходя к координатам $F, G \mapsto \tau, \sigma$, получаем сделанное утверждение.

Аналогично, если $\partial_\xi q\, \partial_\eta q < 0$, то выполнено равенство

$$\frac{\partial q}{\partial(F+G)} = \partial_\xi q \frac{\partial \xi}{\partial(F+G)} + \partial_\eta q \frac{\partial \eta}{\partial(F+G)} = \frac{\partial_\xi q}{F'} + \frac{\partial_\eta q}{G'} = 0. \quad \square \tag{27.40}$$

Теперь из каждого из двух уравнений (27.35) или (27.36) следует одно и то же равенство

$$|\Phi| = \frac{1}{2C} F'G'|q'|,$$

где q' обозначает производную либо по $\tau := \frac{1}{2}(F+G)$, либо по $\sigma := \frac{1}{2}(F-G)$. Постоянная C соответствует растяжке новых координат F, G, и, для упрощения последующих формул, положим $C = 1/2$. При конформном преобразовании конформный множитель преобразуется по правилу (27.38). Поэтому после конформного преобразования (27.38) будет выполнено равенство

$$|\tilde{\Phi}| = |q'|. \tag{27.41}$$

В дальнейшем знак тильды у конформного множителя мы, для краткости, опустим.

Таким образом, координаты всегда можно выбрать таким образом, чтобы функции q и Φ зависели одновременно только от времениподобной или пространственноподобной координаты

$$\zeta = \frac{1}{2}(F \pm G) := \begin{cases} \tau, & \partial_\xi q\, \partial_\eta q > 0 \\ \sigma, & \partial_\xi q\, \partial_\eta q < 0. \end{cases} \tag{27.42}$$

Это значит, что двумерная метрика (27.28) обладает вектором Киллинга, ∂_σ или ∂_τ, как следствие уравнений (27.32) и (27.33). Назовем эти решения соответственно однородными и статическими, хотя это и относится только к определенной системе координат. Существование вектора Киллинга является обобщением *теоремы Биркгоффа* [**?**], утверждающей, что произвольное сферически симметричное решение вакуумных уравнений Эйнштейна должно быть статическим. (Это утверждение было опубликовано ранее в статье [**?**].) Обобщение заключается в том, что наличие вектора Киллинга доказано не только для сферически симметричных решений ($K = 1$), но и для решений, инвариантных относительно групп преобразований $\mathbb{IO}(2)$ ($K = 0$) и $\mathbb{SO}(1,2)$ ($K = -1$).

Окончательно, решением уравнений (27.32), (27.33) в фиксированной системе координат является равенство (27.41) и утверждение о том, что функции q и Φ зависят только от одной переменной ζ (27.42). Осталось решить только одно уравнение (27.34).

В статическом, $q = q(\sigma)$, и однородном, $q = q(\tau)$, случаях уравнение (27.34) принимает вид

$$(q^2)'' = \quad 2(K - \Lambda q^2)\Phi, \qquad q = q(\sigma), \tag{27.43}$$

$$(q^2)'' = -2(K - \Lambda q^2)\Phi, \qquad q = q(\tau). \tag{27.44}$$

Чтобы проинтегрировать полученные уравнения, необходимо выразить Φ через q с помощью уравнения (27.41), а для этого необходимо раскрыть знаки модулей.

Рассмотрим подробно статический случай $q = q(\sigma)$, $\Phi > 0$ и $q' > 0$. Тогда уравнение (27.43) с учетом (27.41) примет вид

$$(q^2)'' = 2(K - \Lambda q^2)q'.$$

Его легко проинтегрировать

$$(q^2)' = 2\left(Kq - \frac{\Lambda q^3}{3} - 2M\right),$$

где $M = \mathsf{const}$ – постоянная интегрирования. В дальнейшем мы увидим, что она совпадает с массой в решении Шварцшильда. Выполнив дифференцирование в левой части и поделив на $2q > 0$, получим уравнение

$$q' = K - \frac{2M}{q} - \frac{\Lambda q^2}{3}.$$

Поскольку в рассматриваемом случае $q' = \Phi$, то отсюда следует выражение для конформного множителя через переменную q:

$$\Phi(q) = K - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \tag{27.45}$$

Если $q = q(\sigma)$, $\Phi > 0$ и $q' < 0$, то аналогичное интегрирование приводит к уравнению

$$q' = -\Phi(q),$$

где в правой части стоит тот же самый конформный множитель (27.45). Этот случай можно объединить с предыдущим, записав уравнение для q в виде

$$|q'| = \Phi(q), \qquad q = q(\sigma), \quad \Phi > 0. \tag{27.46}$$

Аналогично интегрируются статический случай при $\Phi < 0$:

$$|q'| = -\Phi(q), \qquad q = q(\sigma), \quad \Phi < 0. \tag{27.47}$$

Если решение однородно, $q = q(\tau)$ и $\Phi > 0$, $q' > 0$, то интегрирование уравнения (27.44) приводит к равенству

$$q' = -\left(K - \frac{2M}{q} - \frac{\Lambda q^2}{3}\right).$$

То есть в этом случае конформный множитель надо отождествить с правой частью

$$\hat{\Phi} = -\left(K - \frac{2M}{q} - \frac{\Lambda q^2}{3}\right). \tag{27.48}$$

Поскольку выражение конформного множителя в однородном случае через q отличается знаком, то мы пометили его шляпкой. Таким образом, однородные решения уравнений Эйнштейна можно записать в виде

$$|q'| = \hat{\Phi}(q), \qquad q = q(\tau), \quad \hat{\Phi} > 0. \tag{27.49}$$

$$|q'| = -\hat{\Phi}(q), \qquad q = q(\tau), \quad \hat{\Phi} < 0. \tag{27.50}$$

Если конформный множитель отрицателен, то сигнатура метрики равна $(-+--)$. В этом случае, сделав замену $\tau \leftrightarrow \sigma$, мы вернемся к прежней сигнатуре метрики $(+---)$. Это преобразование позволяет объединить стационарные и однородные решения, написав знак модуля у конформного множителя в выражении для метрики (27.29). Тогда общее решение вакуумных уравнений Эйнштейна (27.3) в соответствующей системе координат примет вид

$$ds^2 = |\Phi|(d\tau^2 - d\sigma^2) - q^2 d\Omega, \tag{27.51}$$

где конформный множитель Φ имеет вид (27.45). При этом переменная q зависит либо от σ (статическое локальное решение), либо от τ (однородное локальное решение) через дифференциальное уравнение

$$\left|\frac{dq}{d\zeta}\right| = \pm\Phi(q), \tag{27.52}$$

где выполнено правило знаков:

$$\begin{array}{lll} \Phi > 0: & \zeta = \sigma, & \text{знак } + \text{ (статическое локальное решение)}, \\ \Phi < 0: & \zeta = \tau, & \text{знак } - \text{ (однородное локальное решение)}. \end{array} \tag{27.53}$$

Таким образом, из четырехмерных уравнений Эйнштейна вытекает, что на поверхности $\mathbb{U}$ возникает метрика с одним вектором Киллинга, которая была подробно

рассмотрена в главе 26. Теперь с помощью метода конформных блоков можно построить глобальные (максимально продолженные вдоль экстремалей) решения вакуумных уравнений Эйнштейна. Число особенностей и нулей конформного множителя (27.45) зависит от соотношения между постоянными K, M и Λ. Поэтому возможно существование многих существенно различных глобальных решений, которые мы рассмотрим в следующих разделах.

Конформный множитель (27.45) имеет одну сингулярность: простой полюс при $q = 0$. Поэтому, согласно правилам построения глобальных решений из раздела 26.4 каждое глобальное решение соответствует одному из интервалов $(-\infty, 0)$ или $(0, \infty)$. Из вида конформного множителя (27.58) следует, что эти глобальные решения получаются друг из друга преобразованием $M \mapsto -M$. Поэтому, не ограничивая общности, мы опишем только глобальные решения, соответствующие положительному интервалу $(q_-, q_+) = (0, \infty)$. Впрочем, это предположение уже было сделано при параметризации (27.30).

Поскольку конформный множитель $\Phi(q)$ является гладкой функцией при $q > 0$, то все возникающие лоренцевы поверхности $\mathbb{U}$ и метрика на них, являются гладкими.

Используя уравнения (27.26), (9.2) и (**??**), нетрудно вычислить скалярную кривизну поверхности $\mathbb{U}$

$$R^g = \frac{2}{3}\Lambda + \frac{4M}{q^3}. \tag{27.54}$$

Она не зависит от гауссовой кривизны K римановой поверхности $\mathbb{V}$ и сингулярна при $q = 0$, если $M \neq 0$. Отметим, что сингулярная часть двумерной скалярной кривизны (27.54) пропорциональна собственному значению четырехмерного тензора Вейля (см., например, [**?**]):

$$\frac{1}{48}\widehat{C}_{ijkl}\widehat{C}^{ijkl} = \left(-\frac{M}{q^3}\right)^2. \tag{27.55}$$

Теперь перейдем к описанию всех пространственно симметричных глобальных решений вакуумных уравнений Эйнштейна.

27.4.1 Сферически симметричные решения $K = 1$

При $K = 1$ риманова поверхность $\mathbb{V}$ представляет собой сферу $\mathbb{S}^2$, и все решения сферически симметричны. Для сферы единичного радиуса метрику (27.21) можно записать в сферических координатах

$$d\Omega = d\theta^2 + \sin^2\theta d\varphi^2. \tag{27.56}$$

Сферически симметричную метрику пространства-времени, которая удовлетворяет уравнениям Эйнштейна, можно записать в виде

$$ds^2 = |\Phi(q)|(d\tau^2 - d\sigma^2) - q^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{27.57}$$

где

$$\Phi(q) = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \tag{27.58}$$

Переменная q связана с σ или τ дифференциальным уравнением (27.52), где выполнено правило знаков (27.53).

Определение. Координаты τ, σ, в которых записана сферически симметричная метрика (27.57), называются *черепашьими*. Это название, по-видимому, произошло потому что диаграммы Картера–Пенроуза чем то напоминают рисунок панциря черепахи. □

Обобщение решения Шварцшильда на случай ненулевой космологической постоянной (27.57) было получено Коттлером [**?**].

В рассматриваемом случае все решения параметризуются двумя постоянными: космологической постоянной Λ и массой M. Вторую постоянную мы будем называть массой, хотя она и не имеет физического смысла массы для большинства решений, отличных от решения Шварцшильда.

Как было отмечено в предыдущем разделе, мы рассматриваем глобальные решения вакуумных уравнений Эйнштейна, соответствующих положительному интервалу $(q_-, q_+) = (0, \infty)$. Структура глобальных решений определяется количеством и типом положительных корней уравнения $\Phi(q) = 0$ или эквивалентного кубического уравнения

$$\frac{\Lambda}{3} q^3 - q + 2M = 0. \tag{27.59}$$

При ненулевой космологической постоянной это уравнение может иметь до трех нулей. Элементарный анализ показывает, что по крайней мере один из корней отрицателен. Это значит, что при положительных q возможно существование не более двух горизонтов.

При положительной космологической постоянной $\Lambda > 0$ существуют следующие возможности в зависимости от величины массы M (см. рис.27.1 слева). Если $M > \frac{1}{3\sqrt{\Lambda}}$, то уравнение (27.59) не имеет положительных корней. При $M = \frac{1}{3\sqrt{\Lambda}}$ возникает один положительный корень второго порядка. В интервале значений $0 < M < \frac{1}{3\sqrt{\Lambda}}$ существуют два положительных корня. При неположительных значениях $M \leq 0$ имеется один простой положительный нуль.

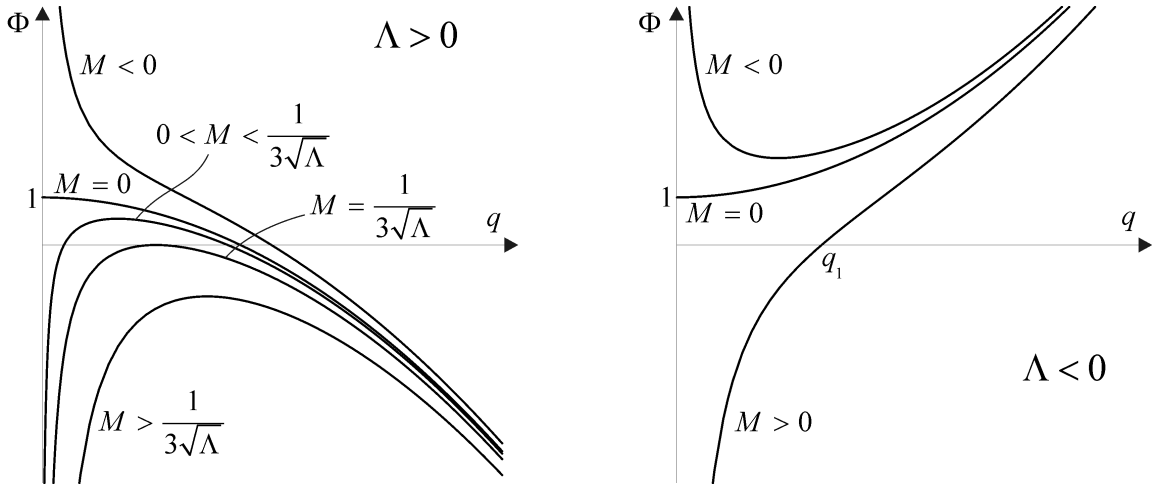


Рис. 27.1: Поведение конформного множителя $\Phi = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}$ в зависимости от величины массы M при положительной (слева) и отрицательной (справа) космологической постоянной Λ.

При отрицательных значениях космологической постоянной $\Lambda < 0$ мы имеем один простой положительный нуль при $M > 0$ и ни одного нуля при $M \leq 0$ (см. рис.27.1 справа).

Перейдем к классификации решений.

Пространство-время Минковского $\Lambda = 0$, $M = 0$

Наиболее простое сферически симметричное решение получается при $\Lambda = 0$ и $M = 0$ (см. рис.27.2). В этом случае $N = 1$, и метрика принимает вид

$$ds^2 = d\tau^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \qquad r \in (0, \infty), \tag{27.60}$$

где мы переобозначили $\sigma \mapsto r$. Точка $r = 0$ является координатной сингулярностью. Переходя к декартовым координатам в четырехмерном пространстве-времени и добавляя мировую линию начала сферической системы координат $r = 0$, мы получим пространство-время Минковского $\mathbb{R}^{1,3}$. При этом пространственная координата r естественным образом отождествляется с радиусом сферической системы координат. В этом случае пространство-время нельзя представить в виде топологического произведения $\mathbb{U} \times \mathbb{S}^2$, и поэтому диаграмма Картера–Пенроуза на рис.27.2 не продолжается.

Черная дыра Шварцшильда $\Lambda = 0$, $M > 0$

Решение Шварцшильда соответствует нулевой космологической постоянной $\Lambda = 0$ и положительной массе $M > 0$. Оно уже обсуждалось в разделе 26.5.1. В этом случае конформный множитель (27.58) имеет один простой нуль в точке $q_1 = 2M$ и, следовательно, один горизонт. Соответствующая диаграмма Картера–Пенроуза также изображена на рис.27.2.

Голая сингулярность $\Lambda = 0$, $M < 0$

При нулевой космологической постоянной $\Lambda = 0$ и отрицательной массе $M < 0$ уравнение (27.59) не имеет положительных корней. Этот случай также был рассмотрен в разделе 26.5.1. При указанных значениях постоянных существуют две диаграммы Картера–Пенроуза, каждая из которых состоит из одного треугольного конформного блока, изображенного на рис.27.2, и его пространственного отражения. Каждый из конформных блоков представляет собой максимально продолженную поверхность $\mathbb{U}$. Сингулярная граница времениподобна и называется *голой сингулярностью*, т.к. не окружена горизонтом.

Решение де Ситтера $\Lambda > 0$, $M = 0$

Решение де Ситтера соответствует положительной космологической постоянной и нулевой массе. При этом пространство-время представляет собой многообразие постоянной (в наших обозначениях положительной) скалярной кривизны и может быть представлено, как четырехмерный гиперболоид, вложенный в пятимерное пространство Минковского $\mathbb{R}^{1,4}$, с индуцированной метрикой. Его группой симметрии является группа Лоренца $\mathbb{O}(1,4)$, а метрика имеет максимальное число – десять – векторов Киллинга, что совпадает с размерностью группы симметрии. Конформный множитель (27.58) имеет один простой положительный нуль, соответствующий горизонту.

