Справочный материал.
Нормированные пространства.

Векторное пространство (над полем k; у нас чаще $k = \mathbb{R}$ и иногда $k = \mathbb{C}$) — это множество E, в котором введены две операции:

— "сложение" $E \times E \to E \qquad (x, y) \mapsto x + y$;

— "умножение на скаляры"[1] $k \times E \mapsto E \qquad (\lambda.x) \mapsto \lambda x$,

причём

1) относительно сложения E является абелевой группой;

2) умножение на скаляры обладает свойствами

а) умножение на единицу — тождественное преобразование: $1x = x$ (подразумевается — для всех $x \in E$; в дальнейшем я обычно считаю подобные уточнения само собою разумеющимися);

б) дистрибутивность: $\lambda(x + y) = \lambda x + \lambda y, \quad (\lambda + \mu)x = \lambda x + \mu y$;

в) ассоциативность: $(\lambda\mu)x = \lambda(\mu x)$.

Отображение $A : E \to F$ одного векторного пространства в другое называется линейным, если оно "сохраняет обе эти операции", т.е. если $A(x + y) = Ax + Ay$ и $A(\lambda x) = \lambda Ax$.

Размерность $\dim E$ векторного пространства E — это максимальное число линейно независимых элементов в E. ($x_1, \ldots, x_n$ называются линейно независимыми, если

$$\sum \lambda_i x_i = 0 \Longleftrightarrow \text{все } x_i = 0.)$$

Если такого максимального числа нет, т.е. если $\forall n \ \exists \geqslant n$ линейно независимых элементов, то говорят, что $\dim E = \infty$.

Далее $k = \mathbb{R}$ или $\mathbb{C}$.

Нормированное пространство — это векторное пространство E, снабжённое нормой, т.е. функцией[2]

$$E \to \mathbb{R}_+ \qquad x \mapsto |x|$$

---

[1]Скаляры — это элементы поля k, т.е. (при нашем k) попросту числа. Слово "скаляр" употребляют обычно тогда, когда где-то рядом фигурируют векторы (т.е. в нашем случае, элементы E) — скаляры как бы противопоставляются векторам. Вначале во избежание путаницы я буду обозначать скаляры греческими, а векторы — латинскими буквами.

[2]$\mathbb{R}_+ = \{\lambda;\ \lambda \in \mathbb{R}, \lambda \geqslant 0\}$. Почему значок "+" употребляется при условии "$\geqslant$", официально объясняется принятым в группе Бурбаки изменением смысла слов "больше", "положительно" и т.д. (что в свою очередь мотивируется какими-то словами), но я подозреваю, что на самом деле всё это сделано затем, чтобы шпион не догадался.

со свойствами:

$$|x| = 0 \Longleftrightarrow x = 0;$$
$$|\lambda x| = |\lambda||x|;$$
$$|x + y| \leqslant |x| + |y| \qquad \text{(неравенство треугольника).}$$

Когда хотят явно упомянуть об используемой норме, то говорят: “нормированное пространство $(E, |\,.\,|)$”, но чаще говорят короче о “нормированном пространстве E”.

Нормы $|\,.\,|$ и $\|\,.\,\|$ называются эквивалентными, если $\exists C_1, C_2 > 0\ \forall x \in E$ $C_1|x| \leqslant \|x\| \leqslant C_1|x|$.

Упражнение. Формально это определение выглядит несимметричным оносительно $|\,.\,|$ и $\|\,.\,\|$, но на самом деле эквивалентность норм является настоящим отношением эквивалентности, т.е. (обозначая его на минуту через $\sim$)

$$|\,.\,| \sim |\,.\,|; \quad |\,.\,| \sim \|\,.\,\| \Longleftrightarrow \|\,.\,\| \sim |\,.\,|;$$
$$\text{если } |\,.\,| \sim \|\,.\,\| \text{ и } \|\,.\,\| \sim |||\,.\,|||, \text{ то } |\,.\,| \sim |||\,.\,|||.$$

В $\mathbb{R}^n$ или $\mathbb{C}^n$ (элементы этих пространств суть конечные наборы $(x_1, \ldots, x_n)$, где все $x_i \in \mathbb{R}$ или все $x_i \in \mathbb{C}$) часто употребляют нормы

$$|x|_1 = \sum |x_i|, \qquad |x|_2 = \sqrt{\sum |x_i|^2}, \qquad |x|_\infty = \max_i |x_i|.$$

Эти обозначения связаны с более общим обозначением

$$|x|_p = \Big(\sum |x_i|^p\Big)^{\frac{1}{p}} \qquad (p \geqslant 1)$$

Сразу даже не очевидно, что $|x|_p$ — норма (неравенство треугольника для неё фактически эквивалентно неравенству Минковского). $|x|_1$ и $|x|_2$ — частные случаи $|x|_p$, а $|x|_\infty = \lim_{p\to\infty} |x|_p$. Действительно, хоть одно из слагаемых в выражении для $|x|_p^p$ равно $\max |x_i|^p = |x|_\infty^p$, поэтому $|x|_\infty^p \leqslant \sum |x_i|^p$, $|x|_\infty \leqslant |x|_p$. С другой стороны, в той же сумме каждое слагаемое $\leqslant |x|_\infty^p$, поэтому $|x|_p^p \leqslant n|x|_\infty^p$,

$$|x|_p \leqslant n^{\frac{1}{p}}|x|_\infty \to |x|_\infty \quad \text{при } p \to \infty.$$

Упражнение. Докажите, что нормы $|x|_1, |x|_2$ и $|x|_\infty$ эквивалентны друг другу (равно как и все нормы $|x|_p$).

Оказывается, вообще в конечномерном векторном пространстве (над $\mathbb{R}$ или $\mathbb{C}$) все нормы эквивалентны друг другу. В бесконечномерном векторном пространстве это не так.

Сходимость: говорят, что последовательность x_n элементов из E сходится или стремится к x_0 и что x_0 является её пределом (запись — как обычно: $x_n \to x_0, x_0 = \lim x_n$), если $|x_n - x_0| \to 0$. Аналогично, для функции $f(t)$ от вещественного аргумента t, определённой возле $t = t_0$ и принимающей значения в E, говорят, что $f(t)$ стремится к $a \in E$ и что a является её пределом (запись:

$$f(t) \to a \quad \text{при } t \to t_0, \quad a = \lim f(t)),$$

если $|f(t) - a| \to 0$ при $t \to t_0$. Очевидная перефразировка относится сходимости и пределу при $t \to \infty$. Без изменений по сравнению с числовым случаем получаются основные свойства пределов — единственность предела, предел суммы равен сумме пределов, предел произведения $\lambda_n x_n$ равен произведению $(\lim \lambda_n)(\lim x_n)$, — и аналогично для функций.

Упражнение. Докажите, что если нормы $|\,.\,|$ и $\|\,.\,\|$, определённые в векторном пространстве E, эквивалентны, то в нормированных пространствах $(E, |\,.\,|)$ и $(E, \|\,.\,\|)$ одни и те же последовательности являются сходящимися или нет, при этом предел сходящейся последовательности в $(E, |\,.\,|)$ совпадает с её пределом в $(E, \|\,.\,\|)$.

Последовательность x_n называется фундаментальной, если

$$\forall \varepsilon > 0 \; \exists N = N(\varepsilon) \; \forall n, m > N \; |x_n - x_m| < \varepsilon.$$

(Можно ввести аналогичное понятие и для функций от числового аргумента.) Если последовательность x_n сходится, то она — фундаментальная.

Известно, что фундаментальная числовая последовательность всегда сходится. Здесь подразумевается, что числа — это числа из $\mathbb{R}$ (или $\mathbb{C}$). Если же мы пожелали бы иметь дело только с рациональными числами (множество таковых обозначается через $\mathbb{Q}$), то убедились бы, что фундаментальная последовательность может и не сходиться. Она, конечно, сходится к некоторому числу из $\mathbb{R}$, но оно может быть иррациональным, и тогда в $\mathbb{Q}$ у неё предела нет.

Нормированное пространство E называется полным, если в нём всякая фундаментальная последовательность сходится.

Упражнение. Докажите, что если нормы $|\,.\,|$ и $\|\,.\,\|$, определённые в векторном пространстве E, эквивалентны, то нормированные пространства $(E, |\,.\,|)$ и $(E, \|\,.\,\|)$ одновременно являются полными или неполными. При этом в них фундаментальными являются одни и те же последовательности.

$\mathbb{R}^n$ и $\mathbb{C}^n$ являются полными при использовании любой нормы. Сходимость последовательности x_k, элементы которой являются наборами n чисел: $x_k = (x_k^1, \ldots, x_k^n)$, равносильна сходимости n числовых последовательностей $x_k^1, \ldots, x_k^n$ и $\lim x_k = (\lim x_k^1, \ldots, \lim x_k^n)$. Сходимость по этой норме (не только многочленов) — это равномерная сходимость на $[0, 1]$. Известна теорема Вейерштрасса: для всякой функции, заданной и непрерывной на каком-нибудь отрезке $[a, b]$, имеется равномерно сходящаяся к ней на $[a, b]$ последовательность многочленов. Это можно перефразировать так. Введём пространство непрерывных функций

$$C[a, b] = \{\text{непрерывные функции } [a, b] \to \mathbb{R}\}, \qquad \|f\| = \sup_{[a,b]} |f(x)|.$$

Тогда E плотно в $C[a, b]$ (т.е. сколь угодно близко к любому элементу $C[a, b]$ имеется элемент из E). Это аналогично плотности $\mathbb{Q}$ в $\mathbb{R}$, и неполнота E аналогична неполноте $\mathbb{Q}$.

Полные нормированные пространства называются банаховыми.

Как известно, прямым (или декартовым) произведением $A \times B$ множеств A и B называют множество всевозможных пар (x, y), где $x \in A$, $y \in B$ (мы уже пользовались этими понятием и обозначением). Прямое произведение $E \times F$ двух векторных пространств E, F естественным образом наделяется труктурой векторного пространства, т.е. в нём вводятся операции сложения и умножения на скаляры, имеющие надлежащие свойства. Эти операции вводятся "покоординатно": принимают, что

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad \lambda(x, y) = (\lambda x, \lambda y).$$

Проверка того, что введённые таким образом операции действительно удовлетворяют условиям, фигурирующим в определении векторного пространства, является автоматической. (Читателю следует её провести, если он никогда не делал этого раньше.) Когда $E \times F$ рассматривается как векторное пространство с только что описанными операциями, на это указывают посредством специального названия и специального обозначения для этого векторного пространства — его называют прямой суммой векторных пространств E, F и обозначают через $E \oplus F$. Если E, F — не только векторные, но и нормированные пространства, то их прямую сумму $E \oplus F$ тоже можно снабдить нормой. Используют одну из эквивалентных друг другу норм

$$|(x,y)|_1 = |x| + |y|, \quad |(x,y)|_2 = \sqrt{|x|^2 + |y|^2}, \quad |(x,y)|_3 = \max(|x|, |y|)$$

(почему они эквивалентны?) или ещё какую-нибудь норму, эквивалентную этим. Вообще говоря, нет оснований предпочесть какую-то из этих норм другим, но в каком-нибудь конкретном случае в связи с его спецификой такие основания могут появиться. Например, если E, F — гильбертовы пространства, то предпочтительна норма $|\,.\,|_2$ (как Вы думаете, почему?).

Упражнение. Если E, F – банаховы пространства, то и $E \oplus F$ — тоже банахово пространство.

Если E, F — нормированые пространства, то можно говорить о непрерывности любого отображения $f : P \to F$, где $P \subset E$ — любое подмножество: f непрерывно в точке x, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall y \in P, |y - x| < \delta \Longrightarrow |f(y) - f(x)| < \varepsilon.$$

Говорят, что f непрерывно, если f непрерывно во всех точках M.

Упражнение. Докажите, что если f является непрерывным в x_0 при использовании каких-нибудь норм в E и F, то f является непрерывным в x_0 при использовании других норм, эквивалентных исходным.

Упражнение. Докажите, что непрерывность линейного отображения (или, как чаще говорят, линейного оператора) $A : E \to F$ в какой-нибудь точке x_0 равносильна его непрерывности во всех точках E и что это равносильно также следующему свойству: имеется такое $C > 0$, что $|Ax| \leqslant C|x|$ при всех $x \in E$.

Последнее свойство называют ограниченностью линейного отображения A. (Здесь имеется некоторая несогласованность терминологии для линейных отображений и для функций (со значениями в $\mathbb{R}, \mathbb{C}$ или вообще в другом нормированном пространстве F). Функция f называется ограниченной, если имеется такое C, что $|f(x)| \leqslant C$ при всех x. Функция

$$f : \mathbb{R} \to \mathbb{R} \qquad x \mapsto x$$

ограничена как линейное отображение, но не как функция.)

“Величину” ограниченного линейного оператора характеризуют посредством его нормы $\|A\|$, определяемой так: $\|A\| = \sup_{x \neq 0} \frac{|Ax|}{|x|}$.

Теорема Хана–Банаха.

Линейный (и однородный) функционал на векторном пространстве E — это линейное отображение $\xi : E \to \mathbb{R}$ (или $E \to \mathbb{C}$). Под ограниченностью линейного функционала понимают его ограниченность как линейного отображения, а не как функции. Иными словами, линейный функционал ξ ограничен, если имеется такое $C > 0$, что $|\xi(x)| \leqslant C|x|$ при всех x. В конечномерном случае все линейные функционалы ограничены, но в бесконечномерном случае это не так. Ниже речь идёт только об ограниченных линейных функционалах. Порой я даже не буду специально упоминать об их ограниченности, подразумевая, что слова "линейный функционал" означают "ограниченный линейный функционал".

Ограниченные линейные функционалы образуют вектоное пространство E^*, которое называется сопряжённым к E. (Сумма функционалов и произведение функционала на число определяются очевидным образом.) Нормой ограниченного линейного функционала называется $\|\xi\| = \sup_{x\neq 0} \frac{|\xi(x)|}{|x|}$. Рассматриваемое с этой нормой, E становится нормированным пространством.

Говоря о сопряжённом пространстве E^*, подразумевают, что в нём используется именно эта норма $\| \,.\, \|$, и поэтому вместо $(E^*, \| \,.\, \|)$ пишут просто E^*.

Упражнение. Докажите, что E^* полно (является банаховым пространством), независимо от полноты или неполноты E.

Теорема Хана–Банаха. Пусть на векторном подпространстве $V \subset E$ нормированного пространства E имеется ограниченный линейный функционал $\xi : V \to \mathbb{R}$. Тогда его можно продолжить на всё E (имеется в виду — продолжить как ограниченный линейный функционал) с сохранением нормы. Иными словами, имеется такой ограниченный линейный функционал $\eta : E \to \mathbb{R}$, что его ограничение (сужение) $\eta|V$совпадает с ξ и $\|\eta\| = \|\xi\|$.

Это единственная нетривиальная теорема в начальной части курса. (Упоминавшаяся выше теорема Вейерштрасса, конечно, тоже нетривиальна, но она упоминалась только попутно при рассмотрении примера и, собственно, немного повозившись, можно было бы и без неё убедиться в в неполноте рассматривавшегося тогда пространства E.)

Следствие. $\forall x \in E, x \neq 0 \quad \exists \xi \in E^* \quad \xi(x) = 1, \|\xi\| = \frac{1}{|x|}$.

Действительно, на векторном пространстве $\mathbb{R}x = \{\lambda x;\ \lambda \in \mathbb{R}\}$ имеется линейный функционал

$$\xi : \mathbb{R}x \to \mathbb{R} \qquad \xi(\lambda x) = \lambda.$$

Ясно, что $\xi(x) = 1$, а $\|\xi\| = \sup_{\lambda \neq 0} \frac{|\lambda|}{|\lambda x|} = \sup \frac{|\lambda|}{|\lambda||x|} = \frac{1}{|x|}$. Продолжение ξ на всё E с сохранением нормы (это продолжение мы снова обозначим через ξ) и является функционалом, существование которого утверждается в следствии.

Метрические и топологические пространства.

Метрическое пространство — это множество M, на котором задана метрика d, т.е. функция $d : M \times M \to \mathbb{R}_+$, обладающая свойствами:

$d(x, y) = 0 \Longleftrightarrow x = y$;

$d(x, y) = d(y, x)$;

$d(x, y) \leqslant d(x, z) + d(z, y)$ ("неравенство или аксиома треугольника").

Нормированное пространство $(E, |\,.\,|)$ является метрическим с $d(x, y) = |x - y|$.