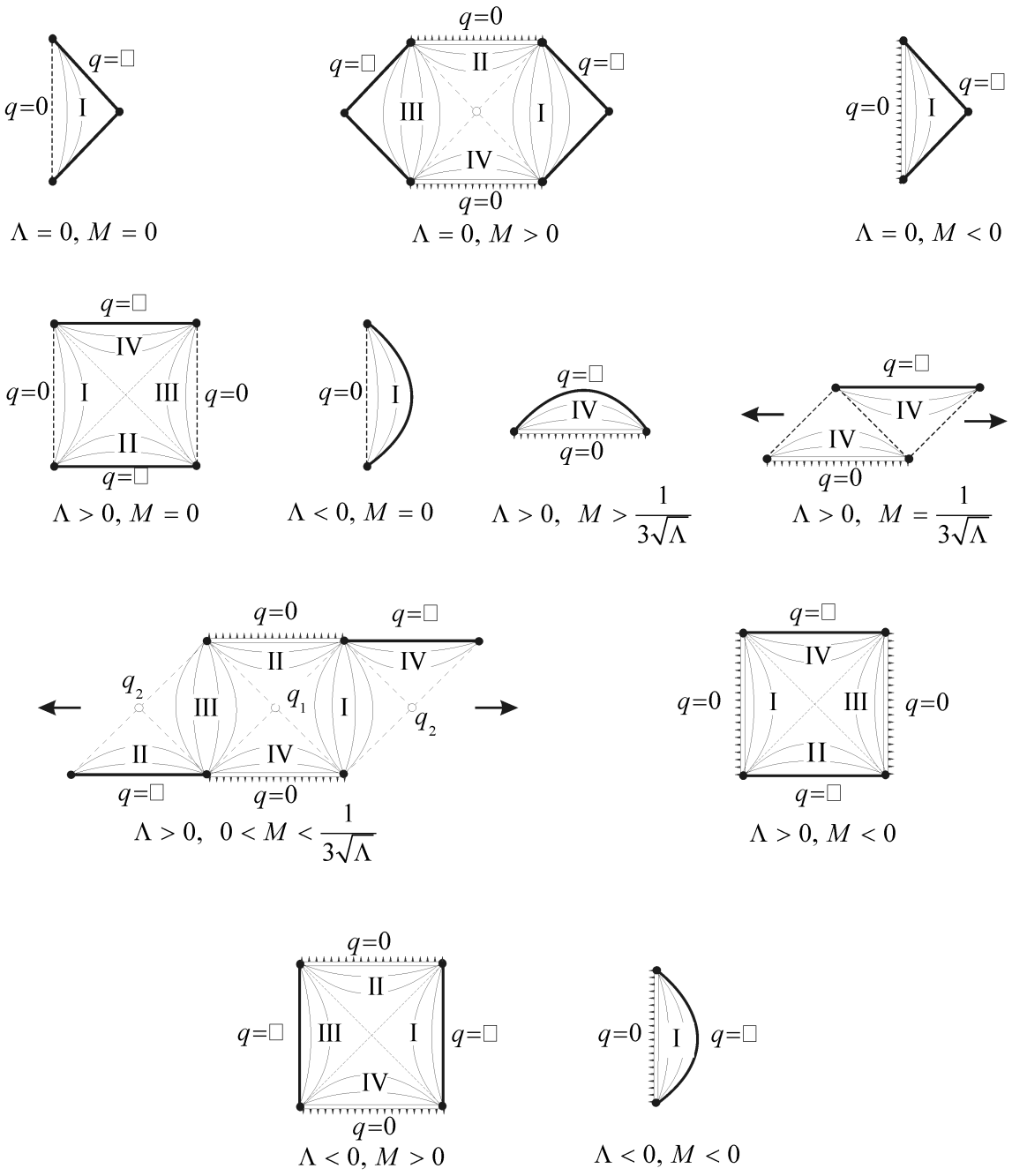


Рис. 27.2: Сферически симметричные решения для конформного множителя $\Phi = 1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}$. Условные обозначения те же, что и на рис.26.3. Стрелки показывают возможное периодическое продолжение решений.

Статические и однородные решения в координатах Шварцшильда имеют вид

$$ds^2 = \left(1 - \frac{\Lambda}{3} r^2\right) d\tau^2 - \frac{dr^2}{1 - \frac{\Lambda}{3} r^2} - r^2 d\Omega^2, \tag{27.61}$$

$$0 < r < \sqrt{\tfrac{3}{\Lambda}}\,, \quad -\infty < \tau < \infty\,,$$

$$ds^2 = -\frac{dt^2}{1 - \frac{\Lambda}{3} t^2} + \left(1 - \frac{\Lambda}{3} t^2\right) d\sigma^2 - t^2 d\Omega^2, \tag{27.62}$$

$$\sqrt{\tfrac{3}{\Lambda}} < t < \infty\,, \quad -\infty < \sigma < \infty\,.$$

Поскольку уравнение (27.52) в этом случае интегрируется явно, то метрику можно записать также в конформно плоском виде

$$ds^2 = \frac{1}{\mathsf{ch}^{\,2}\left(\sqrt{\frac{\Lambda}{3}}\sigma\right)} (d\tau^2 - d\sigma^2) - \frac{3}{\Lambda}\, \mathsf{th}^{\,2}\left(\sqrt{\tfrac{\Lambda}{3}}\sigma\right) d\Omega^2, \tag{27.63}$$

$$ds^2 = \frac{1}{\mathsf{sh}^{\,2}\left(\sqrt{\frac{\Lambda}{3}}\tau\right)} (d\tau^2 - d\sigma^2) - \frac{3}{\Lambda}\, \mathsf{cth}^{\,2}\left(\sqrt{\tfrac{\Lambda}{3}}\tau\right) d\Omega^2, \tag{27.64}$$

соответственно для статического и однородного случая. Область определения $r \in \left(0, \sqrt{3/\Lambda}\right)$ переходит в $\sigma \in (0, \infty)$, а $t \in \left(\sqrt{3/\Lambda}, \infty\right)$ – в $\tau \in (0, \infty)$.

Поведение конформного множителя показано на рис.27.1 слева. Он имеет один простой нуль в некоторой точке, которую мы обозначим q_1. В интервале $(0, q_1)$ конформный множитель положителен и ему ставится в соответствие два треугольных статических конформных блока I и III с метрикой (27.63). В интервале (q_1, ∞) конформный множитель отрицателен, и ему соответствуют два треугольных однородных конформных блока II и IV с метрикой (27.64). Эти четыре конформных блока склеиваются единственным образом по правилам, сформулированным в разделе 26.4. Соответствующая диаграмма Картера–Пенроуза показана на рис.27.2. Решение де Ситтера получится после добавления времениподобной границы $r = 0$ к сплетенному произведению $\mathbb{U} \times \mathbb{S}^2$.

Заметим, что конформный множитель для решения де Ситтера является квадратичным полиномом, $\Phi = 1 - \frac{\Lambda q^2}{3}$. Как показано в разделе 26.5.5, при максимальном продолжении поверхности с такой метрикой возникает однополостный гиперболоид. В рассматриваемом случае мы этого не делаем не потому что возникают какие то проблемы с поверхностью $\mathbb{U}$, а потому что при $q = 0$ коэффициент при угловой части метрики (27.57) становится равным нулю и метрика четырехмерного пространства-времени вырождается. Здесь возникает такая же ситуация, как в пространстве Минковского $\mathbb{R}^{1,3}$ в сферической системе координат.

Решение анти-де Ситтера $\Lambda < 0$, $M = 0$

Изменение знака космологической постоянной в уравнениях Эйнштейна приводит к качественному изменению решений. При нулевой массе мы получаем решение анти-де Ситтера. Соответствующее пространство-время представляет собой многообразие постоянной отрицательной кривизны, которое можно представить в виде четырехмерного гиперболоида, вложенного в плоское пространство $\mathbb{R}^{2,3}$ с метрикой $\eta_{\alpha\beta} = \mathsf{diag}\,(1, 1, -1, -1, -1)$. Метрика анти де Ситтера симметрична относительно действия группы вращений $\mathbb{O}(2, 3)$ и имеет также максимальное число – десять –

векторов Киллинга. Конформный множитель (27.58) не имеет нулей и всегда положителен. Поэтому решение статично и не имеет горизонтов. В координатах Шварцшильда метрика имеет тот же вид (27.61), что и для решения де Ситтера, однако из-за отрицательного знака Λ область изменения r совпадает со всем положительным интервалом $(0,\infty)$. В конформно плоском виде метрика принимает вид

$$ds^2 = \frac{1}{\cos^2\left(\sqrt{\frac{|\Lambda|}{3}}\sigma\right)}(d\tau^2 - d\sigma^2) - \frac{3}{|\Lambda|}\,\mathrm{th}^2\left(\sqrt{\tfrac{|\Lambda|}{3}}\sigma\right)d\Omega. \tag{27.65}$$

При этом координата σ меняется в конечном интервале $\sigma \in \left(0, \frac{\pi}{2}\sqrt{\frac{3}{|\Lambda|}}\right)$. Соответствующая диаграмма Картера–Пенроуза имеет вид линзы и изображена на рис.27.2. Это решение неполно на границе $r = 0$ и не может быть продолжено, что является следствием выбора сферической системы координат. Для получения полного решения анти-де Ситтера мировую линию $r = 0$ необходимо добавить к многообразию, что можно сделать путем перехода к другой системе координат.

Однородная пространственная сингулярность $\Lambda > 0$, $M > \frac{1}{3\sqrt{\Lambda}}$

При $\Lambda > 0$, $M > \frac{1}{3\sqrt{\Lambda}}$ конформный множитель Φ для положительных $q > 0$ не имеет нулей и всегда отрицателен. Решение однородно и не имеет горизонтов. Поверхность $\mathbb{U}$ в этом случае представляется диаграммой Картера–Пенроуза в виде линзы, изображенной на рис.27.2, и ее отражения во времени. При конечном значении времени в прошлом имеется истинная сингулярность и у четырехмерной кривизны, и у тензора кривизны поверхности $\mathbb{U}$. Многообразие не может быть продолжено через эту сингулярность. Космологическая интерпретация этого решения простая. Вселенная рождается в конечном прошлом из сингулярности и развивается вечно. В пределе $t \to \infty$ метрика на поверхности $\mathbb{U}$ стремится к метрике де Ситтера. Это же верно для трех следующих глобальных решений.

Двойной горизонт $\Lambda > 0$, $M = \frac{1}{3\sqrt{\Lambda}}$

Если $\Lambda > 0$, $M = \frac{1}{3\sqrt{\Lambda}}$ то конформный множитель Φ имеет один нуль второго порядка, соответствующий горизонту $q_1 = 1/\sqrt{\Lambda}$. При $q > 0$ функция Φ неотрицательна и, следовательно, все конформные блоки однородны. Два однородных конформных блока соответствуют интервалу $(0, q_1)$ и два – интервалу (q_1, ∞). Они склеиваются в фундаментальную область, показанную на рис.27.2. Эту фундаментальную область можно либо продолжить влево и вправо, что приводит к универсальному накрывающему пространству, либо отождествить противоположные неполные края. В последнем случае возникающая поверхность диффеоморфна цилиндру. В случае одного двойного горизонта имеются также поверхности, возникающие из построенных выше отражением во времени.

Два горизонта $\Lambda > 0$, $0 < M < \frac{1}{3\sqrt{\Lambda}}$

При $\Lambda > 0$, $0 < M < \frac{1}{3\sqrt{\Lambda}}$ конформный множитель Φ имеет два нуля и, значит, два горизонта в точках q_1 и q_2. Двум интервалам $(0, q_1)$ и $(q_2, 0)$, в которых конформный множитель отрицателен, ставится в соответствие по два однородных треугольных конформных блока. В интервале (q_1, q_2) конформный множитель положителен, и

ему ставится в соответствие два статических квадратных конформных блока. Фундаментальная область возникает после склейки всех шести конформных блоков по правилам, описанным в разделе 26.4. Как и в предыдущем примере, фундаментальную область можно либо продолжить влево и вправо, и получить универсальное накрывающее пространство, как показано стрелками на рис.27.2, либо отождествить неполные края.

Две статические сингулярности $\Lambda > 0$, $M < 0$

При положительной космологической постоянной $\Lambda > 0$ и отрицательной массе $M < 0$ конформный множитель Φ имеет один простой нуль в точке q_1. Интервалу $(0, q_1)$, в котором конформный множитель положителен, ставится в соответствие два треугольных статических конформных блока. Интервалу (q_1, ∞), где конформный множитель отрицателен, ставится в соответствие два треугольных однородных конформных блока. После склейки этих блоков возникает диаграмма Картера–Пенроуза, показанная на рис.27.2. Мы имеем две сингулярности в статических областях I, III, разделенных горизонтами. Из области I вдоль времениподобной кривой можно достичь только правую сингулярность. Это же верно для области III и левой сингулярности. Из области II имеется возможность попасть в обе сингулярности. Из области IV никакой сингулярности достичь невозможно, и жизнь там продолжается вечно, если вы родились в этой области.

Черная дыра анти-де Ситтера $\Lambda < 0$, $M > 0$

Если $\Lambda < 0$ и $M > 0$, то конформный множитель Φ имеет один нуль и, соответственно, один горизонт в точке q_1, как показано на рис.27.1 справа. Интервалу $(0, q_1)$, где конформный множитель отрицателен, ставится в соответствие два треугольных однородных конформных блока. В интервале (q_1, ∞) конформный множитель положителен, и ему соответствует два треугольных статических конформных блока. Эти четыре конформных блока склеиваются в диаграмму Картера–Пенроуза, показанную на рис.27.2. Она аналогична диаграмме для шварцшильдовской черной дыры. Отличие заключается в том, что полные левая и правая границы в рассматриваемом случае времениподобны. Сингулярности кривизны расположены в конечном прошлом (белая дыра) и в конечном будущем (черная дыра). Они имеют тот же характер, что и в решении Шварцшильда. Метрика пространства-времени не является асимптотически плоской и стремится к метрике анти-де Ситтера при $q = r \to \infty$.

Голая сингулярность $\Lambda < 0$, $M < 0$

При $\Lambda < 0$ и $M < 0$ имеется голая сингулярность без горизонтов. Соответствующая диаграмма Картера–Пенроуза совпадает с диаграммой Картера–Пенроуза для $\Lambda > 0$, $M > \frac{1}{3\sqrt{\Lambda}}$ изображенной на рис.27.2, но повернутой на угол $\pi/2$, т.к. конформный множитель положителен $\Phi > 0$, и, следовательно, метрика статична. Метрика при $q = r \to \infty$ также стремится к метрике анти-де Ситтера.

27.4.2 Планарные решения $K = 0$

В случае $K = 0$ метрика на римановой поверхности $\mathbb{V}$ (27.21) становится евклидовой

$$d\Omega_{\text{P}} := dy^2 + dz^2. \tag{27.66}$$

Это значит, что соответствующая максимально продолженная поверхность $\mathbb{V}$ является либо евклидовой плоскостью $\mathbb{R}^2$ с группой симметрии Пуанкаре $\mathbb{IO}(2)$, либо ее компактификацией (цилиндр, тор). Будем называть соответствующие четырехмерные глобальные решения вакуумных уравнений Эйнштейна *планарными*. Для решений этого типа метрика в координатах Шварцшильда имеет вид

$$ds^2 = \Phi(q)d\zeta^2 - \frac{dq^2}{\Phi(q)} - q^2\Omega_{\text{P}}, \tag{27.67}$$

где

$$\Phi(q) = -\frac{2M}{q} - \frac{\Lambda q^2}{3}. \tag{27.68}$$

Координаты q и ζ определены в (27.52). Множество планарных решений так же как и в сферически симметричном случае параметризуется двумя постоянными: космологической постоянной Λ и массой M.

Замечание. Для планарных решений координата q не может быть интерпретирована, как радиус пространства. Поэтому мы с самого начала отказались от ее обозначения через r, т.к. решение Шварцшильда является лишь частным случаем описываемого общего подхода. □

Физическая интерпретация планарных решения весьма интересна. Например, при $\mathbb{V} = \mathbb{T}^2$ трехмерное пространство представляет собой прямое произведение тора $\mathbb{T}^2$ и прямой $\mathbb{R}$. Это пространство содержит нестягиваемые замкнутые пространственноподобные кривые, т.е. имеет нетривиальную фундаментальную группу. С физической точки зрения такие пространства описывают кротовые норы. В этом случае все горизонты также представляют собой торы.

Для построения глобальных планарных решений вакуумных уравнений Эйнштейна необходимо знать поведение конформного множителя (27.68) при различных значениях Λ и M. При $\Lambda \neq 0$ оно ясно из рис.27.1 и получается сдвигом оси абсцисс на единицу вверх, как показано на рис.27.3.

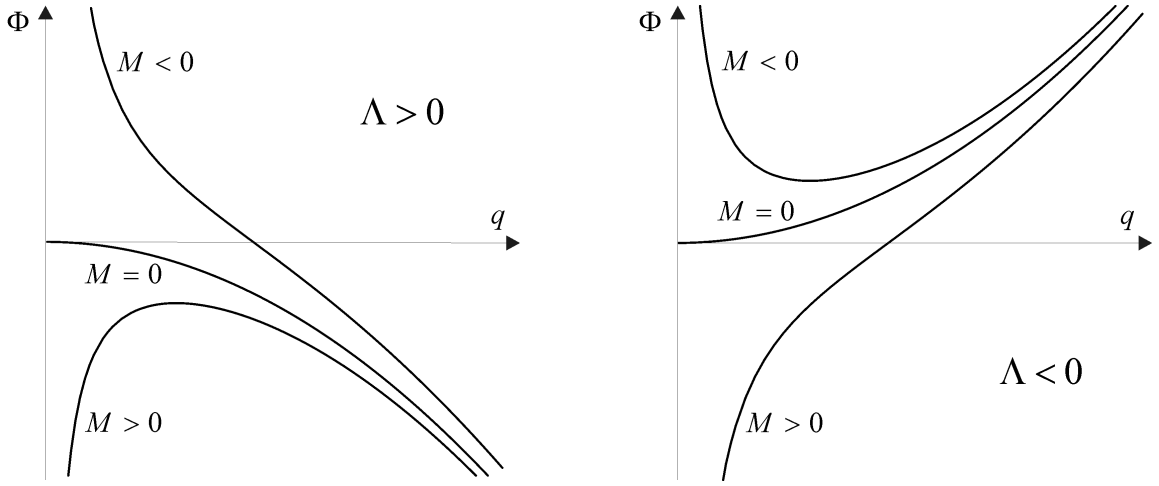


Рис. 27.3: Поведение конформного множителя $\Phi = -\frac{2M}{q} - \frac{\Lambda q^2}{3}$ в зависимости от величины массы M при положительной (слева) и отрицательной (справа) космологической постоянной Λ.

При $M = 0$ конформный множитель является квадратичным полиномом. Этот случай был рассмотрен в разделе 26.5.5 и соответствует поверхностям $\mathbb{U}$ постоянной кривизны.

Если космологическая постоянная положительна, $\Lambda > 0$, то конформный множитель при $M > 0$ не имеет нулей. При $M < 0$ существует один положительный нуль.

Для отрицательной космологической постоянной, $\Lambda < 0$, ситуация противоположна. При $M > 0$ есть один простой нуль, а при $M < 0$ нули отсутствуют.

Перечислим все возможные максимально продолженные поверхности $\mathbb{U}$. Как и ранее, достаточно рассмотреть глобальные решения, соответствующие интервалу $q \in (0, \infty)$.

Однородные и голые сингулярности $\Lambda = 0$, $M \neq 0$

При нулевой космологической постоянной и положительной массе $M > 0$ конформный множитель при $q > 0$ отрицателен и не имеет нулей (горизонтов). Поэтому глобальное решение состоит из одного треугольного конформного блока типа IV, однородно и имеет пространственноподобную сингулярность при $q = 0$. В шварцшильдовских и черепашьих координатах метрика имеет вид

$$ds^2 = \frac{t}{2M}dt^2 - \frac{2M}{t}d\sigma^2 - t^2 d\Omega_{\rm P}, \qquad t \in (0, \infty), \tag{27.69}$$

$$ds^2 = \sqrt{\tfrac{M}{|\tau|}}(d\tau^2 - d\sigma^2) - 4M|\tau| d\Omega_{\rm P}, \qquad \tau < 0 \quad \text{или} \quad \tau > 0. \tag{27.70}$$

Соответствующая диаграмма Картера–Пенроуза изображена на рис.27.4.

Как и ранее, существует также глобальное решение типа II, полученное из показанного на рисунке обращением времени.

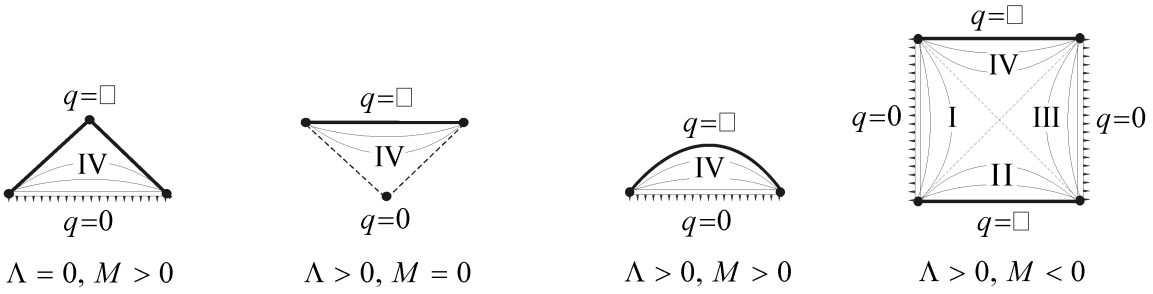


Рис. 27.4: Планарные решения $K = 0$. Условные обозначения те же, что и на рис.26.3.

Если масса отрицательна, $M < 0$, то оба глобальных решения имеют такой же вид как и при $M > 0$, но являются статическими. Поэтому соответствующие диаграммы Картера–Пенроуза необходимо просто повернуть на угол $\pi/2$. В таком случае возникают голые сингулярности.

Решение (анти-) де Ситтера $\Lambda \neq 0$, $M = 0$

При положительной космологической постоянной $\Lambda > 0$ и нулевой массе $M = 0$ конформный множитель

$$\Phi = -\frac{\Lambda t^2}{3}, \qquad q := t \in (0, \infty),$$

всюду отрицателен и не имеет нулей. Соответствующее глобальное решение однородно и не имеет горизонтов. Оно представляется треугольным конформным блоком

типа IV, изображенном на рис.27.4. Есть также глобальное решение типа II, полученное из решения типа IV обращением времени.

В шварцшильдовских и черепашьих координатах метрика имеет вид

$$ds^2 = \frac{3}{\Lambda t^2}dt^2 - \frac{\Lambda t^2}{3}d\sigma^2 - t^2 d\Omega_{\text{P}}, \qquad t \in (0,\infty), \tag{27.71}$$

$$ds^2 = \frac{3}{\Lambda \tau^2}(d\tau^2 - d\sigma^2) - \frac{9}{\Lambda^2\tau^2}d\Omega_{\text{P}}, \qquad \tau \in (0,\infty). \tag{27.72}$$

Вдоль границы $q = 0$ (т.е. $t = 0$) диаграммы Картера–Пенроуза 4-х мерная метрика вырождается и ее нельзя продолжать. Эти светоподобные края показаны на рис.27.4 пунктиром.

Можно проверить, что метрики (27.71) и (27.72) описывают пространства постоянной четырехмерной кривизны. Это означает, что эти метрики представляют собой решение де Ситтера.

В случае отрицательной космологической постоянной $\Lambda < 0$ зависимость компонент метрики от времени в выражениях (27.71) и (27.72) необходимо заменить на зависимость от пространственной координаты, и соответствующие диаграммы Картера–Пенроуза повернуть на угол $\pi/2$. В этом случае пространство-время снова является пространством постоянной кривизны, и возникает новое представление решения анти-де Ситтера.

Однородные и голые сингулярности $\Lambda M > 0$

При положительной космологической постоянной, $\Lambda > 0$, и положительной массе, $M > 0$, конформный множитель

$$\Phi = -\frac{\Lambda t^3 + 6M}{3t} < 0, \qquad q := t \in (0,\infty), \tag{27.73}$$

отрицателен и не имеет нулей (горизонтов). В шварцшильдовских и черепашьих координатах метрика имеет вид

$$ds^2 = \frac{3t}{\Lambda t^3 + 6M}dt^2 - \frac{\Lambda t^3 + 6M}{3t}d\sigma^2 - t^2 d\Omega_{\text{P}}, \qquad t \in (0,\infty), \tag{27.74}$$

$$ds^2 = |\Phi|(d\tau^2 - d\sigma^2) - t^2 d\Omega_{\text{P}}, \qquad -\frac{\pi}{12\sqrt{3}M} < \tau < \frac{\pi}{4\sqrt{3}M}, \tag{27.75}$$

где функция $t = t(\tau)$ определена уравнением (27.52), которое интегрируется и дает неявную связь $t = t(\tau)$ (при $q' > 0$):

$$\frac{\Lambda}{3}(\tau + \text{const}) = \frac{1}{6a}\ln\frac{a^2 - at + t^2}{(a+t)^2} + \frac{1}{a\sqrt{3}}\,\text{arcth}\left(\frac{2t-a}{a\sqrt{3}}\right), \qquad a := 6M/\Lambda. \tag{27.76}$$

Соответствующее максимально продолженное решение представляется диаграммой Картера–Пенроуза в виде линзы типа IV, которая изображена на рис.27.4, и ее отражением во времени (диаграмма типа II). Оно описывает однородную пространственную сингулярность при $t = 0$ и является асимптотически пространством-временем де Ситтера при $t \to \infty$.

Глобальное решения, возникающие при $\Lambda < 0$, $M < 0$, получается из рассмотренного выше решения преобразованием $\tau \leftrightarrow \sigma$. Соответствующие решения статичны. Диаграммы Картера–Пенроуза также имеют вид линзы и получаются из предыдущего случая поворотом на угол $\pi/2$.

Черное кольцо $\Lambda M < 0$

Если $\Lambda > 0$ и $M < 0$, то конформный множитель (27.73) имеет один простой нуль, и пространство-время содержит один горизонт. В этом случае имеется два статических и два однородных конформных блока, которые вместе склеиваются вдоль горизонта в диаграмму Картера–Пенроуза, изображенную на рис.27.4 справа. Она имеет тот же вид, что и в сферически симметричном случае $K = 1$, $\Lambda > 0$ и $M < 0$ (рис.**??**). Метрика для однородного решения имеет вид (27.74), но из-за условия (27.76) области определения координат t и τ меняются. Для статического решения метрика задается уравнением (**??**) с соответствующей областью определения координат r и σ, связанных тем же уравнением, что и t и τ.

Решения, возникающие при $\Lambda < 0$, $M > 0$, получаются из решений, соответствующих $\Lambda > 0$, $M < 0$, путем перестановки пространственной и временной координаты $\tau \leftrightarrow \sigma$. Соответствующую диаграмму Картера–Пенроуза необходимо повернуть на угол $\pi/2$. По виду она совпадает с черной дырой анти де Ситтера. В этом случае после компактификации $\mathbb{V} = \mathbb{T}^2$ горизонт черной дыры имеет вид тора $\mathbb{T}^2$, а пространство-время является асимптотически пространством анти-де Ситтера. Следовательно, в пространстве возникает кольцевая черная дыра (черное кольцо). Напомним, что в этом случае все пространство имеет вид топологического произведения прямой $\mathbb{R}$ и тора $\mathbb{T}^2$.

27.4.3 Гиперболические глобальные решения $K = -1$

При $K = -1$ поверхность $\mathbb{V}$ представляет собой двуполостный гиперболоид $\mathbb{H}^2$ (плоскость Лобачевского). Точнее, верхнюю полу двуполостного гиперболоида (см. раздел 25.2). После компактификации $\mathbb{H}^2$ в качестве поверхности $\mathbb{V}$ получится компактная риманова поверхность рода два и выше. Отметим, что группой изометрий однополостного гиперболоида $\mathbb{H}^2$ является группа преобразований Лоренца $\mathbb{O}(1,2)$. То есть возникают решения вакуумных уравнений Эйнштейна симметричные относительно действия группы Лоренца $\mathbb{SO}(1,2)$ не в пространстве-времени, а на пространственных сечениях $t = \mathsf{const}$.

Метрика двуполостного гиперболоида (27.21) единичного радиуса в гиперболической системе координат имеет вид (25.16)

$$d\Omega_{\text{н}} = d\theta^2 + \mathsf{sh}^{\,2}\theta d\varphi^2. \tag{27.77}$$

Соответствующее вакуумное решение уравнений Эйнштейна можно записать в координатах Шварцшильда

$$ds^2 = \Phi(q)d\zeta^2 - \frac{dq^2}{\Phi(q)} - q^2 d\Omega_{\text{н}}, \tag{27.78}$$

где конформный множитель имеет вид

$$\Phi(q) = -1 - \frac{2M}{q} - \frac{\Lambda q^2}{3}. \tag{27.79}$$

Чтобы описать глобальные гиперболические решения, заметим, что конформный множитель Φ общего вида (27.45), меняет знак на противоположный при преобразовании всех постоянных:

$$K \to -K, \qquad \Lambda \to -\Lambda, \qquad M \to -M.$$

Это означает, что все глобальные решения в случае $K = -1$, можно получить из сферически симметричных решений, если поменять пространственную и временну́ю координату $\tau \leftrightarrow \sigma$ на поверхности $\mathbb{U}$, а также изменить знак космологической постоянной и массы.

Пример 27.4.1. Аналогом решения Шварцшильда является решение с нулевой космологической постоянной $\Lambda = 0$ и отрицательной массой $M < 0$. В этом случае метрика для однородного и статичного конформного блока принимает вид

$$ds^2 = \frac{dt^2}{1+\frac{2M}{t}} - \left(1+\tfrac{2M}{t}\right) d\sigma^2 - t^2 d\Omega_{\text{н}}, \qquad -2M < t < \infty, \tag{27.80}$$

$$ds^2 = \left(1+\tfrac{2M}{r}\right) d\tau^2 - \frac{dr^2}{1+\frac{2M}{r}} - r^2 d\Omega_{\text{н}}, \qquad 0 < r < -2M. \tag{27.81}$$

При этом диаграмму Картера–Пенроуза для решения Шварцшильда необходимо повернуть на угол $\pi/2$, рис.**??**. Соответствующее глобальное решение имеет левую и правую статические сингулярности при $q = r = 0$. Свойства этого глобального решения похожи на свойства двух статических сферически симметричных сингулярностей, рассмотренных в разделе 27.4.1, но с другим асимптотическим поведением. Пространство-время является асимптотически плоским в бесконечно удаленном прошлом и будущем (оба случая соответствуют пределу $q = t \to \infty$). □

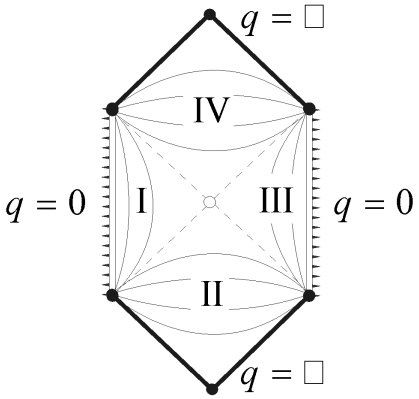


Рис. 27.5: Статические сингулярности в гиперболическом случае.

Остальные глобальные решения для $K = -1$ анализируются аналогично, и мы оставляем это любознательному читателю.

Не так легко представить себе пространство (трехмерное сечение четырехмерного пространства-времени, задаваемое уравнением $t = \mathsf{const}$) при $K = -1$. В рассматриваемом случае оно представляет собой произведение интервала (конечного или бесконечного в зависимости от значений Λ и M) и однополостного гиперболоида $\mathbb{H}^2$. После компактификации, когда поверхность $\mathbb{V}$ представляет собой замкнутую риманову поверхность рода два и выше, глобальное решение можно интерпретировать, как множество кротовых нор, число которых совпадает с родом поверхности (числом ручек). Топология горизонта совпадает с топологией поверхности $\mathbb{V}$.