Пусть (M, d) и (N, r) — метрические пространства. Их прямое произведение можно превратить в метрическое пространство $(M \times N, D)$. Стандартного выбора метрики D в общем случае не имеется, но аналогично тому, что делалось для нормированных пространств, можно предложить некий класс "эквивалентных" метрик. Две метрики ρ_1 и ρ_2 на множестве X мыы назовём эквивалентными, если имеются такие $a, b > 0$, что

$$a\rho_1(x, y) \leqslant \rho_2(x, y) \leqslant \rho_1(x, y) \quad \text{при любых } x, y \in X.$$

Вот примеры эквивалентных друг другу метрик, которые могут играть роль D:

$$D_1((x_1, y_1), (x_2, y_2)) = d(x_1, x_2) + r(y_1, y_2),$$
$$D_2((x_1, y_1), (x_2, y_2)) = \sqrt{d(x_1, x_2)^2 + r(y_1, y_2)^2},$$
$$D_3 = \max(d(x_1, x_2)^2, r(y_1, y_2)^2)$$

(между прочим, $D_i((x_1, y_1), (x_2, y_2)) = |(d(x_1, x_2), r(y_1, y_2))|_i$, где $|\,.\,|_i$ — введённые ранее нормы в $\mathbb{R}^2$).

Упражнение. Проверьте, что D_i — действительно метрики и что они эквивалентны друг другу.

Для метрических пространств очевидным образом вводится понятие предела последовательности x_n элементов этого пространства. Если

$f : (M, d) \to (N, \rho)$[3] — отображение метрических пространств, то очевидным образом определяются понятия непрерывности f в точке $x_0 \in M$ и предела $\lim_{x \to x_0} f(x)$. В обоих случаях (последовательности и отображения) предел (если он существует) является единственным.

С помощью прямого произведения $M_1 \times M_2$ множеств M_1 и M_2 можно рассматривать функцию двух переменных $f(x, y)$ (где $x \in M_1,\ y \in M_2$) со значениями в третьем множестве N как отображение

$$f : M_1 \times M_2 \to N \qquad (x, y) \mapsto f(x, y).$$

Это позволяет в ряде случаев переформулировать вопросы относительно функций нескольких переменныхя как вопросы, касающиеся функций одной переменной — функций, аргументами которых являются элементы соответствующих прямых произведений.

Так, если в каждом из M_i и в N введены метрики d_i и r, то можно говорить о непрерывности функции $f(x, y)$: она непрерывно в точке (a, b), если

$$\forall \varepsilon > 0\ \exists \delta > 0 \quad \text{если } d_1(x, a) < \delta \text{ и } d_2(y, b) < \delta,\ \text{ то } r(f(x, y), f(a, b)) < \varepsilon.$$

Непрерывность этой функции двух переменных (скажем, при $x = a,\ y = b$) равносильна непрерывности f как отображения $M_1 \times M_2 \to N$ в соответствующей точке (в данном случае - в точке (a, b)).

Упражнение. Докажите, что если (M, d) — метрическое пространство, то $d(x, y)$ — непрерывная функция от (x, y).

Фундаментальная последовательность элементов $x_n \in M$ — это такая последовательность, что $\forall \varepsilon > 0\ \exists N = N(\varepsilon)\ \forall n, m \geqslant N \quad d(x_n, x_m) < \varepsilon$. Сходящаяся последовательность всегда является фундаментальной. Если в (M, d) всякая фундаментальная последовательность сходится, то пространство (M, d) называется полным.

Упражнение. Покажите, что если d_1, d_2 — эквивалентнын метрики в M, то одни и те же последовательности элементов M являются сходящимися (или фундаментальными) в (M, d_1) и в $(M.d_2)$, причём в этих двух пространствах они имеют одни и те же пределы. Фундаментальные последовательности в этих пространствах тоже совпадают. Если одно из них полное, то и другое тоже.

[3] В этой записи сказаны две вещи: что f есть отображение $M \to N$ и что M рассматривается с метрикой d, а N — с метрикой ρ.

Рассматривая нормированное пространство $(E, |\,.\,|)$ как метрическое пространство с метрикой $d(x,y) = |y-x|$, мы видим, что для него все эти понятия совпадают с теми, которые были введены ранее применительно к нормированным пространствам. Надо только учесть следующее.

Для нормированных пространств E и F мы говорили о непрерывности отображения $f : P \to F$ подмножества $P \subset E$ в F. Теперь в определении непрерывности мы можем не говорить об отображениях подмножеств. Дело в том, что подмножество P метрического пространства (M, d) само является метрическим пространством с “индуцированной” метрикой $d_P = d|P \times P$ (т.е. с прежней метрикой d, рассматриваемой только для точек P). Непрерывность отображения $f : P \to F$ подмножества $P \subset E$ в F можно понимать просто как непрерывность отображения метрического пространства (P, d_P) (где d_p индуцировано на P метрикой $d(x,y)$, которая в обсуждаемом сейчас случае нормированного пространства E определяется нормой: $d(x,y) = |y-x|$) в метрическое пространство (F, ρ) (где в F сейчас метрика $\rho(u,v) = |v-u|$).

Упражнение. а). Докажите, что если P — замкнутое подмножество[4] полного метрического пространства (M, d), то (P, d_P) тоже является полным метрическим пространством.

б). Пусть X — просто множество без какой-либо дополнительной структуры, а (M, d) — полное метрическое пространство. Обозначим через $B(X, M)$ множество всевозможных отображений $f : X \to M$, которые являются ограниченными в том смысле, что при некотором (а тогда и любом) $a \in M$ все расстояния $d(f(x), a)$ ограничены сверху некоторым не зависящим от x числом, т.е. $\sup_{x \in X} d(f(x), a) < \infty$. (Буква B происходит от bounded.) Введём в $B(X, M)$ расстояние $D(f, g) = \sup_x d(f(x), g(x))$. Покажите, что это действительно метрика в $B(X, M)$ и что $(B(X, M), D)$ — полное метрическое пространство. Какое отношение сходимость $f_n \to g$ в метрике D имеет к известной из анализа равномерной сходимости[5]?

[4]Определение замкнутого множества приводится ниже, но вероятно оно уже известно читателю (сейчас оно нужно только для метрических пространств). В подобных случаях в этом справочном тексте я позволяю себе забегать вперёд.

[5]Там она вводится вначале для функций на отрезке, затем — на подмножестве (возможно, поначалу не на произвольном) $\mathbb{R}^n$. Но оказывается, что по существу её определение не требует ничего от X, тогда как в M должна быть введена метрика.

С другой стороны, стоит напомнить, что о равномерной сходимости можно говорить применительно не только к ограниченным, но и к неограниченным функциям. Но, в). Пусть в дополнение к б) в X введена метрика r. Тогда можно говорить о непрерывных отображениях $f : X \to M$. Множество всех таких отображений обозначим через $C(X, M)$, а его пересечение с $B(X, M)$ из б) — через $BC(X, M)$. Докажите, что $BC(X, M)$ является замкнутым подмножеством пространства $B(X, M)$ (и следовательно полно относительно метрики $R_{BC(X,M)}$).

Определения сходимости, предела и непрерывности можно перефразировать в более геометрических терминах. Открытым шаром $U_r(x_0)$ (имеется в виду — шаром пространства (M, d)) с центром в точке x_0 и радиуса $r > 0$ называется множество

$$U_r(x_0) = \{x \in M;\ d(x, x_0) < r\}.$$

Непрерывность отображения $f : (M, d) \to (N, \rho)$ означает, что для любого открытого шара $V_\varepsilon(f(x_0))$ (шара в N с центром в $f(x_0)$ и какого-то радиуса $\varepsilon > 0$) в M имеется такой шар $U_\delta(x_0)$, что $f(U_\delta(x_0)) \subset V_\varepsilon(f(x_0))$. (Различные буквы V и U для шаров призваны напомнить, что эти шары суть шары в различных пространствах.) Далее, точка $y_0 \in N$ является пределом $\lim_{x \to x_0} f(x)$, если для любого шара $V_\varepsilon(y_0)$ имеется такой шар $U_\delta(x_0)$, что $f(U_\delta(x_0) \setminus \{x_0\}) \subset V_\varepsilon(y_0)$. (В подобных случаях принято ничего не говорить об $f(x_0)$. Может быть, это y_0, а может быть и нет. Может быть, f вообще не определено в самой точке x_0, а определено только для всех достаточно близких к x_0 точек $x \neq x_0$.) Точка $x_0 \in E$ является пределом последовательности x_n, если для любого шара $U_\varepsilon(x_0)$ имеется такое N, что $x_n \in U_\varepsilon(x_0)$ при всех $n \geqslant N$.

Перефразировка определений с помощью шаров — это только первый, довольно очевидный шаг в процессе их геометризации. Главный и априори далеко не столь очевидный шаг состоит в замене шаров на так называемые открытые множества.

Подмножество A метрического пространства (M, d) называется открытым, если вместе с каждой своей точкой оно содержит и все достаточно близкие к ней точки. Подробнее это означает: для любой точки $x \in A$ имеется такое число $\varepsilon > 0$, что все точки y, для которых $d(x, y) < \varepsilon$, тоже содержатся в A, т.е. что шар $U_\varepsilon(x) \subset A$.

Тривиальный пример: всё M открыто, потому что вообще любой шар

в общем, это далеко столь важно, как равномерная сходимость лоя ограниченных функций. Перефразировка для отображений $X \to M$ достаточно очевидна, но мне не кажется нужным на ней останавливаться.

пространства M содержится в M. А вот является ли открытым пустое подмножество $\emptyset$? Условие, определяющее открытость, начинается со слов: "для любой точки $x \in A$", а когда $A = \emptyset$, то таких точек нет. Считать ли, что в таком случае условие автоматически выполняется? Положительный ответ может показаться каким-то отдельным соглашением (и даже не очень-то мотивированным), но давайте перефразируем наше условие так: "если $x \in A$, то ...". В такой формулировке оно выполняется: те x, для которых это "то ..." могло бы нарушаться, должны быть элементами A (от $x \notin A$ ведь ничего и не требуется), а таких x нет.

Упражнение. Докажите, что открытый шар является открытым множеством.

Оказывается, в приведённых выше формулировках понятий предела и непрерывности, использующих шары, можно почти автоматически заменить шары на открытые множества. Например, последовательность $x_n \to x_0$, если для любого открытого множества U, содержащего x_0, найдётся такое N, что $x_n \in U$ при $n \geqslant N$. Ведь это U содержит некоторый открытый шар $U_\varepsilon(x_0)$, и если $x_n \to x_0$, то все x_n с достаточно большими номерами n лежат в этом шаре; тем более они лежат в U. Обратно, пусть для любого открытого множества U, содержащего x_0, существует такое N, что $x_n \in U$ при $n \geqslant N$. К таким открытым множествам относятся и шары $U_\varepsilon(x_0)$, так что выполняется условие, фигурирующее в перефразировке определения предела в терминах шаров.

Упражнение. Докажите, что пересечение двух (а значит, и любого конечного числа) открытых множеств открыто и что объединение любой (даже бесконечной) системы открытых множеств открыто.

Упражнение. а). Докажите, что отображение $f : M \to N$ (где M, N — метрические пространства) непрерывно в точке $x_0 \in M$ тогда и только тогда, когда для любого открытого множества $V \subset N$, содержащего $f(x_0)$, найдётся такое открытое множество $U \subset M$, содержащее x_0, что $f(U) \subset V$. Это отображение всюду непрерывно тогда и только тогда, когда для любого открытого множества $V \subset N$ найдётся такое открытое множество $U \subset M$, что прообраз[6] $f^{-1}(V) \subset U$.

б). Сформулируйте в терминах открытых множеств определение предела отображения.

Последнее упражнение можно рассматривать как формулировку понятий непрерывности и предела, которая внешне выглядит более геометрической, чем обычная формулировка из курса анализа, где пишутся неравенства. Правда, приведённая выше перефразировка этих определений с шарами тоже геометрична. Но шары слишком сильно зависят от метрики. Одной из причин обращения к открытым множествам является стремление достичь некоторой независимости от используемой метрики. Конечно, полной независимости не может быть — ведь сами открытые множества определяются с помощью метрики. Но если при использовании двух метрик получаются одни и те же открытые множества, то все те свойства, которые можно определить в терминах открытых множеств, для этих метрик будут одинаковыми.

Можно посмотреть на всё это ещё с такой стороны. Метрические пространства (M, d) и (N, ρ) называются гомеоморфными, если существует непрерывное взаимно-однозначное отображение M на N, обратное к которому тоже непрерывно. Само это отображение тогда называют гомеоморфизмом (подробнее: гомеоморфизмом пространства M на N). Нам хотелось бы выделить свойства пространств, не меняющиеся при гомеоморфизмах; с этой точки зрения гомеоморфиные пространства одинаковы. Такие вопросы относятся к топологии — математической дисциплине, являющейся, так сказать, геометрией непрерывности, одной только непрерывности, когда игнорируется всё остальное — расстояния, углы, прямолинейность и т.д. Уже на доматематическом интуитивном уровне представляется наглядно ясным различие между непрерывностью и разрывностью. Одно из первых, что делают в курсе математического анализа — это уточняют идею непрерывности (хотя и не в полной общности), давая чёткое формальное определение непрерывности функции (как и

[6]Т.е. множество тех $x \in M$, образы которых при отображении f попадают в V:

$$f^{-1}(V) = \{x;\ x \in M, f(x) \in V\}.$$

Прообраз $f^{-1}(y)$ элемента y — это то же самое, что и прообраз $f^{-1}(\{y\})$ “одноточечного” множества $\{y\}$ (множества, состоящего из одного-единственного элемента y), т.е. множество тех $x \in M$, для которых $f(x) = y$.

Наряду со сказанным, под прообразом элемента y часто понимают любой из элементов множества $f^{-1}(y)$ (о последнем в таком случае говорят как о полном прообразе).

родственного понятия предела) и доказывая несколько примыкающих сюда утверждений. Однако основные объекты математического анализа — производные, ряды и интегралы — связаны преимущественно с другими идеями, а непрерывность (да в какой-то степени и предел) — не более чем необходимый пререквизит. Насколько он необходим для курса анализа, настолько он там и развивается, но не более того. В топологии же идея непрерывности выступает на первый план.

Из данной выше перефразировки непрерывности отображения в терминах открытых множеств явствует, что гомеоморфизм — это такое отображение $f : M \to N$, при котором образы и прообразы открытых множеств являются открытыми. Таким образом, пространства M и N имеют, так сказать, одинаковые системы открытых множеств — если перейти из M в N, заменив каждую точку $x \in M$ её образом $f(x)$, с открытыми множествами как бы не произойдёт изменений. Мы возвращаемся к уже промелькнувшей идее, что всё дело в открытых множествах.

Эти лекции посвящены анализу, а не топологии. Но мы не будем довольствоваться той ограниченной трактовкой топологических понятий, которая свойственна начальному курсу анализа, и даже той, при которой систематически привлекаются метрические пространства, но не более как. В том, что касается непрерывности и т.п., мы приводим формулировки, которые выработаны в топологии. Но это будет сделано только затем, чтобы куда более конкретный материал, составляющий основное содержание данного курса, воспринимался в надлежащей перспективе[7].

В первой трети XX в. постепенно сложилось и стало общепринятым мнение, что топологические свойства пространства M — это те его свойства, которые описываются с помощью его открытых подмножеств. А это привело к определению: топологическим пространством называется множество M вместе с какой-нибудь системой его подмножеств $\mathcal{T}$, име-

[7]В самой же топологии подобные определения и анализ связей между ними, включая освещение роли тех или иных ограничений на рассматриваемые объекты, является, в общем, необходимой частью этой науки. Эту часть называют "теоретико-множественной или общей топологией". Она далеко не исчерпывает всей топологии и даже даёт о её содержании ещё ме́ньшее представление, чем дают о курсе матанализа приводимые в его начале краткие сведения о пределах. Более того: можно довольно далеко продвинуться в топологии, имея дело только с подмножествами привычного $\mathbb{R}^n$. Но даже в той её части, где это возможно, такое самоограничение довольно быстро становится обременительным, вынуждая задерживаться на различных "посторонних" для топологии вещах и тем самым отвлекая внимание от собственно топологических вопросов.

нуемых открытыми, если только эта система удовлетворяет нескольким простым условиям, которые нам уже известны для системы открытых подмножеств метрического пространства: M и $\emptyset$ открытые; пересечение двух (а значит, и любого конечного числа) открытых подмножеств открыто; объединение любой (даже бесконечной) системы открытых подмножеств открыто. Такая система $\mathcal{T}$ может определяться при помощи метрики, но может определяться и как-то иначе, лишь бы она обладала указанными свойствами. Желая явно упомянуть о системе открытых подмножеств, можно говорить о "топологическом пространстве $(E, \mathcal{T})$", но чаще говорят просто о "топологическом пространстве E". Совокупность всех открытых подмножеств $\mathcal{T}$ топологического пространства называют его "топологией" (как мы помним, это слова обозначает также и причастную к топологическим пространствам математическую дисциплину).

Открытой окрестностью точки x_0 пространства M называется любое содержащее эту точку открытое множество. Окрестностью точки x_0 называется любое множество, содержащее некоторую открытую окрестность этой точки [8]. В случае метрического пространства окрестность точки x_0 — это множество, содержащее некоторый открытый шар с центром в x_0[9].

Всё это пока что — не более, чем названия. Но опыт показывает, что окрестности могут с успехом заменять шаровые окрестности и что в терминах окрестностей получаются более короткие перефразировки различных определений и рассуждений (к тому же автоматически имеющие

[8] В литертуре порой под "окрестностями" понимают только "открытые окрестности".

[9] Такой шары, кстати, теперь уместно назвать шаровой окрестностью или, если мы хотим указать радиус шара, шаровой r-окрестностью. Впрочем, раз "на сцене" появилось r, которое ни к каким окрестносям, кроме шаровых, не имеет отношения, то часто говорят об "r-окрестности $U_r(x_0)$," имея в виду шаровую окрестность радиуса r точки x_0. Я последним названием всё-таки не злоупотребляю, но уже начал пользоваться соответствующим обозначением. Замечу кстати, что r-окрестностью $U_r(A)$ множества A (вот тут действительно приходится говорить об r-окрестности, а не о шаре) называется множество точек, каждая из которых находится на расстоянии $< r$ от некоторой точки из A:

$$U_r(A) = \{x; \quad \exists a \in A \quad d(x,a) < r\} = \bigcup_{a \in A} U_r(a).$$

бо́льшую общность). Например: отображение $f : M \to N$ непрерывно в точке x_0, если прообраз любой окрестности V точки $f(x_0)$ в N является окрестностью x_0 в M, т.е. если в M у x_0 имеется такая окрестность U, что $f(U) \subset V$.

Упражнение. Перефразируйте в терминах окрестностей определения предела последовательности точек или отображения.

Упражнение. Докажите, что пересечение двух окрестностей данной точки — окрестность. отсюда следует, что пересечение конечной системы окрестностей — окрестность. Верно ли это для бесконечной системы?

Упражнение. В порядке тривиальной тренировки единственно на понимание названий проверьте, что:

а). В “одноточечном” (т.е. состоящем всего из одной точки) множестве M имеется только одна тополология $\mathcal{T}$.

б). В множестве, состоящем более чем из одной точки, имеются различные топологии. Сколько их имеется в множестве M, состоящем из двух точек? Нет ли среди них двух таких $\mathcal{T}_1$ и $\mathcal{T}_2$, что пространства $(M, \mathcal{T}_1)$ и $(M, \mathcal{T}_2)$ гомеоморфны? (Для топологических пространств гомеоморфизм означает то же самое, что и для метрических — взаимнооднозначное отображение, которое непрерывно и обратное к которому тоже непрерывно.)

в). В любом множестве M можно ввести топологию $\mathcal{T}_{\min}$, для которой $\mathcal{T}_{\min} = \{\emptyset, M\}$ (она является “минимальной среди всех топологий” в том смысле, что любая топология $\mathcal{T} \supset \mathcal{T}_{\min}$) и “дискретную топологию” $\mathcal{T}_d$, содержащую все подмножества множества M[10] (она является “максимальной среди всех топологий”). Проверьте, что это действительно топологии и что дискретная топология определяется метрикой

$$d(x, y) = \begin{cases} 1 & \text{при } x \neq y, \\ 0 & \text{при } x = y \end{cases}$$

(а d — это действительно метрика?). Определяется ли какой-нибудь метрикой топология $\mathcal{T}_{\min}$?

г). Какие последовательности x_n являются сходящимися в этих топологиях? Имеет ли место единственность предела? Какие отображения $(M, \mathcal{T}_d) \to (N, \mathcal{T})$ (где $(N, \mathcal{T})$ — какое-то топологическое пространство) непрерывны?

---

[10]Иногда словосочетанию “дискретная топология” придают более широкий смысл.

Упражнение. Введённые в предыдущем упражнении топологии не играют серьёзной роли (не считая того, что привлечение $\mathcal{T}_d$ позволяет при случае считать, что формулировки, которые не связаны с топологией и потому вроде бы должны даваться отдельно от формулировок, связанных с топологией, на самом деле охватываются последними). Вот пример не менее "странной" топологии $\mathcal{T}$, которая весьма важна, только не в анализе, которым мы интересуемся, а в алгебраической геометрии[11]: это топология в $\mathbb{R}$ или $\mathbb{C}$, содержащая все те множества, дополнения к которым конечны. Проверьте, что это действительно топология.

Упражнение. Присоединим к множеству натуральных чисел $\mathbb{N}$ ещё один элемент — символ ∞ — и введём в множестве $N = \mathbb{N} \cup \{\infty\}$ топологию $\mathcal{T}$, приняв, что элементы $\mathcal{T}$ суть:

любые подмножества $\mathbb{N}$ (иными словами, любые подмножества N, не содержащие ∞);

подмножества N, которые содержат ∞ и имеют конечные дополнения. (Это множества вида $A \cup \{\infty\}$, где A — подмножество $\mathbb{N}$, содержащее все натуральные числа, бóльшие какого-нибудь числа (зависящего от A) и, возможно, ещё какие-нибудь числа.)

а). Проверьте, что $\mathcal{T}$ — топология.

б). Последовательность x_n точек какого-нибудь топологического пространства M с формальной точки зрения является отображением

$$\mathbb{N} \to M \qquad n \mapsto x_n$$

и потому может рассматриваться как определённое всюду вне ∞ отображение N в M. Покажите, что предел этой последовательности — это в точности предел последнего отображения при $n \to \infty$. (Тем самым достигается некоторая унификация понятий.)

в). Соответствует ли топология $\mathcal{T}$ какой-нибудь метрике? (Рассмотрите на числовой прямой множество точек $p_n = \frac{1}{n}$, к которому добавлена точка 0.)

Решая одно из этих упражнений, читатель должен был обнаружить неприятный факт: постулированные нами свойства открытых множеств, позволяя дать вполне разумное определение предела последовательности

[11] Правда, $\mathcal{T}$ важна не сама по себе. Важна так называемая "топология Зарисского", о которой здесь не будет речи, а $\mathcal{T}$ — её частный случай, который представляет интерес только как иллюстрация общего случая, а сам по себе едва ли может претендовать на особое значение.

точек или отображения, не гарантируют его единственности. Чтобы её гарантировать, к условиям, которым должна удовлетворять топология $\mathcal{T}$, добавляют ещё одно. Мы приведём вариант дополнительного условия, "наиболее близкий по духу к повседневной жизни"(понимая под таковой математический анализ — не только как учебный предмет, но и как весь комплекс дисциплин "аналитического характера"). Это условие называется "аксиомой Хаусдорфа" и гласит: у любых двух различных точек топологического пространства $(M, \mathcal{T})$ имеются непересекающиеся окрестности. (В терминах самой топологии $\mathcal{T}$: для любых $x, y \in M, \quad x \neq y$, имеются такие $U, V \in \mathcal{T}$, что $x \in U, y \in V, U \cap V = \emptyset$.) Топологическое пространство, удовлетворяющее аксиоме Хаусдорфа, называют короче хаусдорфовым (а также отделимым) пространством.

Упражнение. Докажите, что:

а). В хаусдорфовом пространстве любое одноточечное множество (т.е. множество, состоящее из одной точки) замкнуто. (Определение замкнутого множества приводится ниже сразу для топологических пространств; если читателю оно известно только для метрических пространств (или, паче чаяния, вообще не известно), то надо познакомиться с этим общим определением — оно достаточно коротко и поэтому такое отклонение от линейной последовательбности изложения, тем более в тексте справочного характера, кажется мне вполне допутимым).

б). Аксиома Хаусдорфа нарушается в приведённом выше примере топологии $\mathcal{T}$ в $\mathbb{R}$, которая, как было сказано, является частным случаем важной для алгебраической геометрии топологии Зарисского[12]. (Напоминаю, что $\mathcal{T}$ принадлежат множества с конечными дополнениями.) Однако одноточечные подмножества замкнуты в этой топологии.

в). Последовательность $x_n = n$ сходится в пространстве $(\mathbb{R}, \mathcal{T})$ из б) к любой его точке. (При желании читатель может убедиться также, что то же самое верно для любой последовательности x_n, принимающей каждое значение только конечное число раз. Когда же некоторые значения принимаются бесконечное число раз, то последовательность имеет единственный предел (какой?), если такое значение только одно, и не имеет предела, если таких значений более одного.)

В свете в) понятно, что от пределов в топологическом пространстве

[12]Поэтому аксиома Хаусдорфа, вполне удовлетворяя нуждам анализа, не подходит для целей алгебры и алгебраической геометрии. Имеются менее ограничительные аксиомы, покрывающие и эти нужды (одна из них состоит в упоминаемой ниже в основном тексте замкнутости одноточечных множеств), но нам они не понадобятся.

$(\mathbb{R}, \mathcal{T})$ мало толку (а ведь, как говорилось, топологии такого типа важны для алгебраической геометрии). Это до некоторой степени объясняет, почему в общей топологии пределы отступают на задний план. ("До некоторой степени" — потому, что (по другим причинам) их роль уменьшается и при ограничении одними хаусдорфовыми пространствами, что, повторяю, приемлемо при ориентации на запросы анализа, а не алгебры или алгебраической геометрии.)

Хотя топологические свойства пространства в конечном счёте определяются системой его открытых подмножеств, отнюдь не очевидно, что стоит принять точку зрения, при которой эта система ставится, так сказать, во главе угла. Учащемуся это так же не очевидно, как это было не очевидно выдающимся учёным в начале XX в., которые пробовали различные другие варианты построения топологии. Во всех вариантах расстояния могли использоваться только при определении какого-то исходного класса объектов, да и то в иных случаях эти объекты могли вводиться и как-нибудь иначе, в дальнейшем же всё делалось на основании одного этого класса. Практика первой трети века показала, что вариант с открытыми множествами всего лучше. С тех пор стало общепринятым понятие "топологического пространства" в том именно виде, как оно определено выше. Для начала ограничиваются только теми условиями на топологию $\mathcal{T}$, которые там указаны, но позднее появляются различные дополнительные условия (например, аксиома Хаусдорфа), обеспечивающие наличие у топологического пространства различных "хороших" свойств, существенных в приложениях. Мы не будем вникать в эти абстрактные дебри — наши топологические пространства будут обычно метрическими пространствами (M, d), для которых часть этих условий выполняется автоматически, а часть условий представляет собой определённые ограничения на (M, d), но для метрических пространств эти условия можно формулировать иначе и в чём-то проще, чем для общих топологических [13].

[13]Бывает, что несколько вариантов того или иного определения или утверждения, эквивалентных применительно к подмножествам $\mathbb{R}^n$, не эквивалентны для более общих метрических пространств, — с этим мы будем вынуждены считаться, — и что несколько вариантов того или иного определения или утверждения, эквивалентных применительно к метрическим пространствам, не эквивалентны для ещё более общих топологических пространств. При разработке общей теории пришлось думать о рациональном выборе вариантов (что не всегда было просто). При изучении этой теории приходится загрузить в память некоторое количество окончательно отобранных вариантов и понимать, почему решено было остановиться именно на них (а также

Я буду придерживаться такого курса: как правило, я довольствуюсь метрическими или даже банаховыми пространствами, но при этом часто привожу формулировки и рассуждения не в привычных терминах ε-δ и соответствующих неравенств, а в терминах открытых множеств и окрестностей.

В заключение необходимо обратить внимание, что не всё, о чём говорилось применительно у метрическим пространствам, можно выразить с помощью открытых множеств, т.е. не всё "улавливается" той системой понятий, которая связана с топологическими пространствами. Эта система понятий позволяет придать точный смысл выражениям, в которых так или иначе говорится о точках x, близких к фиксированной точке a. Но она никак не позволяет выразить что-нибудь вроде "близости друг к другу точек x и y, ни одна из которых не является фиксированной", или вроде того, что "точка x столь же близка к a, как y — к b". Применительно к топологическим пространствам не приходится говорить о какой-либо равномерности (скажем, о равномерной непрерывности), а также о фундаментальных последовательностях и полноте[14]. Заметим, что соответствующие свойства не инвариантны относительно гомеоморфизмов. Так, полное метрическое пространство может быть гомеоморфно неполному. Пример — подмножества $A = [1, \infty)$, $B = (0, 1]$ обычной числовой прямой (рассматриваемые как метрические пространства с индуцированными метриками d_A и d_B). A полно, B неполно, а отображение

$$A \to B \qquad x \mapsto \frac{1}{x}$$

является гомеоморфизмом.

Замкнутые множества, замыкания, внутренности и границы.

При изучении курса матанализа студенту приходится решать ряд задач на пределы, тогда как задач, посвящённых непрерывности, ему предлагают (если вообще предлагают) намного меньше. Возникает впечатление, что пределы "работают по-настоящему", а непрерывность нужна

когда они эквивалентны другим). Имея дело только с подмножествами $\mathbb{R}^n$, мы были бы совсем свободны от таких забот, а имея дело с метрическими пространствами, мы от них не совсем свободны, но всё же менее обременены ими, чем если бы мы занимались топологическими пространствами.

[14]Помимо топологических пространств, в общей топологии вводятся ещё так называемые "равномерные пространства", наделённые некоторой структурой, "улавличающей" равномерность и полноту. Но мы не будем иметь дела с этими пространствами.

чуть ли не “преимущественно для души” и “научного антуража”. В общей топологии дело обстоит наоборот: пределы, оказывается, не всегда достаточно хорошо отражают топологические свойства пространства. Нас это (почти) не касается, раз мы занимаемся анализом, а не топологией, однако в этих лекциях непрерывность, пожалуй, оттесняет пределы с переднего плана. Так вышло вовсе не потому, что пределы важны только для элементарных вопросов из обычного курса матанализа. Нет, в самых “высоких” разделах анализа пределы играют важную роль (и я сам, будучи аналитиком, потратил на них уйму времени). Просто я восстанавливаю баланс, который при начальном ознакомлении с анализом поневоле нарушается.

Сейчас мы остановимся на нескольких понятиях, которые, вероятно, хотя бы отчасти известны читателю применительно к подмножествам $\mathbb{R}^n$ или даже к подмножествам более общих метрических пространств. В этом случае часть их обычно вводится с использованием понятия предела (я повторю соответствующие определения). Однако их можно ввести, и не обращаясь к пределам; в таком виде определения сохраняют смысл и в топологических пространствах. Мы остановимся на этом не столько ради общности, сколько затем, чтобы, так сказать, “придать этим понятиям самостоятельное значение” и приучить читателя к рассуждениям, в которых вместо пределов и метрики привлекаются открытые множества и окрестности (хотя бы при этом речь шла о метрических пространствах).

Итак, сперва напоминается пара определений, которые вначале относятся к подмножествам метрических пространств и даются с использованием понятия предела.

Предельная точка подмножества A метрического пространства (M, d) — это такая точка $x \in M$, к которой сходится некоторая последовательность точек этого подмножества, отличных от неё самой. Замыкание clos A [15] множества A получается присоединением к A всех его предельных точек[16]. Множество A называется замкнутым, если оно содержит все

[15] Замыкание обозначают также посредством черты наверху: $\overline{A}$. Но черта — слишком соблазнительный знак, чтобы не использовать его также и в иных смыслах, например, для обозначения комплексного сопряжения. Поэтому я пишу clos (от closure). Кажется, я встречал также обозначение $[A]$.

[16] Точка a самого множества A может быть, а может и не быть предельной для A — всё равно она включается в clos A. Если она не является предельной, то у неё имеется окрестность U, не содержащая других точек A, кроме a. В последнем случае (т.е. тогда, когда $\exists U \in \mathcal{T} \quad U \cap A = \{a\}$) говорят, что a — это изолированная точка

свои предельные точки, т.е. совпадает со своим замыканием (короче говоря, если $\mathrm{clos} A = A$).

В перефразировке этого определения, не привлекающей пределов и годных для топологических пространств, вместо предельных точек фигурируют "точки прикосновения[17] подмножества A пространства M". Точка $a \in M$ называется точкой прикосновения подмножества $A \subset M$, если в каждой её окрестности U имеется хоть одна точка этого подмножества. Замыкание $\mathrm{clos}\ A$ — это множество всех точек прикосновения множества A. (В частности, если $a \in A$, то саму a можно взять в качестве содержащейся в окрестности U точки из A. Поэтому $A \subset \mathrm{clos}\ A$.) A попрежнему называется замкнутым, если $A = \mathrm{clos}\ A$.

Упражнение. Докажите, что:

а). Если M — метрическое пространство, то точка прикосновения его подмножества A — это или точка из A, или предельная точка множества A (эти "или" не являются взаимно исключительными). Значит, в этом случае второе определение замыкания и замкнутого подмножества совпадает с первым.