27.5 Лоренц-инвариантные решения

В настоящем разделе мы рассмотрим случай C (27.17), когда второе дилатонное поле в сплетенном произведении (27.2) постоянно, $m = \mathsf{const}$. В этом случае, как будет

показано ниже, псевдориманова поверхность $\mathbb{U}$ должна быть поверхностью постоянной кривизны. Следовательно, она представляет собой однополостный гиперболоид, $\mathbb{U} = \mathbb{L}^2$ или его универсальную накрывающую, которые были подробно описаны в разделе 25.3. В этом случае глобальные решения вакуумных уравнений Эйнштейна имеют вид топологического произведения $\mathbb{M} = \mathbb{L}^2 \times \mathbb{V}$. Второй сомножитель $\mathbb{V}$ представляет собой максимально продолженную риманову поверхность с одним вектором Киллинга, которые были рассмотрены в разделе 26.10. Как было показано, поверхность $\mathbb{V}$ может иметь конические сингулярности или сингулярности кривизны вдоль края поверхности $\mathbb{V}$. С физической точки зрения этим сингулярностям соответствуют космические струны или сингулярные доменные стенки, которые эволюционируют во времени.

Случай C похож на пространственно симметричные решения, рассмотренные в случае B, однако имеет также несколько существенно новых черт. Во-первых, мы не можем ограничить себя только положительно определенными метриками $h_{\mu\nu}$ на $\mathbb{V}$, потому что уравнения Эйнштейна (27.14) неинвариантны относительно преобразования $h_{\mu\nu} \to -h_{\mu\nu}$ при заданном m. Отметим, что при $k = \mathsf{const}$ преобразование $g_{\alpha\beta} \to -g_{\alpha\beta}$ всегда можно дополнить перестановкой пространственной и временно́й координаты на $\mathbb{U}$, $\tau \leftrightarrow \sigma$, что вместе оставляют уравнение (27.13) инвариантным. В случае евклидовой метрики на $\mathbb{V}$ это невозможно. Поэтому, не ограничивая общности, мы зафиксируем $m = 1$, но допустим, что метрика $h_{\mu\nu}$ может быть как положительно, так и отрицательно определена. В обоих случаях сигнатура четырехмерной метрики будет лоренцевой: либо $(+---)$, либо $(+--+)$.

Решение уравнений (27.13)–(27.16) проводится так же, как и для метрики лоренцевой сигнатуры, при этом необходимо функцию m заменить на k и метрику $g_{\alpha\beta}$ на $h_{\mu\nu}$. Поэтому мы только кратко обозначим основные этапы вычислений, подчеркнув те моменты, которые специфичны для евклидовой сигнатуры.

При $m = 1$ полная система вакуумных уравнений Эйнштейна (27.13)–(27.16) принимает вид

$$\nabla_\mu \nabla_\nu k - \frac{\nabla_\mu k \nabla_\nu k}{2k} - \frac{1}{2} h_{\mu\nu} \left[\nabla^2 k - \frac{(\nabla k)^2}{2k} \right] = 0, \tag{27.82}$$

$$R^g + \nabla^2 k - 2k\Lambda = 0, \tag{27.83}$$

$$R^h + \frac{\nabla^2 k}{k} - \frac{(\nabla k)^2}{2k^2} - 2\Lambda = 0. \tag{27.84}$$

Как и в случае B, в уравнение (27.83) входит сумма функций от разных аргументов: $R^g = R^g(x)$ и $k = k(y)$. Поэтому скалярная кривизна поверхности $\mathbb{U}$ должна быть постоянна, $R^g = -2K = \mathsf{const}$. Отсюда вытекает, что поверхность $\mathbb{U}$ является однополостным гиперболоидом $\mathbb{L}^2$ или его универсальной накрывающей.

Это – очень важное следствие вакуумных уравнений Эйнштейна, т.к. в рассматриваемом случае C все решения должны быть $\mathbb{SO}(1,2)$ инвариантны, где группа преобразований Лоренца $\mathbb{SO}(1,2)$ действует на однополостном гиперболоиде. Поэтому глобальные решения класса C названы лоренц-инвариантными.

Тогда уравнение (27.83) принимает вид

$$\nabla^2 k - 2(k\Lambda + K) = 0. \tag{27.85}$$

Как и в случае B уравнение (27.84) является следствием уравнений (27.82) и (27.85). Поэтому для нахождения решений вакуумных уравнений Эйнштейна достаточно решить уравнения (27.82) и (27.85).

Следующий шаг состоит в фиксировании координат на поверхности $\mathbb{V}$. Конформно евклидова метрика на поверхности $\mathbb{V}$ имеет вид

$$h_{\mu\nu}dy^\mu dy^\nu = \Phi dz d\bar{z} = \Phi(d\sigma^2 + d\rho^2). \tag{27.86}$$

Здесь $\Phi(z,\bar{z})$ является функцией комплексных координат

$$z = \sigma + i\rho, \qquad \bar{z} = \sigma - i\rho, \tag{27.87}$$

где $\sigma = y^2$, $\rho = y^3$. При этом метрика всего четырехмерного пространства-времени равна

$$ds^2 = k d\Omega_{\text{L}} + \Phi dz d\bar{z}, \tag{27.88}$$

где $d\Omega_{\text{L}}$ – метрика постоянной кривизны на однополостном гиперболоиде $\mathbb{L}^2$, заданная, например, уравнением (27.20).

Не ограничивая общности, рассмотрим положительные $k > 0$. В противном случае можно просто переставить первые две координаты. Тогда удобно ввести параметризацию

$$k = q^2, \qquad q > 0.$$

Для двух неизвестных функций q и Φ вместо уравнений (27.32)–(27.34) возникает следующая система уравнений

$$\partial^2_{zz} q - \frac{\partial_z \Phi \partial_z q}{\Phi} = 0, \tag{27.89}$$

$$\partial^2_{\bar{z}\bar{z}} q - \frac{\partial_{\bar{z}} \Phi \partial_{\bar{z}} k}{\Phi} = 0, \tag{27.90}$$

$$2\frac{\partial_z \partial_{\bar{z}} q^2}{\Phi} - (K + \Lambda q^2) = 0. \tag{27.91}$$

Аналогично случаю B, решением уравнений (27.89) и (27.90) являются функции одного аргумента: $q = q(z \pm \bar{z})$ и $\Phi = \Phi(z \pm \bar{z})$, где функция Φ определяется уравнением

$$|\Phi| = |q'|, \tag{27.92}$$

где штрих обозначает дифференцирование по соответствующему аргументу. В полученной формуле нижний и верхний знаки соответствуют положительно и отрицательно определенной римановой метрике на $\mathbb{V}$. Таким образом, функции q и Φ зависят либо от координаты σ, либо от $i\rho$. Поскольку, благодаря вращательной $\mathbb{SO}(2)$ симметрии конформно евклидовой метрики (27.86), оба выбора равнозначны, то для определенности мы предположим, что функции $q(\sigma)$ и $\Phi(\sigma)$ зависят от σ.

После этого уравнение (27.91) упростится:

$$\frac{1}{2}(q^2)'' = (K + \Lambda q^2)\Phi, \tag{27.93}$$

где штрих обозначает дифференцирование по σ. Чтобы его проинтегрировать, в уравнении (27.92) необходимо раскрыть знаки модулей.

Рассмотрим случай $\Phi q' > 0$. Тогда уравнение (27.93) с учетом (27.92) примет вид

$$\frac{1}{2}(q^2)'' = (K + \Lambda q^2)q',$$

и его легко проинтегрировать:

$$q' = K - \frac{2M}{q} + \frac{\Lambda q^2}{3}, \tag{27.94}$$

где M – произвольная постоянная интегрирования. Хотя в рассматриваемом случае ее нельзя интерпретировать как массу, мы будем использовать старые обозначения, чтобы облегчить сравнение. Учитывая уравнение (27.92) получаем выражение для конформного множителя

$$\Phi(q) = K - \frac{2M}{q} + \frac{\Lambda q^2}{3} \tag{27.95}$$

Случай $\Phi q' < 0$ интегрируется аналогично.

Окончательно, общее решение вакуумных уравнений Эйнштейна в случае C имеет вид

$$ds^2 = q^2 d\Omega_{\text{L}} + \Phi(q)(d\sigma^2 + d\rho^2), \tag{27.96}$$

где конформный множитель задан уравнением (27.95) и функция $q = q(\sigma)$ определяется уравнением (27.92). Таким образом, метрика на поверхности $\mathbb{V}$ имеет один вектор Киллинга ∂_ρ и имеет тот же вид, что и в разделе 26.8. Поэтому мы можем использовать развитую там технику для построения глобальных решений вакуумных уравнений Эйнштейна.

Выбирая функцию $q(\sigma)$ в качестве одной из координат, метрику (27.96) можно записать в виде, напоминающем метрику Шварцшильда,

$$ds^2 = q^2 d\Omega_{\text{L}} + \frac{dq^2}{\Phi(q)} + \Phi(q) d\sigma^2. \tag{27.97}$$

Результирующая метрика имеет три вектора Киллинга, соответствующих группе симметрии $\mathbb{SO}(1,2)$ однополостного гиперболоида постоянной кривизны $\mathbb{L}^2$, и один дополнительный вектор Киллинга ∂_ρ на поверхности $\mathbb{V}$.

Вычисления, аналогичные случаю $k = 1$, приводят к следующему выражению для скалярной кривизны поверхности $\mathbb{V}$

$$R^h = \frac{2}{3}\Lambda + \frac{4M}{q^3}.$$

При этом для инвариантного собственного значения тензора Вейля мы получаем то же выражение (27.55), что и в случае B.

27.5.1 Лоренц-инвариантные решения $K = 1$

Прежде всего отметим, что случаи $K = 1$ и $K = -1$ связаны между собой перестановкой первых двух координат $\tau \leftrightarrow \sigma$. Мы выберем значение $K = 1$, чтобы выражение для конформного множителя Φ имело, с точностью до изменения знака космологической постоянной, тот же вид, что и для сферически симметричного случая. Метрика для однополостного гиперболоида в гиперболической полярной системе координат имеет вид (25.28). Поэтому четырехмерная метрика пространства-времени в координатах Шварцшильда запишется следующим образом

$$ds^2 = q^2(d\theta^2 - \mathsf{ch}^2\theta d\varphi^2) + \frac{dq^2}{\Phi(q)} + \Phi(q) d\rho^2, \tag{27.98}$$

где конформный множитель,

$$\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}, \tag{27.99}$$

имеет тот же вид, что и в решении Коттлера [?], но в рассматриваемом случае этот нетривиальный конформный множитель входит в евклидову часть метрики.

На поверхности $\mathbb{V}$ метрика может быть как отрицательно ($\Phi < 0$), так и положительно ($\Phi > 0$) определена. Для отрицательно определенной метрики сигнатура метрики пространства-времени равна $(+---)$, и роль времени играет координата θ. Поэтому времениподобная координата принимает значения на всей вещественной оси $\theta \in \mathbb{R}$, и трехмерное пространство представляет собой произведение окружности $\varphi \in [0, 2\pi)$ и поверхности $\mathbb{V}$, которая будет построена ниже. Если в качестве $\mathbb{U}$ выбрать универсальную накрывающую однополостного гиперболоида $\mathbb{L}^2$, то трехмерное пространство будет представлять произведение $\mathbb{R} \times \mathbb{V}$. Эволюция этих пространств во времени длится вечно, и если поверхность $\mathbb{V}$ имеет сингулярность, то ей будет соответствовать времениподобная кривая.

Для положительно определенной метрики на $\mathbb{V}$ сигнатура четырехмерной метрики равна $(+--+)$, и времениподобной координатой является угол φ. При $\mathbb{U} = \mathbb{L}^2$ он принимает значения на окружности $\varphi \in [0, 2\pi)$, и трехмерное пространство представляет собой произведение прямой $\theta \in \mathbb{R}$ и поверхности $\mathbb{V}$. Соответствующее пространство-время содержит замкнутые времениподобные кривые (включая экстремали), если только в качестве поверхности $\mathbb{U}$ не выбрано универсальное накрывающее пространство для $\mathbb{L}^2$.

Поскольку метрика на поверхности $\mathbb{V}$ имеет тот же локальный вид, что и в разделе 26.8, то мы может построить максимально продолженные поверхности $\mathbb{V}$. Вид римановой поверхности определяется конформным множителем, показанным на рис.27.1 при $\Lambda \neq 0$. Надо только помнить, что в выражение для конформного множителя космологической постоянная входит с противоположным знаком. Кроме того, при отождествлении $\rho \sim \rho + L$ возможно появление конических сингулярностей, которые соответствуют космическим струнам.

Перейдем к классификации глобальных решений вакуумных уравнений Эйнштейна в рассматриваемом случае C.

Пространство-время Минковского $\Lambda = 0$, $M = 0$

При нулевых значениях постоянных поверхность $\mathbb{V}$ представляет собой полуплоскость $q \in (0, \infty)$, $\rho \in (-\infty, \infty)$. Она неполна при $q = 0$ из-за координатной сингулярности, аналогичной сферически симметричному случаю.

Космическая струна $\Lambda = 0$, $M > 0$

При нулевой космологической постоянной $\Lambda = 0$ и положительных значениях $M > 0$ возникает решение соответствующее решению Шварцшильда. В этом случае существуют две несвязные между собой максимально продолженные римановы поверхности, изображенные на рис.27.6 слева и в центре с положительно и отрицательно определенной метрикой. Для наглядности мы отождествили точки с координатами ρ и $\rho + L$.

Внешнему решению Шварцшильда соответствует интервал $q \in (2M, \infty)$. В этой области конформный множитель положителен и четырехмерная метрика имеет сигнатуру $(+--+)$. Соответствующая поверхность $\mathbb{V}$ является гладким многообразием за исключением возможной конической сингулярности на горизонте $q = 2M$. С топологической точки зрения коническая сингулярность представляет собой плоскость (мировую поверхность космической струны), а все пространство-время описывает

бесконечную эволюцию бесконечной космической струны. Для этого решения кривизна не имеет больше никаких сингулярностей, и если выполнено уравнение (**??**), то пространство-время вообще не имеет сингулярностей. Специфическим свойством этого решения является то, что периметр цилиндра, изображенного на рис.27.6 слева, стремится к постоянной при $q \to \infty$. Это свойство соответствует тому, что решение Шварцшильда является асимптотически плоским.

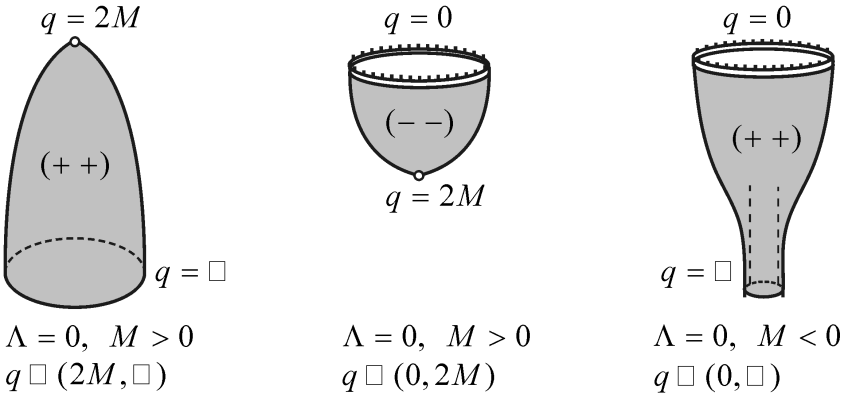


Рис. 27.6: Максимально продолженная риманова поверхность $\mathbb{V}$ при нулевой космологической постоянной $\Lambda = 0$. Условные обозначения те же, что и на рис. 26.3. Незакрашенные окружности обозначают возможные конические сингулярности.

Для внутреннего решения Шварцшильда $q \in (0, 2M)$, конформный множитель отрицателен, и метрика пространства-времени имеет сигнатуру $(+ - --)$. Максимально продолженная поверхность $\mathbb{V}$ представляет собой в этом случае диск конечного радиуса с сингулярным краем $q = 0$, который изображен на рис.27.6 в центре. Если выполнено уравнение (**??**), то поверхность $\mathbb{V}$ гладкая. В противном случае в центре диска $q = 2M$ возникает коническая сингулярность, соответствующая космической струне. Поэтому трехмерное пространство представляет собой космическую струну, окруженную цилиндрической доменной стенкой, на которой кривизна сингулярна. Сингулярный край невозможно адекватно изобразить на рисунке, поскольку он находится на конечном расстоянии, но в то же время имеет бесконечный периметр.