б). Замкнутые множества суть те и только те множества, дополнения к которым открыты.

в). $\mathrm{clos}\ A$ является наименьшим замкнутым множеством, содержащим A.

Упражнение. Из п. б) предыдущего упражнения видно, что совокупность $\mathcal{F}$ замкнутых подмножеств топологического пространства M однозначно определяет его тополгию (т.е. совокупность $\mathcal{T}$ открытых подмножеств). Пусть теперь нам задана какая-то система $\mathcal{F}$ подмножеств M. Обозначим через $\mathcal{T}$ систему подмножеств, являющихся дополнениями подмножеств, входящих в $\mathcal{F}$. Как перефразировть в терминах свойств $\mathcal{F}$ те свойства $\mathcal{T}$, при наличии которых $\mathcal{T}$ является топологией? (Иными словами, когда, объявив какие-то подмножества "замкнутыми", мы действительно получим топологию, в которой именно эти множества будут замкнутыми?)

Упражнение. Докажите, что:

а). $\mathrm{clos}\ \emptyset = \emptyset$.

б). $A \subset \mathrm{clos}\ A$ (подразумевается —для любого $A \subset M$).

множества A (формулировка годится и для топологических пространств).

[17]Слово "прикосновение" наводит на мысль о какой-то связи с касательными или касанием, но это словесное сходство обманчиво.

в). $\operatorname{clos}(A \cup B) = \operatorname{clos} A \cup \operatorname{clos} B$. Верно ли, что $\operatorname{clos}(A \cap B) = \operatorname{clos} A \cap \operatorname{clos} B$?

г). $\operatorname{clos}(\operatorname{clos} A) = \operatorname{clos} A$.

д). $A \subset B \Longrightarrow \operatorname{clos} A \subset \operatorname{clos} B$, причём д) можно получить как следствие из в) и г).

Упражнение. Напрашивается вопрос, нельзя ли определить понятие топологического пространства, исходя из операции замыкания подмножеств? Можно (хотя определение, исходящее из системы открытых множеств, всё-таки кажется более удобным и чаще отправляются именно от него). Пусть каждому подмножеству $A \subset M$ сопоставлено подмножество $\overline{A}$ (т.е. задано отображение множества $\mathcal{P}$ всех подмножеств множества M в себя, причём образ подмножества A обозначается через $\overline{A}$), и пусть при этом выполняются условия , аналогичные условиям а) - г) из предыдущего упражнения, т.е.

$$\overline{\emptyset} = \emptyset, \quad A \subset \overline{A}, \quad \overline{A \cup B} = \overline{A} \cup \overline{B}, \quad \overline{\overline{A}} = \overline{A}.$$

Тогда в M имеется единственная топология $\mathcal{T}$, для которой замыкание $\operatorname{clos} A$ любого подмножества $A \subset M$ совпадает с $\overline{A}$.

Точка $a \in M$ называется внутренней точкой подмножества $A \subset M$, если у неё имеется окрестность, целиком содержащаяся в A. Внутренность или открытое ядро подмножества $A \subset M$ — это множество всех внутренних точек этого подмножества. Она обозначается через $\operatorname{int} A$[18] (от interiour, ср. с "интерьер".) Точка $a \in M$ называпется граничной точкой подмножества $A \subset M$, если во всякой её окрестности имеется как точка из A, так и точка из дополнения $\mathbf{C}A = M \setminus A$. Граница подмножества $A \subset M$ — это множество всех его граничных точек. Она обозначается через $\operatorname{Fr} A$ [19] (от frontier).

Упражнение. Докажите, что:

а). $\operatorname{clos} A = A \cup \operatorname{Bd} A$, $\operatorname{int} A = A \setminus \operatorname{Bd} A$, $\operatorname{Fr} A = \operatorname{clos} A \cap \operatorname{clos} \mathbf{C}A$, $\operatorname{clos} A = \mathbf{C} \operatorname{int} \mathbf{C}A = M \setminus \operatorname{int} \mathbf{C}A$, $\operatorname{int} A = \mathbf{C} \operatorname{clos} \mathbf{C}A$.

б). Открытое множество — это множество, совпадающее со своей внутренностью.

в). В общем случае $\operatorname{int} A$ — это наибольшее открытое множество, содержащееся в A.

[18] Употребляют также обозначение A^O.

[19] А также через $\operatorname{Bd} A$ (от boundary) и ∂A (происхождение последнего обозначения потребовало бы слишком длительных разъяснений).

Простой рисунок на плоскости — замкнутая кривая (например, окружность), ограничивающая некоторую область, даёт наглядное представление о всех этих понятиях. Однако надо предупредить, что уже на плоскости граница области может быть устроена гораздо сложнее. Самой простой книгой, где имеются такие примеры, является книжка

Н.Я.Виленкин. Рассказы о множествах. М.: МЦНМО, 2003.

В своё время соответствующие примеры возникли при продумывании точного смысла тех или иных понятий. Позднее более или менее сходные примеры стали появляться сперва в самой математике при исследовании совсем других (не связанных с анализом смысла понятий) вопросов, а затем и в некоторых вопросах математического естествознания.

Упражнение. Пусть $(X, \mathcal{T})$ — топологическое, а (M, d) — метрическое пространство. Докажите, что пересечение $BC(X, M)$ введённого ранее пространства “ограниченных” отображений $X \to M$ с множеством $C(X, M)$ непрерывных отображений $X \to M$ являеся замкнутым подмножеством в пространстве $B(X, M)$ (которое, как и раньше, снабжается метрикой D) и потому полно в индуцированной метрике $D_{BC(X,M)}$.

В этом упражнении фигурируют две различные структуры — топология в X и метрика в M. Они играют различные роли — топология позволяет говорить о непрерывных отображениях $X \to M$ (при этом вместо метрики в M нам хватило бы и топологии), а метрика в M — о равномерной сходимости отображений $X \to M$ (при этом не используется никакой структуры в X).

Относительные понятия.

Представим себе, что мы сосредоточили внимание на некотором множестве $A \subset M$, которое на какое-то время образует для нас как бы “весь наш мир”. Тогда полезно ввести относительные понятия, которые связаны именно с A и при обращении к которым мы как бы забываем о существовании объемлющего пространства M. Когда M — метрическое пространство с метрикой d, всё сводится к тому, что A рассматривается как метрическое пространство с метрикой d_A, индуцированной на A метрикой d. Все те понятия, о которых мы говорили, можно рассматривать применительно к (A, d_A). При этом, хотя мы и говорим, что A — это “весь наш мир”, всё же стоит посмотреть, какое отношение “события в этом мире” имеют к “событиям во всём M”, т.е. как в терминах (M, d) описываются, скажем, открытые подмножества метрического пространства (A, d_A).

Упражнение. Покажите, что открытые подмножества топологического пространства (A, d_A) — это пересечения $A \cap U$ множества A со всевозможными открытыми подмножествами объемлющего пространства (M, d).

Как видно, в полученном описании топологии $\mathcal{T}_A$ в A, отвечающей метрике d_A, сама эта метрика (как и метрика d в M) явно не участвует — в описании говорится только об A и о топологии $\mathcal{T}$ в M (другое дело, что она-то в данном случае определяется метрикой d). Отправлясь отсюда, мы и в общем случае топологического пространства $(M, \mathcal{T})$ определим индуцированную или относительную топологию $\mathcal{T}_A$ в A как совокупность пересечений $A \cap U$ множества A со всевозможными открытыми подмножествами $U \in \mathcal{T}$. Такие пересечения называют относительно открытыми подмножествами множества A.

Упражнение. Проверьте, что $\mathcal{T}_A$ — действительно топология.

Упражнение. Докажите, что:

а). Окрестность (открытая окрестность) точки $a \in A$ в $(A, \mathcal{T}_A)$ — это пересечение с A окрестности (соответственно, открытой окрестности) этой точки в объемлющем пространстве $(M, \mathcal{T})$.

б). Замкнутые подмножества пространства $(A, \mathcal{T}_A)$ суть пересечения $A \cap F$ множества A со всевозможными замкнутыми подмножествами во всём M. (Такие пересечения, естественно, называют относительно замкнутыми подмножествами множества A.)

в). Замыкание $\mathrm{clos}_A B$ подмножества B пространства $(A, \mathcal{T}_A)$ — это пересечения с A замыкания $\mathrm{clos}\, B$ этого подмножества во всём M.

Компактность.

В самом начале курса матанализа учащийся узнаёт о некоторых особых свойствах конечных замкнутых отрезков $[a, b]$, отличающих их от открытых (или полуоткрытых) интервалов или от бесконечных отрезков (безразлично, бесконечен ли он в обе стороны (так что речь идёт обо всей числовой прямой) или только в одну, а в последнем случае — содержит ли отрезок ограничивающее его с одной стороны число или нет). Всякая бесконечная последовательности точек отрезка $[a, b]$ содержит сходящуюся подпоследовательность; всякое бесконечное подмножество отрезка имеет предельную точку; из всякого покрытия отрезка открытыми интервалами можно выбрать конечное подпокрытие (лемма Гейне–Бореля–Лебега); функция f, непрерывная на таком отрезке, равномерно непрерывна на нём; она ограничена, так что её верхняя и нижняя грани $\sup_{x\in[a,b]} f(x)$ и $\inf_{x\in[a,b]} f(x)$ конечны; обе они достигаются (т.е. имеются такие точки $c_{\max}$ и $c_{\min}$, что $f(c_{\max}) = \sup_{x\in[a,b]} f(x)$ и $f(c_{\min}) = \inf_{x\in[a,b]} f(x)$; иными словами, f принимает где-то на отрезке наибольшее и наименьшее значения). Позднее учащийся узнаёт, что аналогичными свойствами обладает всякое замкнутое ограниченное подмножество $\mathbb{R}^n$ (при этом в лемме Гейне-Бореля-Лебега надо говорить об открытом покрытии, т.е. покрытии открытыми множествами). В общем же случае замкнутое ограниченное подмножество (бесконечномерного) банахова пространства или, тем более, метрического пространства[20] может этими свойствами не обладать. Но может и обладать.

Те метрические пространства, которые обладают перечисленными свойствами, играют столь заметную роль, что заслужили особого названия[21]. Метрическое пространство (M, d) называется компактным, если всякая бесконечная последовательности его точек отрезка содержит сходящуюся подпоследовательность. Такое пространство обладает прочими свойствами, указанными выше для замкнутого отрезка. Повторяю: всякое бесконечное подмножество пространства M имеет предельную точку; из всякого открытого покрытия пространства M (т.е. покрытия его открытыми подмножествами) можно выбрать конечное подпокрытие; функция f, непрерывная на M, равномерно непрерывна на этом пространстве (т.е.

$$\forall\varepsilon > 0\ \exists\delta > 0 \qquad d(x, y) < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon);$$

она ограничена, так что её верхняя и нижняя грани $\sup_{x\in M} f(x)$ и $\inf_{x\in M} f(x)$ конечны; обе они достигаются (т.е. имеются такие точки $c_{\max}$ и $c_{\min}$, что $f(c_{\max}) = \sup_{x\in M} f(x)$ и $f(c_{\min}) = \inf_{x\in M} f(x)$; иными словами, f принимает где-то на M наибольшее и наименьшее значения).

Поскольку нередко пространства возникают как подмножества других пространств, то стоит уделить некоторое внимание компактным подмножествам метрических пространств и некоторым подмножествам родственного характера. Подмножество $A \subset (M, d)$ называется относительно компактным, если его замыкание компактно. ε-сетью подмножества

[20] Ограниченность подмножества $A \subset (M, d)$ естественно понимать в том смысле, что для какой-то (а тогда и любой, — почему?) точки $c \in M$ расстояния $d(c, x)$ от c до всех точек $x \in A$ ограничены сверху одним и тем же числом C (вообще говоря, зависящим от c).

[21] Это относятся и к подмножествам метрических пространств, поскольку эти подмножества можно рассматривать как самостоятельные метрические пространства с индуцированной метрикой.

$A \subset (M, d)$ называется такое подмножество $S \subset M$, что любая точка из A находится на расстоянии $< \varepsilon$ от какой-то точки из S, т.е. что $A \subset U_r(S)$. Подмножество $A \subset (M, d)$ называется вполне ограниченным, если для любого $\varepsilon > 0$ у него имеется конечная ε-сеть. Формально это определение зависит не только от (A, d_A), но и от объемлющего пространства (M, d), — ведь S, вообще говоря, является подмножеством последнего. Однако из-за того, что в нём говорится обо всех $\varepsilon > 0$, ничего не изменится, если дополнительно потребовать, чтобы ε-сеть S была подмножеством самого A. Действительно, если $S \subset M$ является конечной $\frac{\varepsilon}{2}$-сетью для $A \subset M$, то, удалив из S "лишние" точки s — те, которые от любой из точек A отстоят не менее чем на $\frac{\varepsilon}{2}$, — и заменив каждую из оставшихся точек $s \in S$ на какую-нибудь точку a_s из A, для которой $d(s, a_s) < \frac{\varepsilon}{2}$, мы получим конечную ε-сеть $\{a_s\}$ для A, состоящую из точек A (проверьте!). Оказывается, что если подмножество $A \subset (M, d)$ относительно компактно, то оно является вполне ограниченным, а если при этом M полно, то и обратно, вполне полное $A \subset (M, d)$ относительно компактно.

Мы говорили о компактных метрических пространствах. Понятие компактности имеет смысл и для топологических пространств, но оказывается, что "хорошее" свойство таковых получается, если говорить не о сходящихся подпоследовательностях, а о свойстве, обобщающем заключение леммы Гейне-Бореля-Лебега. Топологическое пространство называется компактным, если у любого открытого покрытия этого пространства имеется конечное подпокрытие. Для метрических пространств такое определение оказывается эквивалентным первоначальному, а для топологических — нет. (Первоначально была попытка определять компактность топологического пространства так же, как и в метрическом случае, но в топологическом случае такое свойство оказалось малоинтересным. Однако тем самым название "компактность" оказалось занятым, поэтому хорошее свойство компактности (определяемое через покрытия) назвали иначе — бикомпактностью. Позднее же решили, махнув рукой на прежнего неудачного кандидата, переименовать бикомпактность в компактность.) Те из перечисленных выше свойств компактных метрических пространств, которые не связаны с метрикой (т.е. в которых не говорится о равномерности, полной ограниченности и полноте), присущи и компактным топологическим пространствам.

База топологии.

Условие, определяющее, когда множество является открытым, может формулироваться довольно просто. (См., например, определение открытых подмножеств метрических пространств.) Однако совокупность всех открытых множеств почти во всех мало-мальски содержательных примерах совершенно не обозрима (достаточно подумать о разнообразии открытых подмножеств плоскости) [22]. Зато в топологии $\mathcal{T}$ может иметься такая система $\mathcal{B} \subset \mathcal{T}$ открытых множеств, что каждое открытое множество является объединением множеств из $\mathcal{B}$ — такую систему $\mathcal{B} \subset \mathcal{T}$ называют базой топологии $\mathcal{T}$ (иногда, особенно в более старывх книгах, говорят не о базе, а о базисе), — и в то же время система $\mathcal{B}$ вполне обозрима. (Таким образом, элементы базы — это как бы "стандартные строительные блоки", из которых "собираются" открытые множества, только, в отличие от настоящего строительства, наши "блоки" могут не только примыкать друг к другу, но и перекрываться. Имея блоки всего нескольких типов, можно соорудить из них массу разнообразных построек.)

Например, в метрическом пространстве совокупность открытых шаров со всевозможными центрами и радиусами образует базу соответствующей топологии. Иногда её можно ещё более уменьшить.

Упражнение. Пространство называется сепарабельным, если в нём имеется всюду плотное[23] счётное подмножество. Докажите, что в сепа-

[22]В какой-то степени поддаются единообразному описанию открытые подмножества прямой $\mathbb{R}$. Помимо самой $\mathbb{R}$, открытые подмножества $A \subset \mathbb{R}$ единственным образом представляются как объединения не более чем счётных совокупностей непересекающихся открытых интервалов (a_n, b_n) (один или два из них могут быть полубесконечными). Такое описание выглядит вполне удовлетворительным, однако систематическое обозрение всех открытых подмножеств $A \subset \mathbb{R}$ требовало бы систематического обозрения всех возможностей, которые могут представиться для взаимного расположения этих интервалов, что едва ли можно сделать совсем уж эффективным образом. (Возьмём в каждом интервале (a_n, b_n) какую-нибудь точку c_n, — например, середину интервала, если он конечен. Множество $C = \{c_n\}$ всех этих точек — счётное подмножество $\mathbb{R}$, все точки которого являются изолированными точками этого C. Описание всевозможных множеств C такого типа фактически должно было бы быть существенной частью описания открытых подмножеств $\mathbb{R}$. Уже эта часть сомпительна в отношении того, какой эффективности здесь можно достичь.)

[23]Я уже употреблял этот термин, пояснив (в частном случае), что подмножество A пространства M называется всюду плотным, если сколь угодно близко к любой точке $x \in M$ имеется точка из A. Если понимать "сколь угодно близко к x" как "в любой окрестности точки x", то такая формулировка годится не только для метрических, но и для топологических пространств. Эквивалентные условия: в любом открытом

рабельном метрическом пространстве со всюду плотным подмножеством A открытые шары с центрами из A и радиусами $\frac{1}{n}$ (со всевозможными натуральными n) образуют базу топологии. Значит, в этом случае топология имеет счётную базу.

Покажите также, что если топология имеет счётную базу, то пространство (даже топологическое) сепарабельно.

Для более общих топологических пространств сепарабельность не обязательно влечёт за собой существование счётной базы топологии. Оказывается, что из этих двух свойств — сепарабельности и существования счётной базы — второе является намного более важным. Когда оно нужно, приходится так и говорить: "Пусть топологическое пространство M имеет счётную базу ...". Для метрических же пространств обычно в подобных случаях в формулировках говорят о сепарабельности (хотя скорее всего в конечном счёте нужна счётная база) — это как-то нагляднее (и проверяется более непосредственно). К счастью, нам чаще всего нужны метрические (и даже банаховы) пространства.

Упражнение. Проверьте сепарабельность пространства $C[0,1]$ (как уже говорилось, его элементы — непрерывные функции $f : [0,1] \to \mathbb{R}$, а норма $\|f\| = \sup_{[0,1]} |f(x)|$).

На роль понятия базы можно посмотреть ещё и так: чтобы задать топологию $\mathcal{T}$, можно задать какую-нибудь её базу. Так, определение открытых подмножеств метрического пространства равносильно тому, что открытые шары принимаются за базу соответствующей топологии. Позднее мы встретимся с примерами, когда с самого начала задаётся база будущей топологии (и это не связано с метрикой). Понятно, что при таком образе действий полезно знать, какими свойствами должна обладать система подмножеств $\mathcal{B}$ множества M, чтобы эта система была базой какой-то топологии. Такие условия должны выражать (в терминах самой системы множеств $\mathcal{B}$) тот факт, что если мы возьмём всевозможные объединения множеств из $\mathcal{B}$[24], то полученная совокупность множеств $\mathcal{T}$ будет топологией.

Упражнение. Покажите, что эти условия сводятся к следующим двум: пересечение двух множеств из $\mathcal{B}$ является объединением каких-то

множестве имеется точка из A; замыкание A всюду плотно. (Проверьте!).

[24]Эта формулировка предусматривает и объединение пустой системы множеств из $\mathcal{B}$, которое пусто. Если читателю покажется, что такое умозаключение отдаёт иезуитством, то придётся честно сказать, что помимо объединений множеств из $\mathcal{B}$ мы включаем в $\mathcal{T}$ также и пустое множество.

множеств из $\mathcal{B}$; объединение всех множеств из $\mathcal{B}$ совпадает с M.

Упражнение. а). Пусть $(M_1, \mathcal{T}_1)$ и $(M_2, \mathcal{T}_2)$ — топологические пространства. Докажите, что система множеств $\{U \times V;\ U \in \mathcal{T}_1, V \in \mathcal{T}_2\}$ является базой некоторой топологии $\mathcal{T}$ в $M_1 \times M_2$. Говорят, что $(M_1 \times M_2, \mathcal{T})$ является прямым произведением исходных топологических пространств. Докажите также, что если $\mathcal{B}_i$ — какие-нибудь базы топологий $\mathcal{T}_i$, то $\{U \times V;\ U \in \mathcal{B}_1, V \in \mathcal{B}_2\}$ — база указанной топологии $\mathcal{T}$.

б). Покажите, что если (M_i, d_i) — метрические пространства и $\mathcal{T}_i$ — соответствующие топологии, то в прямом произведении метрических пространств $(M_1 \times M_2, D)$ (где D — какая-нибудь из эквивалентных метрик, вводимых в $(M_1 \times M_2)$ согласно сказанному выше, а $\mathcal{T}$ — определяемая этой метрикой топология, в $(M_1 \times M_2)$, то топологическое пространство $(M_1 \times M_2, \mathcal{T})$ является прямым произведением топологических пространств $(M_i, \mathcal{T}_i)$.

в). Проверьте, что прямое произведение двух хаусдорфовых пространств хаусдорфово, а двух компактных пространств — компактно.

Родственным понятием является база (или, как ещё говорят, фундаментальная система) окрестностей точки x_0 топологического пространства $(E, \mathcal{T})$. Это такая система $\mathcal{B}$ окрестностей точки x_0, что в каждой окрестности содержится некоторая окрестность из $\mathcal{B}$.

Упражнение. а). Докажите, что последовательность x_n тогда и только тогда сходится к x_0, когда для любой окрестности $U \in \mathcal{B}$ (где $\mathcal{B}$ — попрежнему база окрестностей точки x_0) имеется такое N, что $x_n \in U$ при всех $n > N$.

б). Пусть $(F, \mathcal{T}_1)$ — другое топологическое пространство, $\mathcal{B}$ — некоторая база окрестностей точки $x_0 \in E$, $\mathcal{B}_1$ — некоторая база окрестностей точки $y_0 \in F$, и пусть задано отображение $f : E \to F$. причём $f(x_0) = y_0$. Покажите, что непрерывность этого отображения в точке x_0 равносильна тому, что

$$\forall V \in \mathcal{B}_1\ \exists U \in \mathcal{B} \quad f(U) \subset V.$$

Топология Шварца – Уитни.

В основном варианте теории обобщённых функций Л.Шварца эти функции вводятся как непрерывные однородные линейные функционалы на некотором векторном топологическом пространстве $\mathcal{D}$, которое называется пространством “основных” (а иногда — “пробных”) функций. До сих пор мы не вводили понятия векторного топологического пространства. Это, конечно, векторное пространство E, в котором введена также хаусдорфова топология, причём требуется, чтобы алгебраические операции

$$\mathbb{R} \times E \text{ или } \mathbb{C} \times E \to E \qquad (\lambda, x) \mapsto \lambda x$$

и

$$E \times E \to E \qquad (x, y) \mapsto x + y$$

были непрерывными. Имеется также дополнительное свойство “локальной выпуклости”, которое выполняется в обычных примерах и без которого свойства E могут быть “патологическими”. Оно гласит: любая окрестногсть нуля (а тогда и любой другой точки) содержит некоторую выпуклую[25] окрестность нуля.

Естественно, что для теории обобщённых функций нужны различные свойства пространства $\mathcal{D}$, но сейчас нас будет интересовать только топология, поэтому на взаимосвязь топологии и структуры векторного пространства можно пока не обращать серьёзного внимания. Пока что для нас $\mathcal{D}$ будет примером довольно необычного топологического пространства. Однако стоит сразу принять к сведению, что это не просто пример, придуманный для демонстрации того, какой может быть топология, но пример, связанный с важной научной теорией, причём важность этой теории в основном основана на её использовании в теории уравнений в частных производных и некоторых других задачах математической физики (квантовая теория поля), а вовсе не с топологией (хотя кое-что касается и топологии). Читатель, вероятно, об этом слышал, а если не слышал, то прошу принять это на веру.

Я буду рассматривать только частный случай основных (а со временем и обобщённых) функций на прямой $\mathbb{R}$. Но всё, о чём будет идти речь вначале, автоматически переносится на случай функций, заданных на пространстве $\mathbb{R}^n$ или в некоторой области этого пространства.

[25]Подмножество векторного пространства называется выпуклым, если для любых двух его точек соединяющий их прямолинейный отрезок целиком содержится в этом множестве.

Основная функция — это бесконечно дифференцируемая функция f с компактным "носителем" $\text{supp}\, f = \text{clos}\,\{x;\ f(x) \neq 0\}$. Проще говоря, это функция, тождественно равная нулю вне некоторого (своего для каждой функции) конечного отрезка. Ясно, что такие функции образуют векторное пространство, которое и обозначается через $\mathcal{D}$. Но не так уж ясно, какую топологию стоит ввести в $\mathcal{D}$, да и вообще можно ли это сделать "разумным" образом. При наиболее "очевидных" кандидатах в топологию [26] скорее всего получится, что $\mathcal{D}$ неполно — в менее формальных терминах это означает, что предел последовательности функций с компактными носителями может иметь смысл как функция и даже быть бесконечно дифференцируемой фуцнкцей, но иметь некомпактный носитель.

Топология, которую ввёл Л.Шварц ввёл в $\mathcal{D}$, совершенно аналогична топологии, которую в некоторых (других) функциональных пространствах для других целей (относящихся к дифференциальной топологии) ввёл Х.Уитни. Я не знаю, кто из них был первым и были ли их топологии предложены независимо. Вначале они вполне могли не знать друг о друге. Однако странно, что и сейчас, спустя много лет, в известных мне учебниках говорят либо о топологии Шварца в $\mathcal{D}$ (в связи с теорией обобщённых функций), либо о топологии Уитни (в других функциональных пространствах и в другой связи) и не упоминают об их совпадении.

Предупредив, что "Шварц $=$ Уитни", я далее буду говорить о $\mathcal{D}$, ибо, во-первых, как пример это чуть проще (быстрее описывается функциональное пространство) и, во-вторых, в перспективе я надеюсь немного остановиться на обобщённых функциях, а не на задачах дифференциальной топологии. Помимо $\mathcal{D}$, мы рассмотрим ещё два аналогичных, но более простых пространства (см. ниже).

Сперва я скажу, что понимается у Шварца под пределом f последовательности $\{f_n\}$ основных функций. По определению, основная функция f является таким пределом, если все f_n тождественно равны нулю вне некоторого (общего для них) отрезка $[a, b]$ и если последовательность f_n равномерно сходится к f вместе со всеми своими производными, т.е.

[26] Их, кажется, два — топология равномерной сходимости на прямой $\mathbb{R}^n$ последовательности функций f_n вместе со всеми последовательностями их производных $f_n^{(i)}$ и топология равномерной сходимости этих же последовательностей на всевозможных замкнутых отрезках.

каждая последовательность производных $\{f_n^i\}$

$$f_n^{(i)} \rightrightarrows f^{(i)} \quad \text{при } i = 0, 1, 2, \ldots.$$

(Знак $\rightrightarrows$ обозначает равномерную сходимость. Сам по себе этот знак не содержит указания, на каком именно множестве сходимость равномерна. В данном случае имеется в виду равномерная сходимость на $\mathbb{R}$ или (что сейчас равносильно) равномерная сходимость на $[a, b]$.) В приведённом определении никакая топология в E не используется, просто мы сказали, что такие-то последовательности отныне называются сходящимися и что такие-то функции называются их пределами. Назовём на минуту это определение "отдельным определением сходимости". Раз оно не опирается на то, что было установлено для метрических или топологических пространств, то самые привычные свойства сходящихся последовательностей надо проверять заново. Заново надо определять и различные понятия, связанные с непрерывностью, например, ключевое понятие непрерывного однородного линейного функционала, а затем опять-таки надо проверять, что соответствующие объекты имеют надлежащие свойства. Всё это довольно просто, по крайней мере на начальных шагах, но требует времени и места. Возникают два вопроса:

— существует ли такая метрика на E, что сходимость, введённая выше посредством отдельного определения, совпадает со сходимостью в этой метрике?

— если такой метрики нет, то не существует ли такой топологии, что введённая выше сходимость совпадает со сходимостью в этой топологии?

Как мы увидим, ответ на первый вопрос отрицательный, а на второй — положительный. Тем самым заодно мы получим пример топологии, которая не определяется никакой метрикой (неметризуемая топология, неметризуемое пространство), причём это не специально придуманный искусственный пример, а пример, связанный со столь важной теорией, как теория обобщённых функций.

Топология, определяемая метрикой (метризуемая топология) обладает тем свойством, что у каждой точки имеется счётная база (счётная фундаментальная система) окрестностей нуля. Действительно, такой системой для точки x метрического пространства (M, d) является система открытых шаров $U_{\frac{1}{n}}(x)$ (почему?). Мы увидим, что не существует такой топологии, в которой точки E имели бы счётную базу окрестностей, а сходимость была бы такой, как в отдельном определении.

Мы облегчим себе задачу, рассмотрев аналогичную задачу для других пространств, для которых ситуация принципиально такая же, но с которыми иметь дело несколько проще. Во-первых, откажемся от бесконечной дифференцируемости “основных” функций и рассмотрим вевторное пространство $\mathcal{K}(\mathbb{R})$, состоящее из непрерывных функций, каждая из которых имеет компактный носитель. Сходимость в $\mathcal{K}(\mathbb{R})$ определим так: последовательность $\{f_n\}$ сходится к f, если носители всех f_n заключены в некотором конечном отрезке $[a, b]$ и если $f_n \rightrightarrows f$ (обычная равномерная сходимость, безразлично, на всей ли прямой $\mathbb{R}$ или на $[a, b]$). Перейдя к $\mathcal{K}(\mathbb{R})$, мы избавляемся от необходимости думать о производных. Во-вторых, вместо непрерывных функций на $\mathbb{R}$ рассматрим ещё более простые функции на множестве $\mathbb{N}$ натуральных чисел, для каждой из которых носитель $\{k;\ f(k) \neq 0\}$ конечен. Примем, что последовательность таких функций $\{f_n\}$ сходится к f, если носители всех f_n заключены в некотором конечном множестве, т.е. если имеется такое K, что $f_n(k) = 0$ при всех $k \geqslant K$ и всех $n \in \mathbb{N}$, и если $f_n(k) \to f(k)$ при всех k (что содержательно только при $k < K$). Перейдя к пространству $\mathcal{K}(\mathbb{N})$ таких функций, мы избавляемся от необходимости думать о непрерывности.

Упражнение. Докажите, что если в $\mathcal{K}(\mathbb{N})$ введена топология $\mathcal{T}$, в которой точка 0 (т.е. функция $\mathbb{N} \to \mathbb{R}$, тождественно равная нулю) имеет счётную базу окрестностей, то сходимость в этой топологии не совпадает с сходимостью в смысле отдельного определения. Для этого:

а). Покажите, что если точка 0 имеет счётную базу окрестностей, то существует и такая счётная база окрестностей $\{U_n\}$ этой точки, что $U_1 \supset U_2 \supset \ldots \supset U_n \supset U_{n+1} \supset \ldots$.

б). Пусть сходимость в какой-то топологии $\mathcal{T}$ пространства $\mathcal{K}(\mathbb{N})$ равносильна сходимости в смысле отдельного определения. Покажите, что если U — окрестность нуля в топологии $\mathcal{T}$, то для любого $n \in \mathbb{N}$ в U найдётся такая функция f, что $f(n) \neq 0$. (Указание: рассмотрите последовательность функций $\{f_k\}$, у которых $f_k(n) = \frac{1}{k}$, а остальные $f_k(m) = 0$.)

б). Докажите, что существует такая последовательность функций $\{f_k\}$ что при всех n

$$f_n \in U_n \quad \text{и } f_n(n) \neq 0.$$

Выведите отсюда, что $f_k \not\to 0$ в смысле отдельного определения, а в то же время $f_k \to 0$ в смысле сходимости в топологии $\mathcal{T}$.

Но, может быть, в $\mathcal{K}(\mathbb{N})$ вообще нет никакой топологии, сходимость в которой совпадала бы со сходимостью в смысле отдельного определения?

Л.Шварц и Х.Уитни обнаружили, что такая топология существует. В ней некоторая база окрестностей нуля $\mathcal{B}_0$ задаётся следующим способом: берётся последовательность положительных чисел $\{\varepsilon_n\}$ и объявляется, что множества $U_{\varepsilon_n}(0) = \{f; |f(x)| < \varepsilon_n$ при всех $n\}$ со всевозможными $\{\varepsilon_n\}$ суть элементы $\mathcal{B}_0$, причём эти окрестности нуля считаются открытыми. (Пока что название "открытая окрестность" — это тоже "отдельное определение".) Базой $\mathcal{B}_f$ окрестностей (причём открытых) другой точки f объявляется система множеств $U_{\varepsilon_n}(0) + f$, где $A + f$ для подмножества $A \subset \mathcal{K}(\mathbb{N})$ означает $\{a + f;\ a \in A\}$ (иными словами, база окрестностей точки f состоит из множеств вида $\{g;\ g - f \in U_{\varepsilon_n}(0))$. Вот теперь мы можем заговорить о топологии, объявив, что открытым называется такое множество, которые вместе с каждой своей точкой f содержат некоторое множество, являющееся элементом $\mathcal{B}_f$.