Доменная стенка $\Lambda = 0$, $M < 0$

Если космологическая постоянная равна нулю, а масса отрицательна, то решение Шварцшильда описывает голую сингулярность. В этом случае конформный множитель при $q \in (0, \infty)$ положителен и сигнатура метрики равна $(+ - ++)$. Максимально продолженная риманова поверхность $\mathbb{V}$ представляет собой полуплоскость, $q \in (0, \infty)$, $\rho \in \mathbb{R}$. Сингулярность кривизны расположена вдоль края $q = 0$, что можно интерпретировать, как доменную стенку, на которой кривизна сингулярна. С четырехмерной точки зрения это решение описывает эволюцию доменной стенки, которая расположена на конечном расстоянии.

Благодаря трансляционной симметрии, точки с координатами ρ и $\rho + L$ можно отождествить. Результирующая поверхность изображена на рис.27.6 справа. Соответствующее трехмерное пространство представляет собой цилиндрическую доменную стенку.

Решение де Ситтера $\Lambda > 0$, $M = 0$

При нулевой массе и положительной космологической постоянной конформный множитель (27.95) положителен для $q \in (0, \infty)$. Поверхность $\mathbb{V}$ имеет метрику с сигнатурой $(+ - ++)$ и показана на рис. 27.7 слева. Для наглядности она изображены после отождествления точек с координатами ρ и $\rho + L$. Пространство де Ситтера соответствует универсальной накрывающей поверхности $\mathbb{V}$. Поверхность $\mathbb{V}$ является гладкой и не имеет сингулярностей. Однако она неполна при $q = 0$. Ее продолжение через край $q = 0$ невозможно, т.к. четырехмерная метрика пространства-времени вырождается.

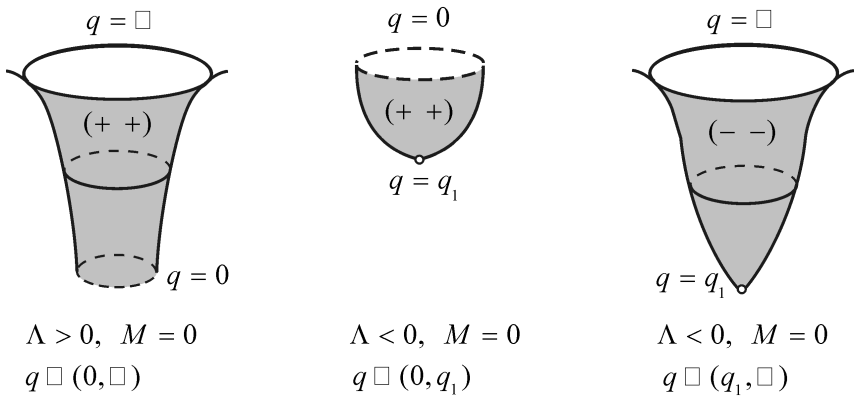


Рис. 27.7: Максимально продолженные римановы поверхности $\mathbb{V}$ при нулевой массе $M = 0$. Условные обозначения те же, что и на рис. 26.3. Незакрашенные окружности обозначают возможные конические сингулярности.

Решение анти-де Ситтера $\Lambda < 0$, $M = 0$

При отрицательной космологической постоянной $\Lambda < 0$ и нулевой константе $M = 0$ возникает решение анти-де Ситтера. Конформный множитель при этом имеет один простой нуль (горизонт) в точке $q_1 = \sqrt{-3/\Lambda}$. Следовательно, существуют две несвязные поверхности $\mathbb{V}$: одна с положительно, а другая с отрицательно определенной метрикой.

При $q \in (0, q_1)$ конформный множитель положителен, метрика пространства-времени имеет сигнатуру $(+ - ++)$, и поверхность $\mathbb{V}$ показана на рис.27.7 в центре. С топологической точки зрения эта поверхность представляет собой диск конечного радиуса с положительно определенной метрикой. В центре диска $q = q_1$ возможно появление конической сингулярности. Ее продолжение через край $q = 0$ невозможно, т.к. четырехмерная метрика вырождается.

Если $q \in (q_1, \infty)$, то конформный множитель отрицателен и метрика пространства-времени имеет сигнатуру $(+ - --)$. Соответствующая поверхность $\mathbb{V}$ показана на рис.27.7 справа. С топологической точки зрения она представляет собой плоскость, в центре которой возможно появление конической сингулярности.

Доменная стенка $\Lambda < 0$, $M > \frac{1}{3\sqrt{-\Lambda}}$

При отрицательной космологической постоянной $\Lambda < 0$ и $M > \frac{1}{3\sqrt{-\Lambda}}$ конформный множитель отрицателен при $q \in (0, \infty)$, не имеет нулей, и сигнатура метрики

пространства-времени равна $(+---)$. С топологической точки зрения поверхность $\mathbb{V}$ представляет собой полуплоскость $q > 0$, $\rho \in \mathbb{R}$ с отрицательно определенной метрикой и сингулярной кривизной на границе $q = 0$. Глобальное решение описывает бесконечную эволюцию бесконечной плоской доменной стенки сингулярной кривизны.

Компактификация по координате ρ дает внутренность цилиндрической доменной стенки, показанной на рис.27.8. Длина направляющей окружности цилиндра стремится к бесконечности при $q \to 0$ и $q \to \infty$. Сингулярный край $q = 0$ находится на конечном расстоянии.

Двойной горизонт $\Lambda < 0$, $M = \frac{1}{3\sqrt{-\Lambda}}$

При указанных значениях постоянных конформный множитель отрицателен и имеет двойной нуль в точке q_1. Метрика пространства-времени в обоих интервалах $q \in (0, q_1)$ и $q \in (q_1, \infty)$ имеет сигнатуру $(+---)$. В этом случае имеем две поверхности, изображенные на рис.27.8. Поверхность $\mathbb{V}$, соответствующая интервалу $q \in (q_1, \infty)$, не является односвязной. Ее можно представить себе в виде плоскости, центр которой $q = q_1$ удален в бесконечность, т.к. он лежит на бесконечном расстоянии.

Поверхность для интервала $(0, q_1)$ представляет собой цилиндрическую доменную стенку сингулярностей кривизны. Сингулярный край $q = 0$ лежит на конечном расстоянии.

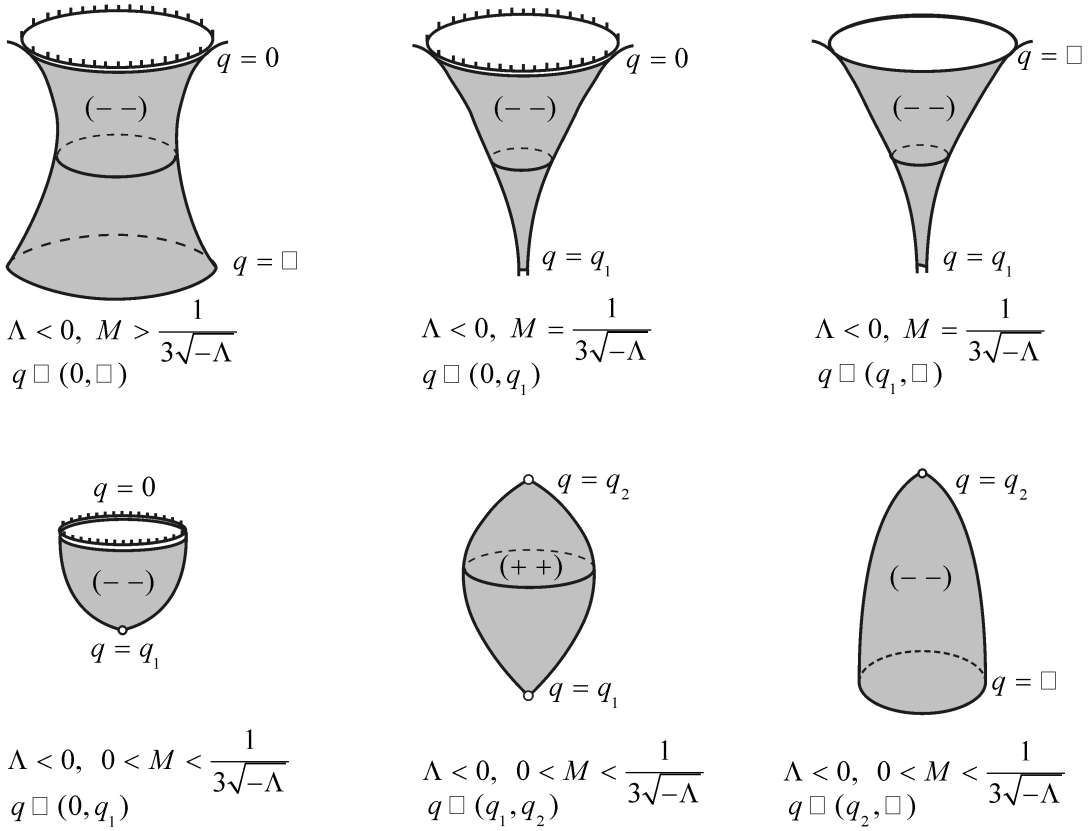


Рис. 27.8: Максимально продолженные римановы поверхности $\mathbb{V}$ при отрицательной космологической постоянной $\Lambda < 0$ и $M \neq 0$. Незакрашенные окружности обозначают возможные конические сингулярности.

Два горизонта $\Lambda < 0$, $0 < M < \frac{1}{3\sqrt{-\Lambda}}$

Если космологическая постоянная отрицательна $\Lambda < 0$ и масса лежит в интервале $0 < M < \frac{1}{3\sqrt{-\Lambda}}$, то конформный множитель имеет максимальное число нулей – два простых нуля – расположенных в точках $q_{1,2}$ (см. рис.27.1 слева). В интервалах $(0, q_1)$ и (q_2, ∞) конформный множитель отрицателен, и сигнатура метрики пространства-времени равна $(+---)$. В интервале (q_1, q_2) конформный множитель положителен, и сигнатура метрики имеет вид $(+--+)$. Максимально продолженная поверхность $\mathbb{V}$ является одной из трех поверхностей, показанных на рис. 27.8 в нижнем ряду.

При $q \in (0, q_1)$ поверхность $\mathbb{V}$ представляет собой диск конечного радиуса, на крае которого кривизна сингулярна. В центре диска, возможно, расположена коническая сингулярность. Это зависит от отождествления $\rho \sim \rho + L$.

Если $q \in (q_1, q_2)$, то поверхность $\mathbb{V}$ топологически представляет собой сферу, в общем случае имеющую две конические сингулярности. Алгебраический анализ системы уравнений $\Phi(q_1) = \Phi(q_2) = 0$ и $\Phi'(q_1) = -\Phi'(q_2)$ при $q_1 \neq q_2$ показывает, что таких решений не существует. Поэтому, подгоняя период компактификации L, можно устранить одну из конических сингулярностей, но не обе одновременно. Это значит, что в рассматриваемом случае всегда должна существовать космическая струна. С топологической точки зрения трехмерное пространство представляет собой произведение сферы, содержащей одну или две конические сингулярности, и прямой.

При $q \in (q_2, \infty)$ поверхность $\mathbb{V}$ – это плоскость, в центре которой, возможно, расположена коническая сингулярность. Экстремали при $q \to \infty$ полны, а кривизна стремится к постоянной.

Космическая струна $\Lambda < 0$, $M < 0$

Если и космологическая постоянная $\Lambda < 0$, и масса $M < 0$, то конформный множитель Φ имеет один простой нуль в точке q_1. В этом случае возможно существование двух поверхностей $\mathbb{V}$.

При $q \in (0, q_1)$ конформный множитель положителен, и сигнатура метрики пространства-времени равна $(+--+)$. Топологически поверхность $\mathbb{V}$ имеет то же вид, что и рассмотренная в предыдущем случае поверхность для $\Lambda < 0$, $0 < M < \frac{1}{3\sqrt{-\Lambda}}$ и $q \in (0, q_1)$, но с отрицательно определенной метрикой.

Вторая поверхность $\mathbb{V}$ соответствует интервалу $q \in (q_1, \infty)$. Конформный множитель отрицателен, и сигнатура метрики пространства-времени равна $(+---)$. Поверхность $\mathbb{V}$ тот же вид, что и в случае $\Lambda < 0$, $M = 0$ и $q \in (q_1, \infty)$ на рис.27.7. Глобальное решение вакуумных уравнений Эйнштейна в общем случае описывает бесконечную космическую струну, причем других сингулярностей у кривизны нет.

Космическая струна $\Lambda > 0$, $M > 0$

При положительной космологической постоянной $\Lambda > 0$ и массе $M > 0$. Конформный множитель имеет один простой нуль в точке q_1, как видно из рис.27.1, справа. Он отрицателен в интервале $q \in (0, q_1)$ и положителен при $q \in (q_1, \infty)$. Этот случай аналогичен предыдущему. Мы имеем две максимально продолженные поверхности $\mathbb{V}$, на которых необходимо просто изменить сигнатуру метрики.

Доменная стенка $\Lambda > 0$, $M < 0$

Если космологическая постоянная положительна, $\Lambda > 0$, а масса отрицательна, то конформный множитель положителен при $q \in (0, \infty)$. Сигнатура метрики пространства-времени равна $(+ - ++)$. В этом случае имеется одна поверхность $\mathbb{V}$, т.к. нули у конформного множителя отсутствуют. Топологически поверхность $\mathbb{V}$ такая же, как и в случае $\Lambda < 0$, $M > 1/3\sqrt{-\Lambda}$, но с положительно определенной метрикой.

27.5.2 Решения с плоскостью Минковского $K = 0$

При $K = 0$ поверхность $\mathbb{U}$ представляет собой или плоскость Минковского $\mathbb{R}^{1,1}$, или цилиндр, или тор. При этом возникают новые решения, интересные с топологической точки зрения. Соответствующая метрика в координатах Шварцшильда имеет вид

$$ds^2 = q^2(dt^2 - dx^2) + \frac{dq^2}{\Phi(q)} + \Phi(q)d\rho^2, \tag{27.100}$$

где

$$\Phi(q) = -\frac{2M}{q} + \frac{\Lambda q^2}{3}.$$

В зависимости от значения постоянных Λ и M, входящих в конформный множитель, возможны четыре качественно отличных случая.

Доменная стенка $\Lambda = 0$, $M \neq 0$

При нулевой космологической постоянной $\Lambda = 0$, но отличной от нуля массе $M \neq 0$ максимально продолженная поверхность $\mathbb{V}$ изображена на рис. 27.9. Метрика на этой поверхности положительно и отрицательно определена соответственно при $M < 0$ и $M > 0$.

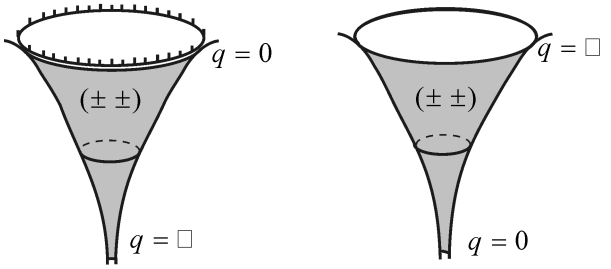


Рис. 27.9: Глобальные решения с плоскостью Минковского, $K = 0$.

Решение (анти-)де Ситтера $\Lambda \neq 0$, $M = 0$

Если масса равна нулю, то возникает еще одно представление решения (анти-)де Ситтера для положительной и отрицательной космологической постоянной. Соответствующая поверхность $\mathbb{V}$ показана на рис. 27.9.

Пусть $\Phi q' > 0$ или, что эквивалентно, $\Lambda q' > 0$. Тогда уравнение (26.99) принимает вид

$$q' = \frac{\Lambda q^2}{3}$$

и легко интегрируется:

$$q = -\frac{3}{\Lambda\sigma},$$

где мы опустили несущественную постоянную интегрирования, соответствующую сдвигу σ. Тогда метрика в черепашьих координатах равна

$$ds^2 = \frac{9}{\Lambda^2\sigma^2}(dt^2 - dx^2) + \frac{3}{\Lambda\sigma^2}(d\sigma^2 + d\rho^2). \tag{27.101}$$

Растяжением координат t и x она приводится к конформно плоскому виду.