Упражнение. Проверьте, что

а) тем самым действительно определяется некоторая хаусдорфова топология $\mathcal{T}$ в $\mathcal{K}(\mathbb{N})$;

б) $(\mathcal{K}(\mathbb{N}), \mathcal{T})$ является локально выпуклым топологическим векторным пространством;

в) сходимость $f_n \to f$ в топологии $\mathcal{T}$ совпадает со сходимостью в смысле отдельного определения.

Объявим, что в $\mathcal{K}(\mathbb{R})$ нулевой элемент 0 имеет базу (открытых) окрестностей $\mathcal{B}_0$, каждая из которых задаётся с помощью заданной на $\mathbb{R}$ непрерывной положительной функции $\varepsilon(t)$, а именно, $U_{\varepsilon(\,.\,)}(0) = \{f \in \mathcal{K}(\mathbb{R});$ $\forall t\ |f(t)| < \varepsilon(t)\}$. Базой (открытых) окрестностей $\mathcal{B}_f$ элемента $f \in \mathcal{K}(\mathbb{R})$ объявим $\mathcal{B}_0 + f$, а открытыми подмножествами $\mathcal{K}(\mathbb{R})$ объявим те подмножества, которые вместе с каждым своим элементом f содержат некоторую "окрестность" из $\mathcal{B}_f$.

Упражнение — аналог для $\mathcal{K}(\mathbb{R})$ предыдущего упражнения.

Взглянем на эту топологию с такой точки зрения. Поскольку никаких условий на $\varepsilon(t)$ не наложено (кроме положительности и непрерывности), то $\varepsilon(t)$ может сколь угодно быстро убывать при $|t| \to \infty$. Характерной особенностью соответствующей окрестности нуля $U_{\varepsilon(\,.\,)}(0)$ является то, что по мере увеличения $|t|$ условие $|f(t)| < \varepsilon(t)$ становится всё более ограничительным. Это замечание объясняет идею топологии Шварца – Уитни в пространстве $\mathcal{D}$. В нём условия, конечно, должны налагаться не только на функцию f, но и на её производные. "Увеличение ограничительности условия при увеличении $|t|$" в этом случае состоит также и в том, что в каждом конечном отрезке условия налагаются только на несколько производных, но по мере увеличения $|t|$ число “контролируемых” производных возрастает. Можно сказать, например, так. Помимо $\varepsilon(t)$ задают ещё неубывающую функцию $m : \mathbb{R}_+ \to \mathbb{N}$ и полагают

$$U_{\varepsilon(\,.\,),m(\,.\,)} = \{f \in \mathcal{D};\ |f^{(i)}(t)| < \varepsilon(t) \ \text{ при } i \leqslant m(|t|).\}$$

Упражнение — закончите определение топологии в $\mathcal{D}$ и докажите утверждения, аналогичные содержащимся в последних двух упражнениях. Проверьте также, что в этой топологии у точки 0 имеется ещё и следующая база окрестностей. Принадлежащие ей окрестности “нумеруются” при помощи невозрастающих последовательностей положительных чисел $\{\varepsilon_n\}$ и неубывающих последовательностей неотрицательных целых чисел $\{m_n\}$, а именно, паре таких последовательностей сопоставляется окрестность

$$U_{\{\varepsilon_n\},\{m_n\}} = \{f \in \mathcal{D};\ |f^i(x)| < \varepsilon_n \text{ при } |x| > n, i \leqslant m_n.\}$$

Связность.

Первоначальное “интуитивное” и наглядное представление о связности таково: фигура несвязна, если она, подобно букве Ы, состоит из двух отдельных, не связанных между собой частей; в противном случае фигура называется связной. Несомненно, буква И — связная фигура. Но является ли связной фигура F (если её можно назвать фигурой), состоящая из графика функции $y = \sin\frac{1}{x},\ x > 0$, и вертикального отрезка $x = 0,\ -1 \leqslant y \leqslant 1$? Эти две части — график и вертикальный отрезок — не так сильно отделены друг от друга, как части Ь и I фигуры Ы: вертикальный отрезок лежит в замыкании графика. И всё же это вроде бы отдельные части, во всяком случае, двигаясь по фигуре F, нельзя перейти из графика в вертикальный отрезок. Можно сделать вывод, что фигура F менее связна, чем буква И, но более связна, чем буква Ы; получается, что связность (в отличие от свежести, — впрочем, это мнение Воланда) бывает разных сортов. Идея одной “разновидности” связности — та, с которой мы начали (“две отдельные части”), идея другой — от одной точки фигуры можно перейти к любой другой точке, не выходя из самой фигуры. Вначале может показаться, что это одно и то же, но пример фигуры F показывает, что это не так.

Формализация этих идей приводит к следующим определениям.

а). Пространство M (здесь можно говорить о топологическом пространстве, но обсуждение оказывается содержательным уже уже для подмножеств $\mathbb{R}^n$; при этом фигурирующие ниже открытые и замкнутые подмножества в M надо понимать как подмножества, открытые и замкнутые в индуцированной топологии) открытые и замкнутые) называется несвязным, если его можно представить в виде объединения двух непересекающихся непустых открытых подмножеств A и B. "Открытых" здесь можно заменить на "замкнутых". Более того, если такое представление возможно, то соответствующее A является открыто-замкнутым (т.е. и открытым, и замкнутым) подмножеством пространства M, не совпадающим ни с $\emptyset$, ни с M; обратно, если существует открыто-замкнутое $A \subset M$, $A \neq \emptyset, M$, то M несвязно. Если M — подмножество бóльшего пространства N, то (как уже говорилось применительно к $N = \mathbb{R}^n$) M рассматривается как "самостоятельное" пространство с индуцированной топологией. Но можно дать и формулировку, в которой не прибегают к относительным понятиям: M несвязно, если M можно представить в виде объединения таких двух непустых его подмножеств A и B, каждое из которых не содержит ни одной точки прикосновения другого, т.е.

$$M = A \cup B, \quad A \cap \operatorname{clos} B = \emptyset, \quad \operatorname{clos} A \cap B = \emptyset, \qquad A, B \neq \emptyset$$

(замыкания здесь берутся в объемлющем пространстве N). Если такое представление невозможно, то пространство M (или подмножество M пространства N) называется связным.

Например, отрезок $[a, b]$ связен. Действительно, пусть $[a, b] = A \cup B$, где A, B — непересекающиеся открыто-замкнутые подмножества $[a, b]$ и, скажем, $a \in A$ (последнее ведь зависит просто от выбора обозначений для двух открыто-замкнутых частей $[a, b]$). Докажем, что тогда $A = [a, b]$. Допустим, что $B \neq \emptyset$, и обозначим $c := \inf B$. Будучи замкнутым подмножеством замкнутого отрезка, B является ограниченным замкнутым подмножеством $\mathbb{R}$, и потому $c \in B$. Далее, все точки x, лежащие строго левее c на отрезке $[0, 1]$, должны принадлежать A; а такие точки существуют, ибо из $a \notin B$ следует, что $a \neq c$ и, значит, $a < c$. Итак, $[a, c)$ — непустое подмножество A. Но тогда, ввиду замкнутости A, $c \in A$, и мы приходим к противоречию.

Аналогичным образом доказывается связность открытого интервала (a, b), открытого с одной стороны и замкнутого с другой стороны полуинтервала $[a, b)$ или $(a, b]$, открытой или замкнутой полупрямой $(-\infty, a)$,

$(-\infty, a]$, $(a, \infty]$ или $[a, \infty)$, прямой $\mathbb{R}$. Впрочем, формально проще всего вывести их связность из их выпуклости, а в конечном счёте — из связности $[a, b]$; см. ниже.

б). (Непрерывным) путём в M называется непрерывное отображение $f : [a, b] \to M$ отрезка $[a, b]$ в M (в порядке стандартизации обычно берут отрезок $[0, 1]$). Наглядно можно представлять себе, что в M движется точка, занимающая в момент времени $t \in [a, b]$ положение $f(t)$. Точка $x = f(a)$ называется началом пути f, а $y = f(b)$ — его концом; говорят также, что путь f соединяет точки x и y (друг с другом). Пространство M называется линейно связным, если любые две его точки можно соединить путём (в M). В противном случае его можно назвать линейно несвязным.

Обратите внимание, что смысл термина "путь" не совсем согласуется со значением этого слова в обычном языке: наш путь — это не только "дорога", по которой движется (с изменением "времени" t) точка $f(t)$, но и описание того, как это движение происходит с учётом времени. Понятие "пути" даёт некоторую формализацию наглядному представлению о "параметризованной кривой". Наглядной интуиции понятие "кривой, т.е. линии" может представляться более привычным, чем понятие "параметризованной кривой" (когда мы вводим на кривой параметр, мы как бы добавляем к простой геометрической картине некий новый и не очень существенный элемент). Но на самом деле при наличии у кривой самопересечений (даже совсем несложных) не так-то просто дать чёткое определение понятию "кривой"[27] (тогда как определение пути очень просто). Заметим, что поскольку ничего, кроме непрерывности, мы не требуем, "кривая" может быть, с точки зрения "наивной интуиции", довольно патологической (напр., она может быть "кривой Пеано", проходящей через каждую точку квадрата, см.

В.Г.Болтянский, В.А.Ефремович. Наглядная топология. М.: Наука, 1982.

[27]См. параграф о параграф о непрерывных кривых в метрических пространствах в известном учебнике

А.Н.Колмогоров, С.В.Фомин. Элементы теории функций и функционального анализа.

Не воспроизводя имеющегося там определения, приведу две фразы оттуда, в которых выражена его идея: "Порядок, в котором проходятся точки кривой, мы будем считать существенным свойством самой кривой." Этот порядок авторы вводят с помощью параметризации кривой (при этом точки самопересечения как бы считаются несколько раз). "Однако при одинаковом порядке прохождения точек пространства выбор "параметра" t мы будем считать несущественным."

или

Н.Я.Виленкин. Рассказы о множествах. М.: Наука, 1969.)

С геометрической точки зрения, идея линейной связности состоит в том, что любые две точки $x, y \in M$ можно соединить линией (отсюда и прилагательное “линейная”). Не уточняя понятия “линии”, можно согласиться, что во всяком случае, раз x и y соединены “линией”, то от x можно перейти к y, двигаясь вдоль этой линии, причём процесс движения непрерывен. Мы приходим к (тоже достаточно наглядному) “кинематическому” представлению о движущейся точке, которая в момент времени t занимает положение $f(t)$, причём при $t = 0$ она “выходит” из x, а при $t = 1$ “приходит” в y. А это и значит, что существует непрерывное отображение $f : [0, 1] \to M$ с $f(0) = x,\ f(1) = y$. Итак, мы можем не вникать в точный смысл понятия кривой, а иметь дело только с простым понятием пути[28].

Несвязное множество M не может быть линейно связным. Действительно, если $M = A \cup B$ с непересекающимися открытыми A, B, то никакой путь $f : [a, b] \to M$ не может соединять точку из A с точкой из B, ибо в противном случае множества $f^{-1}(A)$ и $f^{-1}(B)$ были бы непересекающимися открытыми подмножествами отрезка $[a, b]$, вопреки связности последнего.

Стало быть, линейно связное множество является связным. Обратное же не обязательно, как показывает приведённый в самом начале пример подмножества F плоскости $\mathbb{R}^2$.

Очевидна линейная связность (а значит, и связность) любого выпуклого подмножества M евклидова пространства (по определению, любые две точки $x, y \in M$ можно соединить путём $(1 - t)x + ty,\ t \in [0, 1]$).

Упражнение. Докажите, что непрерывная числовая функция на связном пространстве M, принимающая только целочисленные значения, постоянна.

Упражнение. Докажите следующие утверждения. Образ связного (или линейно связного) множества при непрерывном отображении свя-

[28]Другая причина, по которой в ряде случаев можно не уточнять понятия кривой, а говорить о путях, состоит в том, что обычно начинают с параметризованной кривой, но затем оказывается, что результаты не меняются не только при изменениях параметризации кривой в духе уточнений из книги Колмогорова – Фомина, но и при гораздо более общих деформациях пути — обычно таких, когда он как бы “шевелится” и только его начало и конец остаются неподвижными. (Точного определения я не даю.)

зен (линейно связен). Если любые две точки множества M содержатся в некотором связном (линейно связном) подмножестве (так сказать, соединяются таким подмножеством), то M связно (линейно связно). Произведение нескольких связных (линейно связных) множеств связно (линейно связно). Объединение любой системы связных (или линейно связных) подмножеств, имеющей непустое пересечение, связно (линейно связно).

Из последнего утверждения следует, что для любой точки $x \in M$ существует наибольшее связное (линейно связное) подмножество M, содержащее эту точку. Оно называется компонентой связности (линейной связности) точки x (в M). Компонента линейной связности x состоит из всех тех точек, которые можно соединить путём с x.

Упражнение. Докажите, что замыкание связного подмножества связно (откуда явствует, что компонента связности любой точки $x \in M$ является замкнутым подмножеством M), и покажите (например, с помощью рассматривавшейся выше фигуры F), что замыкание линейно связного подмножества не обязательно линейно связно.

Упражнение. Докажите, что компоненты связности (линейной связности) двух точек множества M либо совпадают, либо не пересекаются. Отсюда следует, что M является объединением непересекающихся подмножеств — различных компонент связности (линейной связности). В связи с этим их называют также компонентами связности (линейной связности) пространства M или его связными (линейно связными) компонентами.

Когда какое-нибудь множество M (не обязательно топологическое пространство) представлено как объединение его непустых непересекающихся подмножеств, то, как известно, говорят, что эти подмножества образуют разбиение исходного множества M. Таким образом, мы имеем два разбиения M — на его компоненты связности и на его компоненты линейной связности. Второе из них "мельче" первого, а первое "крупнее" второго в том смысле, что каждый элемент второго разбиения целиком содержится в некотором элементе первого разбиения. Мы видели, что иногда эти два разбиения не совпадают. Всё же они совпадают в наиболее интересном для нас случае.

Упражнение. Докажите, что у открытого подмножества $U \subset \mathbb{R}^n$ (а также у открытого подмножества банахова (или даже топологического векторного) пространства) компоненты линейной связности являются открытыми множествами и что в этом случае наши два разбиения совпадают (так что в данном случае нетренированная интуиция, не различающая между связностью и линейной связностью, не делает ошибки).

Упражнение. В некоторых учебниках (например, в упомянутом выше учебнике Колмогорова – Фомина) открытое подмножество $U \subset \mathbb{R}^n$ назвают связным, если любые две его точки x, y можно соединить ломаной, целиком лежащей в U. Докажите, что если точки $x, y \in U$ можно соединить в U путём, то их можно соединить и ломаной, так что определение с ломаными совпадает с нашим. То же справедливо и для открытых подмножеств банахова (или даже топологического векторного) пространства E.

Связные открытые подмножества $\mathbb{R}^n$ (или указанного более общего E) называются областями. Однако порой так называют вообще открытые множества (быть может, даже не подмножества $\mathbb{R}^n$), независимо от того, связны они или нет. (И, конечно, слово "область" может использоваться совсем в другом смысле, — напр., "область определения функции".)

Гладкие подмногообразия евклидова пространства.

k-мерное гладкое подмногообразие M (подробнее M^k — верхний индекс, таким образом, указывает размерность) евклидова пространства $\mathbb{R}^n$ — это обобщение привычных понятий гладкой кривой или гладкой поверхности. Определение данного понятия, которое будет дано, внешне не согласуется с данным в начале §4 описанием "гладкого "k-мерного гладкого локального подмногообразия пространства $\mathbb{R}^n$", хотя последнее должно было бы описывать "маленькие кусочки" подмногообразия M. Но вскоре мы увидим, что эта несогласованность — только внешняя.

Подмножество $M \subset \mathbb{R}^n$ называется k-мерным гладким подмногообразием евклидова пространства, если для каждой точки $a \in M$ существуют такая система декартовых координат $(y_1, \dots, y_n)$ с началом координат в a и такая окрестность U_a этой точки (во всём $\mathbb{R}^n$), что попадающая в U_a часть M (т.е. $M \cap U_a$) описывается в терминах этих координат уравнениями вида

$$y_{k+1} = f_1(y_1, \dots, y_k), \dots, x_n = f_{n-k}(y_1, \dots, y_k),$$

где $f_1, \dots, f_{n-k}$ — гладкие функции своих аргументов, определённые в некоторой окрестности W нуля пространства $\mathbb{R}^k$. Обозначим координатные подпространства, отвечающие группам координат $(y_1, \dots, y_k)$ и $(y_{k+1}, \dots, y_n)$, через E_a^k и F_a^{n-k}, так что, например, E_a^k состоит из точек x

с координатами $(y_1, \dots, y_k, 0, \dots, 0)$. Мы часто позволяем себе отождествлять E_a^k с $\mathbb{R}^k$ и F_a^{n-k} — с $\mathbb{R}^{n-k}$, считая, что $(y_1, \dots, y_k)$ — это та самая точка x, которая имеет координаты $(y_1, \dots, y_k, 0, \dots, 0)$, а $(y_{k+1}, \dots, y_n)$ (это набор из $n - k$ чисел, т.е. точка $\mathbb{R}^{n-k}$) есть точка x с координатами $(0, \dots, 0, y_{k+1}, \dots, y_n)$, и в соответствии с этим рассматривать $\mathbb{R}^n$ как прямое произведение $E_a^k \times F_a^{n-k}$ и даже как $\mathbb{R}^k \times \mathbb{R}^{n-k}$ (надо только помнить, что в таком виде наше евклидово пространство представляется не в терминах координат x_i, а в терминах координат y_i, зависящих от a). Тогда можно сказать, что $M \cap U_a$ является графиком гладкой функции $v = f_a(w)$, отображающей W в $\mathbb{R}^{n-k}$ (мы, значит, обозначили $(y_1, \dots, y_k)$ через w и $(y_{k+1}, \dots, y_n)$ — через v). Если для всех $a \in M$ соответствующие функции f_a имеют класс гладкости C^r, то и Mназывают подмногообразием класса C^r или C^r-подмногообразием. Не искючается ни такая возможность, что $r = \infty$, ни такая, что функции f_a являются вещественно-аналитическими (т.е. возле каждой точки своей облоасти определения разлагаются в сходящиеся степенные ряды); в последнем случае часто пишут $r = \omega$.

Приведённое определение, повидимому, в наибольшей степени отвечает наглядному представлению, что стоит считать обобщением понятий гладкой кривой или гладкой поверхности. Во-первых, график гладкой функции является гладким многообразием — это простейший случай. Во-вторых, в общем случае гладкое подмногообразие возле любой своей точки устроено так же, как в этом простейшем случае.

Положение любой точки $x \in M \cap U_a$ полностью характеризуется набором чисел $w = (y_1, \dots, y_k)$, поскольку он определяет и остальные её координаты $v = (y_{k+1}, \dots, y_n) = f_a(w)$. Поэтому набор чисел $w = (y_1, \dots, y_k)$ естественно назвать координатами точки x. Эти координаты являются локальными — они определены только в окрестности $W = M \cap U_a$ точки a. Координатные линии — те линии в W, где изменяется одна из локальных координат y_i, а другие постоянны, — являются, вообще говоря, кривыми, поэтому сами координаты можно охарактеризовать как "криволинейные". Когда мы представляем R^n как $E_a^k \times F_a^{n-k}$, точка $x \in M \cap U$ с локальными криволинейными координатами w представляется парой $(w, f_a(w))$. Поэтому точка X гладко зависит от своих локальных координат w. (От них гладко зависят координаты y_i этой точки, поскольку для неё $(y_1, \dots, y_n)$ как раз и есть $(w, f(w))$. А переход от координат y_i к исходным координатам x_i в $\mathbb{R}^n$ осуществляется с помощью неоднородного невырожденного линейного преобразования.)

В данном случае координаты точки x получаются при её проектировании на E_a^k параллельно F_a^{n-k}. Спроектировать таким образом на E_a^k можно любую точку из M, но, во-первых, различные точки M могут проектироваться в одну и ту же точку подпространства E_a^k (представьте себе ортогональное проектирование окружности на её диаметр) и, вовторых, если точка $x(w) \in M \cap U$ определена при всех $w \in W$ и гладко зависит от v, то вдали от a ни то, ни другое не обязаны иметь место. (При проектировании на ось y графика Γ функции $y = x^3 - x$, которая монотонно возрастает на $\left(-\infty, -\frac{1}{\sqrt{3}}\right]$ и на $\left[\frac{1}{\sqrt{3}}, \infty\right)$, а на $\left[-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]$ монотонно убывает, получается, что возле точки $y = -\frac{1}{\sqrt{3}}$ или $y = \frac{1}{\sqrt{3}}$ по одну сторону от неё точки имеют по три прообраза, а по другую — по одному. А проектирование кривой $\Gamma = \{(x, y);\ y = x^3\}$ на ось y биективно, но в точке $y = 0$ нарушается гладкая зависимость проектируемой точки от y. В обоих примерах Γ является гладким одномерным подмногообразием плоскости, что сразу видно, если проектировать Γ на ось x-ов. В данном примере на подходящее проектирование приводит всё-таки к “хорошей” координате на всей кривой. Но на окружности никакое проектирование к такой координате не приводит.)

Ещё до того, как школьник услышит о координатах на плоскости, он узнаёт о координатах на земном шаре. Например, Москва имеет координаты: $55\frac{1}{2}^o$ северной широты и $37\frac{1}{2}^o$ восточной долготы. Зная об этом, можно сразу найти на немой карте, где нет никаких названий, но нанесена сетка параллелей и меридианов, то место, где находится Москва. Исторически и были сперва введены географические координаты на Земле (и родственные координаты нескольких типов на небесной сфере), а только потом появились более простые декартовы координаты на плоскости. Однако эти старейшие координаты на поверхности (на сфере) не получаются таким образом, как описано выше, даже если рассматривать их только в окрестности какой-нибудь точки a сферы. Ведь если бы они получались таким образом, то при ортогональном проектировании на плоскость E_a^2 параллельно прямой F_a^1 дуги параллелей и меридианов, попадающие в окрестность W точки a на сфере, должны были бы переходить в прямолинейные отрезки. Однако это невозможно. Ведь меридианы и параллели суть некоторые окружности. А если при проектировании окружности на плоскость E_a^2 параллельно прямой F_a^1 получается не эллипс, а прямая, то эта окружность лежит в плоскости, параллельной F_a^1. Значит, прямая F_a^1 параллельна экваториальной плоскости (плоскости земного экватора). Но тогда только два меридиана (являющиеся половинами окружности, лежащей в плоскости,содержащей земную ось и параллельной перпендикулярной этой плоскости прямой F_a^1) проектируются в отрезки прямых.

Примеры такого рода наводят на мысль, что понятию "криволинейные локальные координаты" нецелесообразно ограничивать только такими координатами, которые получаются как декартовы координаты проекции точки x на какое-то линейное подпространство. В этом месте удобно несколько изменить точку зрения и говорить не о координатах в некоторой окрестности W точки a на M (при этом точке $x \in W$ должен был бы сопоставляться какой-то набор чисел), а, наоборот, о сопоставлении набору чисел какой-то точки. Традиционное название для той ситуации, когда точка x зависит от чисел — скажем, от $w = (w_1, \ldots, w_k)$, — параметризация. На параметризацию нужно наложить некоторые условия "невырожденности" (помимо гладкости), потому что без них множество $\{x(w)\}$ может никак не подходить под роль "k-мерного обобщения глакдой кривой или поверхности". Если, к примеру, $x(w) = \text{const}$, то $\{x(w)\}$ сводится к одной точки, сколько бы параметров w_i ни входило в w. Читателю, вероятно, известны также примеры, когда при одном (числовом) параметре t "кривая" $\{x(t)\}$ никак не кажется гладкой, — например, она может иметь угловую точку. Скажем, с ростом t точка $x(t)$ может пробегать множество $([0,1] \times \{0\}) \cup (\{0\| \times [0,1])$, полностью изменив направление своего движения при прохождении через начало координат. Это связано с тем, что в тот момент, когда $x(t) = (0,0)$, скорость $\dot{x}(t) = 0$. Поэтому при определении понятия (параметризованной) гладкой кривой требуют, чтобы скорость $\dot{x}(t)$ нигде не обращалась в нуль. Это одно гарантирует, что маленькая дуга кривой (скажем, дуга вида $x([t_0 - \varepsilon, t_0 + \varepsilon])$ действительно соответствует наглядному представлению о гладкой кривой и формально подпадает под данное выше определение гладкого k-мерного подмногообразия $\mathbb{R}^n$ (с $k = 1$). Однако возможности самопересечений это ещё не исключает (вполне может случиться, что $x(t) = x(t_0)$ при каком-то $t \notin [t_0 - \varepsilon, t_0 + \varepsilon]$). Их отсутствие надо оговаривать специально.

Что следовало бы понимать в данном случае под "невырожденностью" гладкого отображения $g : G \to \mathbb{R}^n$ открытого подмножества G пространства $\mathbb{R}^k$ в $\mathbb{R}^n$? Если бы это было линейное отображение $g(w) = Aw + b$ (где матрица A имеет k столбцов и n строк), то мы потребовали бы, чтобы его образ — некоторое относительно открытое подмножество пространства $A\mathbb{R}^k + b$[29] — был k-мерным, т.е. чтобы k-мерным было само $A\mathbb{R}^k + b$. Это эквивалентно тому, что ранг матрицы A равен k [30]. Нелинейное же отображение g возле какой-нибудь точки $w_0 \in G$ мы аппроксимируем линейным отображением $w \mapsto g(w_0) + g'(w_0)(w - w_0)$, где g' — производная Фреше, т.е. попросту матрица

$$\begin{pmatrix} \frac{\partial g_1(w_0)}{\partial w_1} & \dots & \frac{\partial g_1(w_0)}{\partial w_k} \\ \dots & \dots & \dots \\ \frac{\partial g_n(w_0)}{\partial w_1} & \dots & \frac{\partial g_n(w_0)}{\partial w_k} \end{pmatrix}.$$

Поэтому разумно назвать ранг $\mathrm{Rk}\, g'(w_0)$ этой матрицы рангом отображения g в точке w_0 и понимать "невырожденность" g как условие, что это отображение всюду (во всех точках $w_0 \in G$) ранг k.

Подобные соображения подсказывают, что *подмножество $M \subset \mathbb{R}^n$ является гладким k-мерным подмногообразием, если у каждой точки $a \in M$ имеются такие окрестность U_a этой точки во всём объемлющем пространстве $\mathbb{R}^n$ и такое гладкое отображение $g : V \to \mathbb{R}^n$ некоторого открытого подмножества V пространства $\mathbb{R}^k$ в $\mathbb{R}^n$, которое всюду имеет ранг k, инъективно (никакие две точки не переходят в одну: $g(w) \neq g(w')$ при $w \neq w'$) и образ которого есть попадающая в U_a часть M: $M \cap U_a = g(V)$.* В окрестности $W = M \cap U_a$ точки a на M тем самым вводятся локальные координаты — точке $x \in W$ сопоставляется тот набор чисел $w = (w_1, \dots, w_k)$, который переходит в x при отображении g; иными словами, $w = g^{-1}(x)$. Говорят, что пара (W, g^{-1}) является картой (точнее, гладкой картой) подмногообразия M, что W есть соответствующая координатная окрестность, а $g^{-1} : W \to \mathbb{R}^k$ — координатное отображение, сопоставляющее точке x её локальные координаты $(w_1, \dots, w_k)$[31].

[29] А почему оно относительно открыто?

[30] Одно из равносильных друг другу определений ранга матрицы A состоит в том, что это есть максимальное число линейно независимых столбцов A. А эти столбцы суть образы базисных векторов отображаемого пространства $\mathbb{R}^k$ (в "дифференциальных" обозначениях, векторов $\frac{\partial}{\partial w_i}$), порождающих " векторное подпространство $A\mathbb{R}^k$, при "сдвиге" (трансляции) которого (на вектор b) и получается $A\mathbb{R}^k + b$.

[31] Мы могли бы вместо понятия параматризации исходить из понятия карты — пары (W, h), где W — открытое подмножество в M, а $h : W \to \mathbb{R}^k$ — отображение, гомеоморфно отображающее W на некоторое открытое подмножество G пространства $\mathbb{R}^k$ и .... Тут надо сказать нечто о гладкости и ранге h, но как их понимать? Единственная возможность — перейти к отображению $h^{-1} : G \to \mathbb{R}^n$, применительно

Докажем утверждение, выделенное курсивом. Пусть $a = (a_1, \ldots, a_n) \in M$ и U_a, G, g — такие, как сказано в этом утверждении, причём $a = g(w_0)$. Раз $\operatorname{Rk} g'(w_0) = k$, то у матрицы $g'(w_0)$ имеется ненулевой минор k-го порядка. Пусть его элементы стоят в строках с номерами $i_1 < i_2 < \ldots < i_k$, так что

$$\begin{vmatrix} \frac{\partial g_{i_1}(w_0)}{\partial w_1} & \ldots & \frac{\partial g_{i_1}(w_0)}{\partial w_k} \\ \ldots & \ldots & \ldots \\ \frac{\partial g_{i_k}(w_0)}{\partial w_1} & \ldots & \frac{\partial g_{i_k}(w_0)}{\partial w_k} \end{vmatrix} \neq 0 \ .$$

Номера, оставшиеся в множества $\{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\}$, обозначим через $j_1 < j_2 < \ldots < j_{n-k}$. Обозначим переменные $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ через $y = (y_1, y_2, \ldots, y_k)$, а остальные x_i, пронумерованные в порядке возрастания индексов — через $z = (z_1, \ldots, z_{n-k})$. Сообразно с этим, обратим внимание на "координатиное подпространство" с координатами y, которое обозначим через E_a^k, и на "координатиное подпространство" с координатами z, которое обозначим обозначим через F_a^{n-k}. (В первом $x_{j_1} = a_{j_1}, x_{j_2} = a_{j_2}, \ldots, x_{j_{n-k}} = a_{j_{n-k}}$, а во втором $x_{i_1} = a_{i_1}, x_{i_2} = a_{i_2}, \ldots, x_{i_k} = a_{i_k}$.) $\mathbb{R}^n$ представляется как $E_a^k \times F_a^{n-k}$; g теперь становится отображением $W \to E_a^k \times F_a^{n-k}$, которое можно записать как

$$(y, z) \mapsto g(y.z) = (\overline{g}(w), \widehat{g}(w)),$$

$$\text{где } \overline{g}(w) = (g_{i_1}(w), \ldots, g_{i_k}(w)),\ \widehat{g}(w) = (g_{j_1}(w), \ldots, g_{j_{n-k}}(w));$$

наконец, точка a записывается как некоторая пара $(\overline{a}, \widehat{a}) = (\overline{g}(w_0), \widehat{g}(w_0))$. В новых обозначениях $\det \overline{g}'(w_0) \neq 0$, и по теореме о неявных функциях, в пространстве E_a^k существует такая окрестность W_1 точки $\overline{a}$, в этой окрестности имеется такая гладкая функция $h : W_1 \to \mathbb{R}^k$ и в V имеется такая подобласть V_1, что $h(W_1) \subset V_1$, $\overline{g}(h(y)) = y$ при всех $y \in W_1$ и при этом $h(y)$ — единственная точка V_1, для которой $\overline{g}(w) = y$. Положим при всех $y \in W_1$

$$f_a(y) = \widehat{g}(h(y)).$$

Уменьшив, если потребуется, W_1, можем считать образ $f(W_1)$ гладкого отображения $f_a : W_1 \to F_a^{n-k}$ содержащимся в столь малом шаре D пространства f_a^{n-k} с центром в $\widehat{a}$, что для окрестности $U'_a = W_1 \times D$ точки a в $\mathbb{R}^n$ справедливо следующие два условия малости: $U'_a \subset U_a$ и

к которому можно говорить и о гладкости, и о ранге. Получается, что в решающий момент мы всё-таки вынуждены фактически обращаться к параметризации.

окрестность $U'_a \cap M$ точки a на M содержится в $g(V_1)$. Тогда часть M, попадающая в U'_a, совпадает с графиком функции f_a. Действительно, этот график содержится в M, поскольку

$$(y, f_a(h(y)) = (\overline{g}(h(y)), \widehat{g}(h(y))) = g(h(y)).$$

В то же время все точки x из $M \cap U'_a$ лежат на этом графике. Ведь такая точка, с одной стороны, является образом $g(w) = (\overline{g}(w), \widehat{g}(w))$ с $w \in V_1$ и, с другой стороны, является парой (y, z) с $y \in W_1,\ z \in D$. Раз $y = \widehat{(w)}$ с $y \in W_1,\ w \in V_1$, то $w = h(y)$, а тогда $z = \widehat{g}(h(y)) = f_a(y)$, так что точка x лежит на графике функции f_a.

———————————————————————————

Специальная ортогональная группа.

Матрица n-го порядка A называется ортогональной, если определяемое ею линейное преобразование в $\mathbb{R}^n$ не меняет стандартного скалярного произведения[32]: $\langle Ax, Ay \rangle = \langle x, y \rangle$ при любых $x, y \in \mathbb{R}^n$. Как хорошо известно и легко доказывается,“в чисто-матричных терминах” это равносильно тому, что сопряжённая матрица A' является обратной к той же A: $A'A = I$ (единичная матрица) и $AA' = I$ (каждое из этих двух равенств влечёт за собой другое). Определитель ортогональной матрицы равен ± 1.