Космическая струна $\Lambda > 0$, $M > 0$

Конформный множитель Φ при $\Lambda > 0$ и $M > 0$ имеет один простой нуль в точке q_1. При $q \in (0, q_1)$ и $q \in (q_1, \infty)$, конформный множитель, соответственно, отрицателен и положителен. Поэтому метрики пространства-времени имеют сигнатуры $(+---)$ и $(+-++)$. Максимально продолженные поверхности $\mathbb{V}$ топологически такие же как и для конформного множителя $\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}$, который был рассмотрен ранее. Меняется только величина q_1. Глобальные решения описывают космическую струну и космическую струну, окруженную доменной стенкой сингулярностей кривизны.

Изменение знака обеих постоянных $\Lambda < 0$, $M > 0$ приводит к изменению сигнатуры метрики. Качественные свойства решений остаются прежними.

Доменная стенка $\Lambda > 0$, $M < 0$

В рассматриваемом случае конформный множитель Φ положителен и не имеет нулей. Сигнатура метрики пространства-времени равна $(+-++)$. Существует всего одна максимально продолженная поверхность $\mathbb{V}$. Топологически она такая же, как и для конформного множителя $\Phi = 1 - \frac{2M}{q} + \frac{\Lambda q^2}{3}$, рассмотренного ранее.

При $\Lambda < 0$ и $M > 0$ необходимо просто изменить сигнатуру всей метрики.

27.6 Итоги главы

В заключение главы отметим следующее. Проведенный анализ и классификация глобальных решений, исчерпывает все вакуумные решения уравнений Эйнштейна с космологической постоянной, которые имеют вид сплетенного произведения двух поверхностей. Мы видим, что требование максимального продолжения решений практически однозначно определяет глобальную структуру пространства-времени. Важно отметить, что при решении уравнений движения мы не ставим никаких граничных условий. Явное построение и классификация решений проведена в зависимости от значений постоянной скалярной кривизны одной из поверхностей, значения космологической постоянной и единственной постоянной интегрирования, которая для решения Шварцшильда имеет физический смысл массы черной дыры. Подчеркнем, что решение уравнений Эйнштейна в какой то фиксированной системе координат дает не так уж много. Для того, чтобы дать физическую интерпретацию решений необходимо знать глобальную структуру пространства-времени. Эта задача сложна, но обойти ее, по-видимому, невозможно.

Построенные решения представляют значительный физический интерес. Мы показали, что вакуумные решения уравнений Эйнштейна описывают черные дыры, космические струны, кротовые норы, доменные стенки сингулярностей кривизны и

другие. В настоящей монографии мы лишь кратко обсудили свойства построенных глобальных решений.

Глава 28

Дополнение

28.1 Матрицы

Мы рассматриваем вещественные или комплексные матрицы. Бóльшую часть утверждений и формул настоящего дополнения можно найти, например, в [**?**, **?**].

Для любого натурального $n \in \mathbb{N}$ справедливо разложение

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^{n} C_n^k a^{n-k} b^k \\ &= a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + \ldots + nab^{n-1} + b^n,\end{aligned} \tag{28.1}$$

которое называется *биномом Ньютона*, где

$$C_n^k = \frac{n!}{k!(n-k)!} \tag{28.2}$$

– *биномиальные коэффициенты*. В частности, при $a = b = 1$ и $a = -b = 1$ справедливы равенства

$$2^n = \sum_{k=0}^{n} C_n^k, \qquad 0 = \sum_{k=0}^{n} (-1)^k C_n^k. \tag{28.3}$$

Пусть дана 2×2 матрица

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Если ее определитель отличен от нуля, $\det A = ad - bc \neq 0$, то у нее существует обратная матрица, которая имеет вид

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Для двух 2×2 матриц A и B справедлива формула

$$\det(A+B) = \det A + \operatorname{tr} A \operatorname{tr} B - \operatorname{tr}(AB) + \det B. \tag{28.4}$$

Пусть A – произвольная 3×3-матрица. Тогда справедлива формула (Bengtsson gr-qc/0703114)

$$A^3 - A^2 \operatorname{tr} A - \frac{1}{2} A \left(\operatorname{tr} A^2 - (\operatorname{tr} A)^2 \right) - \det A = 0. \tag{28.5}$$

След этого выражения имеет вид

$$\mathsf{tr}\, A^3 - \frac{3}{2}\,\mathsf{tr}\, A^2\,\mathsf{tr}\, A + \frac{1}{2}(\,\mathsf{tr}\, A)^3 - 3\,\mathsf{det}\, A = 0.$$

Пусть $A = (A_{ab})$ – произвольная квадратная невырожденная $n \times n$ матрица, $\mathsf{det}\, A \neq 0$. Тогда элементы обратной матрица A^{-1ab} равны

$$A^{-1ab} = \frac{A^{ba}}{\mathsf{det}\, A},$$

где A^{ba} – алгебраическое дополнение элемента A_{ba}. Если матрица A симметрична, то обратная матрица A^{-1} также симметрична. Для произведения двух невырожденных квадратных матриц справедлива формула $(AB)^{-1} = B^{-1}A^{-1}$. Кроме того $\mathsf{det}\, A^{-1}\,\mathsf{det}\, A = 1$.

Определитель квадратной $n \times n$ матрицы A является полиномом от n^2 элементов матрицы. Поэтому его можно дифференцировать по каждому элементу. Если матрица A невырождена, то справедлива формула

$$A^{-1ba} = \frac{1}{\mathsf{det}\, A}\frac{\partial\,\mathsf{det}\, A}{\partial A_{ab}}. \tag{28.6}$$

Пусть A и B – две квадратные обратимые матрицы размеров $m \times m$ и $n \times n$, соответственно, и пусть C – произвольная матрица размера $n \times m$. Тогда для квадратных $(m+n) \times (m+n)$-матриц справедливо равенство

$$\begin{aligned}\mathsf{det}\begin{pmatrix} A & -C^{\mathrm{T}} \\ C & B \end{pmatrix} &= \mathsf{det}\, A\ \mathsf{det}\,(B + CA^{-1}C^{\mathrm{T}}) \\ &= \mathsf{det}\, B\ \mathsf{det}\,(A + C^{\mathrm{T}}B^{-1}C).\end{aligned} \tag{28.7}$$

Определитель нижне треугольной матрицы равен произведению определителей диагональных блоков:

$$\mathsf{det}\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \mathsf{det}\, A\,\mathsf{det}\, B. \tag{28.8}$$

Рассмотрим квадратную $m \times m$ матрицу C, которая равна произведению двух прямоугольных матриц

$$C = AB.$$

Размеры матриц A и B равны, соответственно, $m \times n$ и $n \times m$. Тогда справедлива формула Бине–Коши [**?**]

$$\mathsf{det}\, C = \sum_{1 \le \beta_1 < \beta_2 < \ldots < \beta_m \le n} \mathsf{det}\begin{pmatrix} A_{1\beta_1} & \ldots & A_{1\beta_m} \\ \vdots & \vdots & \vdots \\ A_{m\beta_1} & \ldots & A_{m\beta_m} \end{pmatrix} \mathsf{det}\begin{pmatrix} B_{\beta_1 1} & \ldots & B_{\beta_m 1} \\ \vdots & \vdots & \vdots \\ B_{\beta_1 m} & \ldots & B_{\beta_m m} \end{pmatrix}, \tag{28.9}$$

где сумма берется по всем упорядоченным выборкам m индексов $\beta_1, \ldots, \beta_m$ из $1, \ldots, n$. Если $m > n$, то сумма считается равной нулю. Согласно формуле Бине–Коши определитель матрицы C равен сумме произведений всех возможных миноров максимального m-того порядка матриц A и B. При $m > n$ $\mathsf{det}\, C = 0$.

Если матрица A невырождена, то справедливо тождество

$$\begin{aligned}\mathsf{det}\,(A + B) &= \mathsf{det}\, A\,\mathsf{exp}\,\Big(\mathsf{tr}\,\ln(1 + A^{-1}B)\Big) \\ &= \mathsf{det}\ A\ \mathsf{exp}\left[\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\,\mathsf{tr}\,(A^{-1}B)^n\right].\end{aligned} \tag{28.10}$$

В частном случае, если $A = 1$, то

$$\det(1 + B) = 1 + \operatorname{tr} B + \frac{1}{2}(\operatorname{tr}^2 B - \operatorname{tr} B^2) + \ldots + \det B, \tag{28.11}$$

где точки обозначают сумму определителей матриц, получающихся из B всеми возможными вычеркиваниями одинаковых строк и столбцов.

Если норма матрицы A меньше единицы, то матрица $1 - A$ обратима, и справедлива формула

$$\sum_{k=1}^{\infty} A^k = (1 - A)^{-1}, \tag{28.12}$$

которая является аналогом формулы для суммы членов геометрической прогрессии.

Для произвольных некоммутирующих матриц A и B справедливо равенство

$$\mathrm{e}^{-B} A\, \mathrm{e}^{B} = A + [A, B] + \frac{1}{2}[[A, B], B] + \ldots = \sum_{k=0}^{\infty} \frac{1}{k!}[A, B]_{(k)}, \tag{28.13}$$

где члены ряда определены рекуррентными соотношениями

$$[A, B]_{(0)} = A, \qquad [A, B]_{(k+1)} = \Big[[A, B]_{(k)}, B\Big].$$

В частности, если $[B, A] = \lambda A$, то равенство (28.13) принимает вид

$$\mathrm{e}^{B} A\, \mathrm{e}^{-B} = \mathrm{e}^{\lambda} A. \tag{28.14}$$

Если коммутатор двух матриц пропорционален единичной матрице,

$$[A, B] = c\mathbb{1},$$

то ряд (28.13) обрывается и справедливы следующие формулы:

$$\mathrm{e}^{A}\, \mathrm{e}^{B} = \mathrm{e}^{[A,B]}\, \mathrm{e}^{B}\, \mathrm{e}^{A} = \mathrm{e}^{\frac{1}{2}[A,B]}\, \mathrm{e}^{A+B}, \tag{28.15}$$

$$[A,\, \mathrm{e}^{B}] = [A, B]\, \mathrm{e}^{B}. \tag{28.16}$$

Последнее равенство можно переписать в эквивалентном виде

$$\mathrm{e}^{-B} A\, \mathrm{e}^{B} = A + [A, B]. \tag{28.17}$$

Равенство (28.15) известно, как *формула Хаусдорфа.*

Теорема 28.1.1 (Троттер). *Для двух произвольных вещественных или комплексных квадратных матриц A и B справедлива формула*

$$\mathrm{e}^{t(A+B)} = \lim_{n\to\infty} \left(\mathrm{e}^{\frac{t}{n}A}\, \mathrm{e}^{\frac{t}{n}B}\right)^n. \tag{28.18}$$

Доказательство. [?]. □

Пусть заданы две произвольные симметричные матрицы A и B. Тогда, если равенство

$$A_{ab}B_{cd} - A_{ad}B_{cb} - A_{cd}B_{ab} + A_{cb}B_{ad} = 0 \tag{28.19}$$

выполнено для всех значений индексов, то матрицы A и B пропорциональны.

Если элементы квадратной матрицы $A_{ab}(x)$ являются функциями, то производная от определителя равна

$$\partial_\alpha \det A = \det A\, A^{-1ab}\partial_\alpha A_{ba} = -\det A\, \partial_\alpha A^{-1ab} A_{ba}. \tag{28.20}$$

Приведем без доказательства несколько теорем о свойствах матриц с вещественными элементами, которые часто используются в приложениях. Напомним, что *собственными значениями* λ_i, $i = 1, \ldots, n$ квадратной $n \times n$ матрицы A называются корни характеристического (векового) уравнения

$$\det (A - \lambda \mathbb{1}) = 0.$$

Справедливы формулы

$$\det A = \lambda_1 \ldots \lambda_n, \qquad \mathrm{tr}\, A^k = \sum_{i=1}^{n} \lambda_i^k, \tag{28.21}$$

для любого натурального $k \in \mathbb{N}$. У симметричной матрицы все собственные значения вещественны. При этом симметричная матрица положительно определена тогда и только тогда, когда все собственные значения вещественны и положительны $\lambda_i > 0$.

Если симметричная 2×2 матрица $A = A^T$ имеет положительные собственные значения, то из нее можно извлечь квадратный корень, т.е. представить в виде $A = B^2$. Этот корень не является единственным. Однако, с точностью до знака существует единственный симметричный корень $B = B^T$. Поскольку собственные значения матрицы A положительны, то на диагонали матрицы должны стоять положительные числа. Пусть

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \qquad \det A = ad - b^2 > 0, \qquad a, d > 0.$$

тогда квадратный корень имеет вид

$$B = \begin{pmatrix} \sqrt{a}\cos\theta & \sqrt{a}\sin\theta \\ \sqrt{a}\sin\theta & \sqrt{d}\cos\psi \end{pmatrix},$$

где

$$\sin^2\theta = \frac{b^2}{a(a + d + 2\sqrt{ad - b^2})},$$
$$\sin^2\psi = \frac{b^2}{d(a + d + 2\sqrt{ad - b^2})},$$

и для углов выбираются положительные корни: $0 \le \theta, \psi \le \pi/2$. Отметим, что справедливо тождество

$$\sqrt{a}\sin\theta = \sqrt{d}\sin\psi.$$

Теорема 28.1.2. *Пусть A – произвольная квадратная невырожденная матрица. Тогда матрица $A^T A$ симметрична и положительно определена.*

Пусть $\lambda_i(A^T A) \ge 0$ – собственные значения симметричной матрицы $A^T A$. Тогда *сингулярными числами* $\nu_i(A)$ матрицы A называется набор чисел

$$\nu_i(A) = \sqrt{\lambda_i(A^T A)}.$$

Теорема 28.1.3. О сингулярном разложении матриц. *Пусть A произвольная квадратная матрица и $\nu_i(A)$ – ее сингулярные числа, занумерованные произвольным образом, тогда найдутся ортогональные матрицы R и S, такие что*

$$A = R^T \operatorname{diag}\{\nu_1(A), \ldots, \nu_n(A)\}S,$$

где $\operatorname{diag}\{\nu_1(A), \ldots, \nu_n(A)\}$ – диагональная матрица, у которой на диагонали стоят сингулярные числа.

Теорема 28.1.4. *Пусть A и B произвольные квадратные матрицы с сингулярными числами $\nu_i(A)$ и $\nu_i(B)$, занумерованными в порядке возрастания, тогда справедлива оценка*

$$|\operatorname{tr}(AB)| \leq \sum_{i=1}^{n} \nu_i(A)\nu_i(B).$$

Теорема 28.1.5. *Пусть B – произвольная симметричная положительно определенная матрица. Тогда существует единственная симметричная положительно определенная матрица A, такая что $A^2 = B$. При этом мы пишем $A = \sqrt{B}$.*

Теорема 28.1.6. Полярное разложение вещественных матриц. *Любая вещественная матрица A может быть представлена в виде произведения*

$$A = RS, \tag{28.22}$$

где R – неотрицательно определенная симметричная матрица и S – ортогональная матрица. Матрица R всегда однозначно определена соотношением

$$R = \sqrt{AA^T}.$$

Если матрица A невырождена, то симметричная матрица R является положительно определенной, а ортогональная матрица S также определена однозначно, $S = R^{-1}A$.

Доказательство. См., например, [**?**]. □

Пример 28.1.1. Вещественное число $a \in \mathbb{R}$ является 1×1 матрицей и его можно представить в виде $a = |a| \operatorname{sign} a$, где

$$\operatorname{sign} a := \begin{cases} 1 & a > 0, \\ -1, & a < 0, \end{cases}$$

– ортогональная 1×1 матрица. Если $a \neq 0$, то представление существует и единственно. При $a = 0$ модуль числа также равен нулю, $|a| = 0$, а матрица $S = \operatorname{sign} a$ неопределена. □

Полярное разложение вещественных матриц имеет простую геометрическую интерпретацию. Рассмотрим $n \times n$ матрицу A, как линейный оператор, действующий в евклидовом пространстве $\mathbb{R}^n$. Тогда ортогональной матрице соответствует некоторое вращение евклидова пространства вокруг начала координат. Симметричная положительно определенная матрица осуществляет дилатацию евклидова пространства вдоль n взаимно перпендикулярных направлений с различными в общем случае коэффициентами растяжения. Следовательно, полярное разложение матриц означает последовательное выполнение некоторого вращения и некоторой дилатации.