Ортогональные матрицы n-го порядка образуют группу, которая называется ортогональной группой и обозначается через $O(n)$. а ортогональные матрицы с определителем 1 — её подгруппу, которая называется специальной ортогональной группой и обозначается через $SO(n)$. В соответствии с этим ортогональные матрицы с определителем 1 тоже называют специальными ортогональными матрицами, но, кажется, прилагательное “специальная” всё-таки чаще употребляют применительно к группе $SO(n)$, чем к её элементам. Обе группы $O(n)$ и $SO(n)$ суть некоторые подмножества множества $Mat(n)$ всех квадратных матриц n-го порядка [33]. Квадратная матрица $A = (a_{ij})$ n-го порядка имеет n^2 ко-

[32] Для $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n)$ (как мы помним, лучше считать, что на самом деле это столбцы, что, впрочем, в данный момент несущественно) скалярное произведение $\langle x, y \rangle = \sum x_i y_i$.

[33] Сейчас речь идёт только о матрицах с вещественными коэффициентами. Если бы у нас фигурировали матрицы то с вещественными, то с комплексными коэффициентами, то надо было бы детализировать обозначения и писать, соответственно, что-нибудь вроде $Mat(n, \mathbb{R})$ или $Mat(n, \mathbb{C})$.

эффициентов a_{ij}, которые независимо друг от друга могут принимать любые вещественные значения; поэтому $Mat(n)$ можно отождествить с пространством $\mathbb{R}^{n^2}$. Всё сводится к тому, чтобы заново пронумеровать коэффициенты a_{ij} — первоначально они нумеруются двумя индексами (i,j), принимающими значения от 1 до n, а теперь вместо этого надо их пронумеровать одним индексом, принимающим значения от 1 до n^2. В $\mathbb{R}^{n^2}$ определено расстояние между точками и тем самым топология. Тем самым эти понятия переносятся и в множество матриц $Mat(n)$. Расстояние между точками, кстати, не зависит от выбранной нумерации координат. В терминах коэффициентов матриц $A=(a_{ij}),\ B=(b_{ij})$ оно равно просто $\sqrt{\sum_{i,j}(a_{ij}-b_{ij})^2}$ (если исходить из "евклидова" расстояния в $\mathbb{R}^{n^2}$). "Достаточная близость" B к A означает просто "достаточную близость" каждого из n^2 коэффициентов b_{ij} к a_{ij}.

Упражнение. $Mat(n)$ является конечномерным нормированным пространством — это, очевидно, векторное пространство и в нём имеется "операторная" норма $\|A\|=\sup_x \frac{|Ax|}{|x|}$. $\sqrt{\sum_{ij} a_{ij}^2}$ —тоже норма в этом пространстве (ведь это обычная евклидова норма в соответствующем $\mathbb{R}^{n^2}$). Обозначим её через $\left|A\right|$. Докажите, что

$$\|A\| \leqslant \left|A\right| \leqslant \sqrt{n}\|A\|$$

и что эти неравенства не могут быть улучшены.

(Указание. Точность одного из неравенств видна на примере матрицы, в которой только один столбец ненулевой, а другого — на примере единичной матрицы. Чтобы доказать левое неравенство, установите, что $|Ax|^2 \leqslant \left|A\right|^2|x|^2$ для любого $x\in\mathbb{R}^n$. Что касается правого неравенства, то, вероятно, наиболее "идейное" его доказательство, хотя и идущее несколько кружным путём, таково.

а). Наряду с нормой $\left|.\right|$ введём в $Mat(n)$ скалярное произведение $\langle A,B\rangle=\sum_{ij}a_{ij}b_{ij}$ (едва ли надо объяснять использованные обозначения). Получите для неё других выражения. Во-первых, если a_i и b_i — столбцы (соответственно, строки) этих матриц, то $\langle A,B\rangle=\sum\langle a_i,b_i\rangle$. Во-вторых, $\langle A,B\rangle=\mathrm{tr}\,(AB')=\mathrm{tr}\,(A'B)$, где $\mathrm{tr}\,C$ обозначает след матрицы C, а штрих — транспозицию.

б). Определим линейные отображения

$$L_C: Mat(n)\to Mat(n), \qquad R_C: Mat(n)\to Mat(n)$$

как
$$A \mapsto CA \quad \text{и} \quad A \mapsto AC$$
(L и R происходят, конечно, от left и right). Докажите, что если $C \in O(n)$, то $\langle L_C A, L_C B\rangle = \langle A, B\rangle = \langle R_C A, R_C B\rangle$ и что в общем случае $|L_c A| \leqslant \|C\||A|, \quad |R_c A| \leqslant \|C\||A|$.

в). Проверьте, что достаточно доказать нераенство $|A|^2 \leqslant n\|A\|$ для положительно определённых симметрических матриц A (для них $(Ax, x) \geqslant 0$ при всех $x \in \mathbb{R}^n$[34]). (Указание. Представьте рассматриваемую матрицу как произведение ортогональной и положительно определённой симметрической.)

г). Завершите доказательство неравенства $|A|^2 \leqslant n\|A\|$. (Указание. Докажите, что след $\operatorname{tr} B$ положительно определённой симметрической матрицы B равен сумме её собственных значений, а норма $\|B\|$ равна наибольшему собственному значению B.)

Две нормы $|\,.\,|_1$ и $|\,.\,|_2$ в векторном пространстве V (над $\mathbb{R}$) называются *эквивалентными*, если имеются такие $C_1, C_2 > 0$, что
$$C_1|x|_1 \leqslant |x|_2 \leqslant C_2|x|_2 \qquad \text{при всех } x \in V.$$
(А это действительно отношение эквивалентности?) Эквивалентные нормы определяют одну и ту же топологию (почему?). Более того, можно доказать, что вообще в конечномерном векторном пространстве имеется только одна хаусдорфова топология, совместимая со структурой векторного пространства (т.е. такая, что по отношению к ней и к обычной топологии в $\mathbb{R}$ непрерывны отображения
$$V \times V \to V \qquad (x, y) \mapsto x + y$$
и
$$\mathbb{R} \times V \to V \qquad (\lambda, x) \mapsto \lambda x.)$$

Таким образом, выше мы конкретизировали утверждение об эквивалентности норм $|.|$ и $\|\,.\,\|$ в $Mat(n)$ — эквивалентность является проявлением общего факта, но мы указали конкретные значения соответствующих C_i.

[34] В общем случае матрица A называется положительно определённой, если аналогичное неравенство имеет место в $\mathbb{C}^n$ при использовании стандартного эрмитова произведения $\langle z, w\rangle = \sum \overline{z_i} w_i$. Это условие сильнее, чем аналогичное условие в вещественной области. Но для вещественных симметрических матриц оно эквивалентно данному выше.

В R^{n^2} можно говорить о гладкости, что позволяет пользоваться соответствующими понятиями и для матриц — скажем, говорить о гладкой зависимости матрицы от параметров или о том, что некоторая величина, зависящая от матрицы, зависит от неё гладко. В терминах коэффициентов матриц всё это хорошо известно, я просто хочу отметить, что по существу здесь нет ничего нового — в этих терминах просто повторяются понятия и их свойства, обсуждающиеся в начальном курсе матанализа применительно к $\mathbb{R}^m$. Раз в $Mat(n)$ имеется топология, то соответствующая топология ("относительная" или "индуцированная") определена и для любых подмножеств $Mat(n)$, в частности, для $SO(n)$. Несколько менее тривиально обстоят дела с "гладкими" понятиями. Ясно, как понимать гладкую зависимость матрицы от параметра — надо просто рассматривать эту матрицу как точку в $Mat(n)$ (хотя бы мы и знали, что это матрица какого-то специального типа, скажем, ортогональная). Если угодно, это сводится к гладкому характеру зависимости коэффициентов этой матрицы от параметра. Но что понимать под гладкой зависимостью какой-то "величины" (может быть, числовой, а может быть, и нет) от матрицы $A \in SO(3)$? Если бы группа $SO(n)$ была просто подмножеством пространства $Mat(n)$, вопрос остался бы без ответа. Но, к счастью, $SO(n)$ является гладким подмногообразием этого пространства.

Действуя эвристически, попытаемся сперва угадать, каким должно было бы быть касательное пространство $T_I(SO(n))$ к группе $SO(n)$ в точке I, если бы $SO(n)$ действительно была гладким подмногообразием в $Mat(n)$. Затем мы введём декартовы координаты в $Mat(n)$, часть которых пробегает найденное нами гипотетическое касательное пространство $T_I(SO(n))$, а другая часть — перпендикулярное к нему векторное пространство, и убедимся, что в терминах этих координат часть $SO(n)$, близкая к I, действительно выглядит как график гладкой функции. После этого уже легко будет получить аналогичное описание части $SO(n)$, близкой к любой другой специальной ортогональной матрице I.

Пусть $A(t)$ — параметризованная гладкая кривая в $SO(n)$, определённая при малых $|t|$ и проходящая при $t = 0$ через I. Дифференцируя по t тождественное равенство $A'(t)A(t) = I$, получим:

$$\dot{A}'(t)A(t) + A'(t)\dot{A}(t) = 0, \qquad \dot{A}'(0) + \dot{A}(0) = 0,$$

т.е. матрица $\dot{A}(0)$ — кососимметрическая (или, как ещё говорят, антисимметрическая). Если предположить, что $SO(n)$ является гладким подмногообразием в $Mat(n)$, то $\dot{A}(0)$ является касательным вектором к $SO(n)$ в точке $A(0) = I$ и в направлении любого касательного вектора $X \in T_I(SO(n)$ можно провести гладкую параметризованную кривую, говоря более формально — существует гладкая функция $A(t)$ со значениями в $SO(n)$, для которой $A(O) = I$ и $\dot{A}(0) = X$. Итак, касательные векторы к $SO(n)$ в единице этой группы являются кососимметрическими матрицами.

Отсюда ещё не следует, что любая кососимметрическая матрица A является касательным вектором к $SO(n)$. Но можно заподозрить, что это так, на основании следующего соображения. Возьмём параметризованную кривую $A(t) = I+tA$ в $Mat(n)$. Для неё $A'(t)A(t) = (I-tA)(I+tA) = I - t^2A^2$ — она не лежит в $SO(n)$, для неё условие ортогональности $A'(t)A(t) = I$ нарушается, но это нарушение имеет всего второй порядок малости по t, а не первый. Естественно думать, что наша кривая $A(t)$ отходит от $SO(n)$ тоже на расстояние порядка t^2 и что возле этой кривой имеется какая-то другая (“немного подправленная”) кривая $B(t)$, которая лежит в $SO(n)$ и отличается от $A(t)$ только на величину порядка t^2. Тогда $B(0) = I,\ \dot{B}(0) = A$, так что $A \in T_I(SO(n))$.

(На самом деле можно взять $B(t) = e^{At}$, где используется матричный экспоненциал. Но всё равно заключение, что $A \in T_I(SO(n))$, остаётся пока что условным — его можно сделать при условии, что $SO(n)$ является гладким подмногообразием в $Mat(n)$ и потому касательное пространство $T_I(SO(n))$ существует.)

Итак, мы подозреваем, что касательным пространством $T_I(SO(n))$ служит пространство всех кососимметрических матриц n-го порядка. Обозначим его на минуту через $\mathfrak{A}$.

Упражнение. Покажите, что подпространство $\mathfrak{S} \in Mat(n)$, перпендикулярное $\mathfrak{A}$ в смысле скалярного произведения $\langle , \rangle$, , состоит из всевозможных симметрических матриц (отчего я м выбрал на минуту обозначение $\mathfrak{S}$).

Теперь надо доказать главное: у единичной матрицы n-го порядка I имеется такая окрестность U в $Mat(n)$, что $U \cap SO(n)$ есть график некоторой C^∞-гладкой функции $S = S(A)$, где $A \in U \cap \mathfrak{A},\ S(A) \in \mathfrak{S}$. Это значит, что $U \cap SO(n)$ состоит из сумм $I + A + S(A)$ с $A \in U \cap \mathfrak{A}$.

Условие ортогональности матрицы V состоит в том, что $V'V = I$. (Для $SO(n)$ есть ещё дополнительное условие, что $\det V = 1$, а не -1, но $\det V$ не может равняться -1 возле I.) Взяв $A \in \mathfrak{A},\ A \approx 0$, ищем такую малую $S \in \mathfrak{S}$, чтобы матрица $V = I + A + S$ была ортогональной:

$$(I - A + S)(I + A + S) = I,$$

$$I - A + S + A - A^2 + SA + S - AS + S^2 = I,$$

$$\begin{gathered} 2S + S^2 - A^2 + [S, A] = 0 \\ ([S, A] \text{ —коммутатор матриц } S \text{ и } A\ , [S, A] = SA - AS). \end{gathered} \tag{1}$$

Забавный факт: это ур-ие эквивалентно более простому

$$2S + S^2 - A^2 = 0. \tag{2}$$

Эквивалентность этих уравнений означает, что (1) $\Longrightarrow [S, A] = 0$ и (2) $\Longrightarrow [S, A] = 0$. Докажем, например, первую импликацию. По A и S строим ортогональную матрицу $V = I+A+S$. Тогда $V^{-1} = V' = I-A+S$,

$$V + V^{-1} = 2I + 2S, \quad V - V^{-1} = 2A,$$

и так как V, V^{-1} и I коммутируют, то S и A тоже коммутируют.

Можно было бы доказать существование решения у (1) с помощью теоремы о неявной функции. (При этом можно было бы также усмотреть, что $S(A)$ представляется в виде некоторого степенного ряда от A со скалярными коэффициентами, а тогда понятно, что $S(A)$ и A коммутируют.) Но проще сразу найти решение (2): Из этого уравнения видно, что

$$S^2 = I - A^2.$$

Справа стоит симметрическая матрица, а поскольку A малая (не то чтобы уж очень малая — $\|A\| < 1$), то эта правая часть положительно определена. Из такой симметрической матрицы извлекается квадратный корень (а): приведём её к главным осям и ...; б) биномиальный ряд Ньютона). Ясно также, что других решений возле 0 нет. Заметим ещё, что $S(A) = o(A)$ и потому производная Фреше $S_A(0) = 0$. (Подробности всего этого предоставляются читателю).

Итак, возле I в терминах координат с началом в I и координатными пространствами $\mathfrak{A}, \mathfrak{S}$ группа $SO(n)$ действительно является графиком некоторой гладкой функции $S = S(A)$. В терминах тех же "координат" (A, S) её касательное пространство в I, т.е. в точке $(0, 0)$, состоит из векторов $(A, S_A(0)A) = (A, 0)$ со всевозможными $A \in \mathfrak{A}$. Это значит, что данное касательное пространство, "перенесённое" из I в 0, есть $\mathfrak{A}$, как и утверждалось.

Последний шаг: надо посмотреть, как устроено $SO(n)$ возле любой своей матрицы $A_0 \in SO(n)$. Идея состоит в том, чтобы "перенести" всё в окрестность U единичной матрицы — ту окрестность, где $SO(n) \cap U$ описывается в терминах "координат" $A \in \mathfrak{A}, S \in \mathfrak{S}$ (и при сдвиге начала координат из O в I) как график функции $S(A)$. Для этого используем правый сдвиг $Mat(n)$ на элемент A_0^{-1}:

$$R_{A_0^{-1}} : Mat(n) \to Mat(n) \qquad g \mapsto gA_0^{-1}.$$

Если $A_0 \in SO(n)$, то $R_{A_0}^{-1} SO(n) = SO(n)$. Сдвинутая на A_0 прежняя окрестность U становится окрестностью $U_{A_0} = R_{A_0}U$ точки A_0, подпространства $\mathfrak{a}$ и $\mathfrak{S}$ переходят в новые подпространства $\mathfrak{A}A_0$ и $\mathfrak{S}A_0$, попрежнему перпендикулярные друг другу и в сумме порождающие всё $Mat(n)$. При этом

$$R_{A_0^{-1}}(SO(n) \cap U_{A_0}) = R_{A_0}(SO(n) \cap U) = R_{A_0}\{I + A + S(A);\ A \in \mathfrak{A}, |A| < 1\} =$$

$$= \{A_0 + AA_0 + S(A)A_0;\ AA_0 \in \mathfrak{A}A_0, |AA_0| =$$
$$= |A| < 1\} = \{A_0 + B + S(BA_0^{-1})A_0;\ B \in \mathfrak{A}A_0, |B| < 1\}.$$

Как видно, в терминах разложения пространства $Mat(n)$ в прямую сумму "координатных подпространств" $\mathfrak{A}A_0$ и $\mathfrak{S}A_0$, рассматриваемая часть группы $SO(n)$ является графиком определённой при $B \in \mathfrak{A}A_0$, $|B| < 1$ гладкой (почему?) функции $B \mapsto S(BA_0^{-1})A_0$, принимающей значения в $\mathfrak{S}A_0$.

———————————————————————————