Теорема 28.1.7. Полярное разложение комплексных матриц. *Любая комплексная матрица A может быть представлена в виде произведения*

$$A = RU, \tag{28.23}$$

где R – неотрицательно определенная эрмитова матрица и U – унитарная матрица. Матрица R всегда однозначно определена соотношением

$$R = \sqrt{AA^\dagger}.$$

Если матрица A невырождена, то эрмитова матрица R является положительно определенной, а унитарная матрица U также определена однозначно, $U = R^{-1}A$.

Доказательство. См., например, [**?**]. □

Пример 28.1.2. Комплексные числа $z = x + iy$ являются комплексными 1×1 матрицами. Для них полярное разложение имеет хорошо известный вид $z = r\,\mathrm{e}^{i\varphi}$. Если $z \neq 0$, то модуль r и аргумент φ комплексного числа определены однозначно. При $z = 0$ модуль комплексного числа также равен нулю, $r = 0$, а аргумент неопределен. □

Замечание. Полярные разложения матриц справедливы также для другого порядка матриц в (28.22) и (28.23): $A = S'R'$ и $A = U'R'$. Вообще говоря, новые матрицы будут отличаться от старых. Сомножители R, S и R, U в полярном разложении матриц перестановочны между собой тогда и только тогда, когда матрицы A, A^T и $A, A^\dagger$ коммутируют, соответственно, для вещественного и комплексного случая. □

Определение. Две квадратные $n \times n$ матрицы A и B называются *подобными*, если существует такая невырожденная матрица S, что

$$B = S^{-1}AS. \tag{28.24}$$

Отображение $A \mapsto S^{-1}AS$ называется *преобразованием подобия.* □

Легко проверить, что преобразование подобия является отношением эквивалентности в множестве квадратных $n \times n$ матриц. Если матрицы A и B подобны, то матрица S, с помощью которой проводится преобразование, определена неоднозначно. Ее можно умножить слева на произвольную невырожденную матрицу, коммутирующую с A.

С помощью преобразования подобия любую матрицу можно преобразовать к некоторому каноническому виду. Одним из таких видов является жорданова форма матрицы.

Определение. *Жордановым блоком* или *жордановой клеткой* $J_k(\lambda)$ называется верхняя треугольная $k \times k$ матрица вида

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & & & & 0 \\ & \lambda & 1 & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & 1 \\ 0 & & & & & \lambda \end{pmatrix}, \tag{28.25}$$

у которой на главной диагонали стоят одинаковые числа λ, а над ними – единицы. Все остальные элементы матрицы равны нулю. По определению $J_1(\lambda) = \lambda$. *Жордановой* $n \times n$ *матрицей* называется любая прямая сумма жордановых клеток

$$J := \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_k}(\lambda_k) \end{pmatrix}, \qquad n_1 + n_2 + \ldots + n_k = n, \tag{28.26}$$

где порядки n_i каких то клеток могут совпадать и числа λ_i не обязательно различны. □

Можно рассматривать, например, вещественные или комплексные жордановы матрицы.

Все собственные числа жордановой клетки одинаковы и равны λ, т.е. λ является собственным числом жордановой клетки (28.25) кратности k. λ_i, $i = 1, \ldots, n$ – собственные числа матрицы A. Если $\lambda_i = \lambda_j$ и отличается от всех остальных собственных чисел, то кратность собственного числа λ_i равна $n_i + n_j$.

Теорема 28.1.8. *Пусть задана комплексная $n \times n$ матрица A. Тогда существует невырожденная $n \times n$ матрица S такая, что*

$$A = SJS^{-1}, \tag{28.27}$$

где J – жорданова матрица (28.26). Матрица A определяет подобную ей жорданову матрицу однозначно с точностью до перестановки жордановых клеток на ее главной диагонали. Собственные значения λ_i не обязательно различны. Если матрица A вещественна и обладает только вещественными собственными значениями, то подобие может быть реализовано с помощью вещественной матрицы S.

Число k жордановых клеток матрицы A (с учетом возможных повторов) равно максимальному числу ее линейно независимых векторов. При этом для каждой жордановой клетки существует только один независимый собственный вектор. Матрица A диагонализируема тогда и только тогда, когда $k = n$. В частности, если все собственные числа различны, то матрицу A можно привести к диагональному виду с помощью преобразования подобия.

Если порядок всех жордановых клеток равен единице, $n_i = 1$, $i = 1, \ldots, k$, то жорданова матрица является диагональной. Если один из жордановых блоков имеет бóльший порядок, $n_i > 1$, то жорданова матрица не может быть приведена к диагональному виду преобразованием подобия.

Предложение 28.1.1. *Любая комплексная матрица A подобна своей транспонированной A^T.*

Любая жорданова клетка записывается в виде суммы единичной и нильпотентной матриц:

$$J_k(\lambda) = \lambda \mathbb{1} + N_k, \qquad (N_k)^k = 0. \tag{28.28}$$

Любая жорданова матрица записывается в виде суммы диагональной и нильпотентной матриц:

$$J = D + N, \qquad N^k = 0, \tag{28.29}$$

где k – порядок наибольшей жордановой клетки. Если задана произвольная матрица и ее жорданова форма, то

$$A = SJS^{-1} = SDS^{-1} + SNS^{-1} := A_D + A_N,$$

где A_D диагонализируемая и A_N нильпотентная матрицы. При этом $[A_D, A_N] = 0$.

Приведем некоторые результаты для эрмитовых и симметричных матриц.

Определение. Комплексная $n \times n$ матрица A называется *эрмитовой*, если выполнено равенство $A^\dagger = A$, где символ $\dagger$ обозначает эрмитово сопряжение: $A^\dagger := (\bar{A})^T$. Если выполнено равенство $A^\dagger = -A$, то матрица A называется *антиэрмитовой*. □

Если матрица A вещественна, то эрмитовы и антиэрмитовы матрицы соответствуют симметричным и антисимметричным матрицам.

Теорема 28.1.9. *Все собственные значения эрмитовой $n \times n$ матрицы A вещественны. Матрица A эрмитова тогда и только тогда, когда существует унитарная матрица U такая, что выполнено равенство*

$$A = U\Lambda U^{-1}, \qquad U \in \mathbb{U}(n), \tag{28.30}$$

где Λ – вещественная диагональная матрица. Матрица Λ составлена из собственных значений матрицы A. Унитарную матрицу U можно выбрать из односвязной подгруппы $\mathbb{SU}(n) \subset \mathbb{U}(n)$.

Все собственные значения вещественной симметричной $n \times n$ матрицы B вещественны. Матрица B симметрична тогда и только тогда, когда существует ортогональная матрица S такая, что выполнено равенство

$$B = S\Lambda S^{-1}, \qquad S \in \mathbb{O}(n), \tag{28.31}$$

где Λ – вещественная диагональная матрица. Матрица Λ составлена из собственных значений матрицы B. Ортогональную матрицу S можно выбрать из связной компоненты единицы $\mathbb{SO}(n)$.

С помощью одной унитарной матрицы иногда удается диагонализировать сразу несколько эрмитовых матриц.

Теорема 28.1.10. *Пусть $\mathcal{F}$ – семейство эрмитовых матриц. Унитарная матрица U такая, что матрица UAU^{-1} – диагональна для всех $A \in \mathcal{F}$, существует тогда и только тогда, когда $[A, B] = 0$ для всех $A, B \in \mathcal{F}$.*

В заключение рассмотрим некоторые свойства суперматриц, элементами которых являются коммутирующие и антикоммутирующие объекты. Пусть задана квадратная суперматрица

$$M_M{}^N = \begin{pmatrix} A_m{}^n & B_m{}^\nu \\ C_\mu{}^n & D_\mu{}^\nu \end{pmatrix}, \tag{28.32}$$

где A и D – квадратные невырожденные матрицы с коммутирующими элементами, а B и C – матрицы с антикоммутирующими элементами, которые могут быть прямоугольными. Суперслед и супероопределитель матрицы M определены соотношениями:

$$\mathsf{str}\, M := \mathsf{tr}\, A - \mathsf{tr}\, D, \tag{28.33}$$

$$\mathsf{sdet}\, M := \frac{\mathsf{det}\, A}{\mathsf{det}\,(D - CA^{-1}B)} = \frac{\mathsf{det}\,(A - BD^{-1}C)}{\mathsf{det}\, D}. \tag{28.34}$$

Из определения вытекает, что ряд свойств обычного следа и определителя наследуется суперматрицами:

$$\begin{aligned} \mathsf{str}\,(M+N) &= \mathsf{str}\,M + \mathsf{str}\,N, \\ \mathsf{str}\,(MN) &= \mathsf{str}\,(NM), \\ \mathsf{sdet}\,M^{-1} &= \frac{1}{\mathsf{sdet}\,M}, \\ \mathsf{sdet}\,(MN) &= \mathsf{sdet}\,M\ \mathsf{sdet}\,N, \\ \mathsf{sdet}\,M &= \mathsf{exp}\ \mathsf{str}\ \mathsf{ln} M. \end{aligned} \tag{28.35}$$

28.2 Матрицы Паули

Определение. Три эрмитовых 2×2 матрицы $\boldsymbol{\sigma} = \{\sigma_i\}$, $i = 1, 2, 3$:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{28.36}$$

называются *матрицами Паули.* □

Они обладают следующими свойствами:

$$\begin{aligned} \sigma_1\sigma_2 &= -\sigma_2\sigma_1 = i\sigma_3, \\ \sigma_2\sigma_3 &= -\sigma_3\sigma_2 = i\sigma_1, \\ \sigma_3\sigma_1 &= -\sigma_1\sigma_3 = i\sigma_2, \\ \sigma_i^2 &= 1 \quad \text{(суммирования по } i \text{ нет)}, \\ \sigma_1\sigma_2\sigma_3 &= i. \end{aligned} \tag{28.37}$$

Свойства (28.37) можно записать в виде коммутатора и антикоммутатора матриц Паули:

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ij}{}^k\sigma_k, \qquad \{\sigma_i\sigma_j\} = 2\delta_{ij} \tag{28.38}$$

или

$$\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ij}{}^k\sigma_k, \tag{28.39}$$

где ϵ_{ijk} – полностью антисимметричный тензор третьего ранга и подъем индексов осуществляется с помощью евклидовой метрики δ_{ij}.

Матрицы Паули встречаются в различных контекстах. Во-первых, они являются генераторами группы двумерных унитарных матриц $\mathbb{SU}(2)$. Во-вторых, второе из соотношений (28.38) означает, что матрицы Паули являются образующими алгебры Клиффорда $\mathbb{CL}(3)$. В-третьих, матрицы Паули, умноженные на $-i$ (28.47), можно рассматривать, как представление образующих алгебры кватернионов.

Для следов произведения матриц Паули справедливы тождества:

$$\mathsf{tr}\,\sigma_i = 0, \qquad \mathsf{tr}\,(\sigma_i\sigma_j) = 2\delta_{ij}, \qquad \mathsf{tr}\,(\sigma_i\sigma_j\sigma_k) = 2i\varepsilon_{ijk}. \tag{28.40}$$

Рассмотрим вектор $\boldsymbol{x} = \{x^i\}$ в трехмерном евклидовом пространстве с обычным скалярным произведением. Тогда справедлива формула

$$e^{i(\boldsymbol{x}\boldsymbol{\sigma})} = \mathsf{cos}\,|\boldsymbol{x}| + i\frac{(\boldsymbol{x}\boldsymbol{\sigma})}{|\boldsymbol{x}|}\,\mathsf{sin}\,|\boldsymbol{x}|, \tag{28.41}$$

где

$$(\boldsymbol{x}\boldsymbol{\sigma}) = x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3, \qquad |\boldsymbol{x}| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}.$$

Свертка равенства (28.39) с произвольными векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ приводит к тождеству

$$(\boldsymbol{x}\boldsymbol{\sigma})(\boldsymbol{y}\boldsymbol{\sigma}) = (\boldsymbol{x}\boldsymbol{y}) + i([\boldsymbol{x}\boldsymbol{y}]\boldsymbol{\sigma}), \tag{28.42}$$

где $[\boldsymbol{x}\boldsymbol{y}]^i = x^j y^k \varepsilon_{ik}{}^i$ – векторное произведение векторов. Из свойства (28.39) следует, что всякая функция от матриц Паули сводится к линейной функции.

С групповой точки зрения матрицы Паули являются генераторами группы $\mathbb{SU}(2)$. Скалярное произведение $(\boldsymbol{x}\boldsymbol{\sigma})$ и экспонента $e^{i(\boldsymbol{x}\boldsymbol{\sigma})}$ являются соответственно элементами алгебры Ли $\mathfrak{su}(2)$ и группы Ли $\mathbb{SO}(2)$. Отображение (28.41) представляет собой экспоненциальное отображение алгебры Ли в группу Ли.

28.3 Кватернионы

Определение. Рассмотрим четырехмерное векторное пространство $\mathbb{H}$ над полем вещественных чисел, которое содержит поле вещественных чисел как подпространство. Обозначим базис этого пространства символами $1, i, j, k$, где 1 – единица из поля вещественных чисел. Тогда произвольный элемент $q \in \mathbb{H}$ представим в виде

$$q = a + bi + cj + dk, \tag{28.43}$$

где a, b, c, d – вещественные числа. Введем билинейное умножение в $\mathbb{H}$, положив следующее правило умножения базисных векторов

$$\begin{aligned} ij &= -ji = k, \\ jk &= -kj = i, \\ ki &= -ik = j, \\ i^2 &= j^2 = k^2 = -1. \end{aligned} \tag{28.44}$$

Элементы вида (28.43) с операциями сложения, умножения на действительные числа и умножения (28.44) называются *кватернионами.* □

С помощью прямых вычислений можно проверить, что кватернионы с так определенным умножением образуют ассоциативную, но не коммутативную алгебру $\mathbb{H}$ с единицей над полем вещественных чисел. Эту алгебру можно представить в виде комплексных матриц $A(q) \in \mathbb{GL}(2, \mathbb{C})$

$$A(q) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}, \tag{28.45}$$

где i – мнимая единица. Прямые вычисления показывают, что отображение $q \mapsto A(q)$ является гомоморфизмом, то есть

$$A(q_1 q_2) = A(q_1)A(q_2).$$

Для доказательства этого утверждения достаточно проверить его для $q = i, j, k$. В частности, при $q = i, j, k$ имеем

$$A(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad A(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad A(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{28.46}$$

Эти матрицы с точностью до множителя $-i$ совпадают с матрицами Паули (28.36)

$$\sigma_1 = -iA(k), \quad \sigma_2 = -iA(j), \quad \sigma_3 = -iA(i). \tag{28.47}$$

Введем операцию сопряжения в $\mathbb{H}$, то есть поставим каждому кватерниону вида (28.43) его сопряженный элемент

$$\bar{q} = a - bi - cj - dk. \tag{28.48}$$

Это отображение $q \to \bar{q}$ является антиавтоморфизмом алгебры кватернионов, то есть

$$\overline{q_1 + q_2} = \bar{q}_1 + \bar{q}_2, \quad \overline{q_1 q_2} = \bar{q}_2 \bar{q}_1.$$

Число a называется вещественной частью $\mathsf{re}\, q = a$, а остаток $\mathsf{im}\, q = bi + cj + dk$ – мнимой частью кватерниона. В алгебре кватернионов можно ввести симметричное скалярное произведение

$$(q_1, q_2) = \mathsf{re}\, q_1 \bar{q}_2 = a_1 a_2 + b_1 b_2 + c_1 c_2 + d_1 d_2, \tag{28.49}$$

которое совпадает с обычным евклидовым произведением в $\mathbb{R}^4$. Поскольку каждый кватернион представим в виде $q = \mathsf{re}\, q + \mathsf{im}\, q$, то алгебра кватернионов распадается в прямую сумму ортогональных подпространств $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}^3$, где $\mathbb{R}$ – прямая, состоящая из вещественных кватернионов, а $\mathbb{R}^3$ – ортогональное трехмерное евклидово пространство мнимых кватернионов.

Определим норму кватерниона

$$\| q \|^2 = q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2. \tag{28.50}$$

Нетрудно проверить, что

$$\| q \|^2 = \mathsf{det}\, A(q).$$

Отсюда следует, что $\| q_1 q_2 \|^2 = \| q_1 \|^2 \| q_2 \|^2$.

Каждый отличный от нуля кватернион обладает единственным обратным элементом

$$qq^{-1} = q^{-1}q = 1,$$

где

$$q^{-1} = \frac{\bar{q}}{\| q \|^2}.$$

Это означает, что множество кватернионов относительно операции сложения и умножения (28.44) (без умножения на действительные числа) образует тело, так как умножение некоммутативно.

Алгебра кватернионов $\mathbb{H}$ допускает естественное отождествление с двумерным комплексным пространством $\mathbb{C}^2$. Используя определение умножения в алгебре кватернионов $k = -ji$ (28.44), запишем кватернион в виде

$$q = (a + ib) + j(c - id) = z_1 + j\bar{z}_2,$$

где линейные комбинации $z_1 = a + ib$ и $z_2 = c + id$ можно отождествить с парой комплексных чисел. При этом 1 и j рассматриваются, как базис в $\mathbb{C}^2$.

Рассмотрим мнимый кватернион, не содержащий вещественной части,

$$r = bi + cj + dk, \quad \| r \| = b^2 + c^2 + d^2.$$

Его можно представить, как вектор в трехмерном евклидовом пространстве $r \in \mathbb{R}^3$. Пусть $q \in \mathbb{H}$ произвольный обратимый (в смысле кватернионного умножения) кватернион. Тогда кватернион qrq^{-1} снова является чистым кватернионом с той же нормой

$$qrq^{-1} \in \mathbb{R}^3, \quad \| qrq^{-1} \| = \| r \| .$$

Это значит, что каждому обратимому кватерниону, т.е. отличному от нуля, ставится в соответствие некоторое вращение. При этом двум кватернионам q и $-q$, отличающимся знаком, соответствует одно и то же вращение. Можно проверить, что любое вращение можно реализовать с помощью кватерниона единичной нормы. Поскольку кватернионы единичной нормы параметризуют трехмерную сферу единичного радиуса, то существует гомеоморфизм

$$\mathbb{SO}(3) \approx \frac{\mathbb{S}^3}{\mathbb{Z}_2}.$$

Рассмотрим отображение кватернионов $q \in \mathbb{H}$, задаваемое двумя кватернионами $a, b \in \mathbb{H}$ единичной нормы

$$q \mapsto aqb^{-1}, \quad \| a \| = \| b \| = 1.$$

Это отображение задает вращение четырехмерного евклидова пространства $\mathbb{H} = \mathbb{R}^4$, поскольку сохраняет норму q. Можно проверить, что любое вращение определяет пару a и b с точностью до знака, так как кватернионы a, b и $-a, -b$ определяют одно и то же вращение. Это значит, что существует гомеоморфизм

$$\mathbb{SO}(4) \approx \frac{\mathbb{S}^3 \times \mathbb{S}^3}{\mathbb{Z}_2}.$$

28.4 Полностью антисимметричные тензоры

Полностью антисимметричный тензор ранга n в n-мерном евклидовом пространстве единственен с точностью до умножения на произвольную ненулевую постоянную и имеет всего одну независимую компоненту. Его всегда можно представить в виде константы, умноженной на тензор, все ненулевые компоненты которого равны по модулю единице. Этот тензор часто используется в приложениях и имеет специальное обозначение

$$\begin{aligned} \varepsilon^{a_1 \ldots a_n} &= \varepsilon^{[a_1 \ldots a_n]}, \quad & \varepsilon^{12\ldots n} &= 1, \\ \varepsilon_{a_1 \ldots a_n} &= \varepsilon^{a_1 \ldots a_n}, \quad & \varepsilon_{12\ldots n} &= 1. \end{aligned} \tag{28.51}$$

Здесь подъем индексов произведен с помощью евклидовой метрики. Произведения и свертки двух таких тензоров выражаются через символы Кронекера

$$
\begin{aligned}
\varepsilon^{a_1\dots a_n}\varepsilon_{b_1\dots b_n} &= n!\;\delta^{[a_1\dots a_n]}_{b_1\dots b_n} &&= \mathsf{det}\begin{pmatrix} \delta^{a_1}_{b_1} & \delta^{a_2}_{b_1} & \cdots & \delta^{a_n}_{b_1} \\ \delta^{a_1}_{b_2} & \delta^{a_2}_{b_2} & \cdots & \delta^{a_n}_{b_2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta^{a_1}_{b_n} & \delta^{a_2}_{b_n} & \cdots & \delta^{a_n}_{b_n} \end{pmatrix} \\
\varepsilon^{b_1 a_2\dots a_n}\varepsilon_{b_1\dots b_n} &= (n-1)!\;\delta^{[a_2\dots a_n]}_{b_2\dots b_n} &&= \frac{n!}{C_n^1}\delta^{[a_2\dots a_n]}_{b_2\dots b_n} \\
\varepsilon^{b_1 b_2 a_3\dots a_n}\varepsilon_{b_1\dots b_n} &= 2(n-2)!\;\delta^{[a_3\dots a_n]}_{b_3\dots b_n} &&= \frac{n!}{C_n^2}\delta^{[a_3\dots a_n]}_{b_3\dots b_n} \\
\varepsilon^{b_1\dots b_k a_{k+1}\dots a_n}\varepsilon_{b_1\dots b_n} &= k!(n-k)!\;\delta^{[a_{k+1}\dots a_n]}_{b_{k+1}\dots b_n} &&= \frac{n!}{C_n^k}\delta^{[a_{k+1}\dots a_n]}_{b_{k+1}\dots b_n} \\
\varepsilon^{b_1\dots b_{n-1} a_n}\varepsilon_{b_1\dots b_n} &= (n-1)!\;\delta^{a_n}_{b_n} &&= \frac{n!}{C_n^{n-1}}\delta^{a_n}_{b_n} \\
\varepsilon^{b_1\dots b_n}\varepsilon_{b_1\dots b_n} &= n! &&
\end{aligned}
\tag{28.52}
$$

Полностью антисимметричный тензор максимального ранга в пространстве с метрикой произвольной сигнатуры определяется соотношениями

$$
\varepsilon_{12\dots n} = 1, \qquad \varepsilon^{12\dots n}\,\mathsf{sgn}\,, \tag{28.53}
$$

где множитель $\mathsf{sgn} = \mathsf{sign}\,(\,\mathsf{det}\, g_{\alpha\beta})$ равен знаку определителя метрики. Этот множитель возникает при подъеме индексов. Для таких пространств правые части (28.52) необходимо умножить на множитель sgn.

Определитель произвольной матрицы $A_a{}^b$ в евклидовом пространстве можно записать в виде

$$
\begin{aligned}
\mathsf{det}\,(A_a{}^b) &= \varepsilon_{b_1 b_2\dots b_n} A_1{}^{b_1} A_2{}^{b_2}\dots A_n{}^{b_n} = \varepsilon^{a_1 a_2\dots a_n} A_{a_1}{}^1 A_{a_2}{}^2\dots A_{a_n}{}^n \\
&= \frac{1}{n!}\varepsilon^{a_1 a_2\dots a_n}\varepsilon_{b_1 b_2\dots b_n} A_{a_1}{}^{b_1} A_{a_2}{}^{b_2}\dots A_{a_n}{}^{b_n}.
\end{aligned}
\tag{28.54}
$$

Полностью антисимметричный тензор в n-мерном евклидовом пространстве инвариантен относительно $\mathbb{SO}(n)$ вращений с матрицей $S_a{}^{a'} \in \mathbb{SO}(n)$:

$$
\varepsilon^{a'_1\dots a'_n} = \varepsilon^{a_1\dots a_n} S_{a_1}{}^{a'_1}\dots S_{a_n}{}^{a'_n} = \mathsf{det}\,(S_a{}^{a'})\varepsilon^{a'_1\dots a'_n}. \tag{28.55}
$$

Полностью антисимметричный тензор в голономном базисе имеет следующие компоненты:

$$
\varepsilon^{\alpha_1\dots\alpha_n} = e^{\alpha_1}{}_{a_1}\dots e^{\alpha_n}{}_{a_n}\varepsilon^{a_1\dots a_n}, \tag{28.56}
$$

и в геометрии Римана–Картана является ковариантно постоянным:

$$
\nabla_\beta \varepsilon^{\alpha_1\dots\alpha_n} = 0. \tag{28.57}
$$

В аффинной геометрии ???

По-определению полностью антисимметричный тензор в пространстве Минковского четного числа измерений меняет знак при пространственном отражении $(x^0, x^i) \to (x^0, -x^i)$, то есть является псевдотензором. Приставку псевдо- мы, как правило, опускаем.

Полностью антисимметричная тензорная плотность $\hat{e}^{\alpha_1\ldots\alpha_n}$ веса -1 определяется соотношениями

$$\begin{aligned}\hat{\varepsilon}^{\alpha_1\ldots\alpha_n} &= \sqrt{|g|}\varepsilon^{\alpha_1\ldots\alpha_n},\\ \hat{\varepsilon}_{\alpha_1\ldots\alpha_n} &= \frac{1}{\sqrt{|g|}}\varepsilon_{\alpha_1\ldots\alpha_n}.\end{aligned} \tag{28.58}$$

Ее ненулевые компоненты равны по модулю единице в любой системе координат,

$$\hat{\varepsilon}^{\alpha_1\ldots\alpha_n} = \begin{cases} +1 & \text{– четная перестановка индексов,}\\ -1 & \text{– нечетная перестановка индексов,}\\ 0 & \text{– совпадение любой пары индексов.}\end{cases} \tag{28.59}$$

Отсюда следует, что все частные производные антисимметричной тензорной плотности равны нулю:

$$\partial_\beta\hat{\varepsilon}^{\alpha_1\ldots\alpha_n} = 0. \tag{28.60}$$

Приведем явные формулы для трехмерного евклидова пространства $\mathbb{R}^3$, которые часто используются в физических приложениях,

$$\varepsilon_{123} = 1, \qquad \varepsilon^{123} = 1. \tag{28.61}$$

Свертки:

$$\begin{aligned}\varepsilon^{lmn}\varepsilon_{ijk} &= \delta^{lmn}_{ijk} + \delta^{mnl}_{ijk} + \delta^{nlm}_{ijk} - \delta^{lnm}_{ijk} - \delta^{mln}_{ijk} - \delta^{nml}_{ijk} = \det\begin{pmatrix}\delta^l_i & \delta^m_i & \delta^n_i\\ \delta^l_j & \delta^m_j & \delta^n_j\\ \delta^l_k & \delta^m_k & \delta^n_k\end{pmatrix}\\ \varepsilon^{klm}\varepsilon_{ijm} &= \delta^{kl}_{ij} - \delta^{lk}_{ij},\\ \varepsilon^{jkl}\varepsilon_{ikl} &= 2\delta^j_i,\\ \varepsilon^{ijk}\varepsilon_{ijk} &= 6.\end{aligned} \tag{28.62}$$

Полностью антисимметричный тензор второго ранга в евклидовом пространстве имеет две ненулевые компоненты

$$\varepsilon_{12} = 1, \qquad \varepsilon^{12} = 1 \tag{28.63}$$

или

$$\varepsilon_{ab} = \varepsilon^{ab} = \varepsilon^a{}_b = \varepsilon_a{}^b = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}. \tag{28.64}$$

Свертки:

$$\begin{aligned}\varepsilon^{ab}\varepsilon_{cd} &= \delta^{ab}_{cd} - \delta^{ba}_{cd} = \det\begin{pmatrix}\delta^a_c & \delta^b_c\\ \delta^a_d & \delta^b_d\end{pmatrix}\\ \varepsilon^{ab}\varepsilon_{ad} &= \delta^b_d,\\ \varepsilon^{ab}\varepsilon_{ab} &= 2.\end{aligned} \tag{28.65}$$

В пространстве Минковского нечетного числа измерений формулы для сверток полностью антисимметричных тензоров те же, что и в евклидовом пространстве. В пространстве Минковского четного числа измерений полностью антисимметричные

тензоры с верхними и нижними индексами отличаются знаком. Определим антисимметричный тензор четвертого ранга в пространстве Минковского следующим образом

$$\varepsilon_{0123} = 1, \qquad \varepsilon^{0123} = -1. \tag{28.66}$$

Свертки:

$$\begin{aligned}
\varepsilon^{abcd}\varepsilon_{efgh} =& -\delta^{abcd}_{efgh} - \delta^{acdb}_{efgh} - \delta^{adbc}_{efgh} + \delta^{acbd}_{efgh} + \delta^{abdc}_{efgh} + \delta^{adcb}_{efgh} \\
&+ \delta^{bcda}_{efgh} + \delta^{bdac}_{efgh} + \delta^{bacd}_{efgh} - \delta^{bdca}_{efgh} - \delta^{bcad}_{efgh} - \delta^{badc}_{efgh} \\
&- \delta^{cdab}_{efgh} - \delta^{cabd}_{efgh} - \delta^{cbda}_{efgh} + \delta^{cadb}_{efgh} + \delta^{cdba}_{efgh} + \delta^{cbad}_{efgh} \\
&+ \delta^{dabc}_{efgh} + \delta^{dbca}_{efgh} + \delta^{dcab}_{efgh} - \delta^{dbac}_{efgh} - \delta^{dacb}_{efgh} - \delta^{dcba}_{efgh} \\
\varepsilon^{abcd}\varepsilon_{afgh} =& -\delta^{bcd}_{fgh} - \delta^{cdb}_{fgh} - \delta^{dbc}_{fgh} + \delta^{cbd}_{fgh} + \delta^{bdc}_{fgh} + \delta^{dcb}_{fgh} \\
\varepsilon^{abcd}\varepsilon_{abgh} =& -2(\delta^{cd}_{gh} - \delta^{dc}_{gh}), \\
\varepsilon^{abcd}\varepsilon_{abch} =& -6\delta^{d}_{h}, \\
\varepsilon^{abcd}\varepsilon_{abcd} =& -24.
\end{aligned} \tag{28.67}$$

Полностью антисимметричные тензоры в трехмерном евклидовом пространстве и пространстве Минковского связаны соотношениями

$$\varepsilon_{ijk} = \varepsilon_{0ijk}, \qquad \varepsilon^{ijk} = -\varepsilon^{0ijk}. \tag{28.68}$$

Полностью антисимметричный тензор второго ранга в пространстве Минковского $\mathbb{R}^{1,1}$ имеет две ненулевые компоненты

$$\varepsilon_{01} = 1, \qquad \varepsilon^{01} = -1 \tag{28.69}$$

или

$$\varepsilon_{ab} = -\varepsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \varepsilon^{a}{}_{b} = -\varepsilon_{a}{}^{b} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{28.70}$$

Свертки:

$$\begin{aligned}
\varepsilon^{ab}\varepsilon_{cd} &= -\delta^{ab}_{cd} + \delta^{ba}_{cd}, \\
\varepsilon^{ab}\varepsilon_{ad} &= -\delta^{b}_{d}, \\
\varepsilon^{ab}\varepsilon_{ab} &= -2.
\end{aligned} \tag{28.71}$$

В двумерном пространстве Минковского $\mathbb{R}^{1,1}$ можно ввести проекционные операторы

$$\Pi^{\pm b}_{a} = \frac{1}{2}(\delta^{b}_{a} \pm \varepsilon_{a}{}^{b}).$$

Отметим, что аналогичная конструкция в евклидовом пространстве не приводит к проекционным операторам.

Пусть в римановом пространстве $(\mathbb{M}, g)$, $\dim \mathbb{M} = n$, задано тензорное поле с компонентами $K_{\alpha\beta}$ типа $(0,2)$ такое, что $\det K_{\alpha\beta} \neq 0$. Тогда существует обратное тензорное поле $K^{\alpha\beta}$ типа $(2,0)$,

$$K^{\alpha\beta}K_{\beta\gamma} = \delta^{\alpha}_{\gamma},$$

и его можно представить в виде

$$K^{\alpha\beta} = \frac{1}{|\det g|(n-1)!}\hat{\varepsilon}^{\alpha\gamma_2\ldots\gamma_n}\hat{\varepsilon}^{\beta\delta_2\ldots\delta_n}M_{\gamma_2\delta_2}\ldots M_{\gamma_n\delta_n}, \tag{28.72}$$

Если метрика имеет лоренцеву сигнатуру, то в пространстве четного числа измерений в правой части появляется знак минус.