Федеральное государственное бюджетное учреждение науки
Математический институт им. В.А. Стеклова Российской академии наук

На правах рукописи

Малыхин Юрий Вячеславович

Оценки колмогоровских поперечников и связанных с ними величин

Специальность 1.1.1 — вещественный, комплексный и функциональный анализ

Диссертация на соискание ученой степени
доктора физико-математических наук

Москва — 2025

Оглавление

Введение **3**

1 Усреднённый колмогоровский поперечник **32**
1.1 Предварительные сведения 32
1.2 Определение и свойства d_n^{avg} 44
1.3 Нижние оценки поперечников в L_q и ℓ_q^N 50
1.4 Приближение специальных систем функций 67

2 Поперечники классов функций критической гладкости **79**
2.1 Класс Соболева W_1^1 79
2.2 Классы Бесова $B_{1,\theta}^1$ 85

3 Поперечники классов функций высокой гладкости **100**
3.1 Относительные поперечники 100
3.2 Классы W_∞^r на отрезке и чебышёвские сплайны 111

4 Полилинейные приближения **135**
4.1 Векторы знаков для полиномиального отображения 135
4.2 Жёсткость и сигнум ранг тензоров 137
4.3 Приближение по словарю полилинейных функций 141

5 Избранные задачи о поперечниках в $\mathbb{R}^N$ **149**
5.1 Поперечники шаров в смешанных нормах 149
5.2 Свойства матриц, приближающих единичную 162
5.3 “Свидетели” поперечника 168

Заключение **172**

Список литературы **174**

Введение

Актуальность темы. Основной предмет диссертации — поперечник по Колмогорову. Это понятие ввёл А.Н. Колмогоров в 1936 году в работе [45]. Для множества K в линейном метрическом пространстве (X, ρ) поперечник порядка n определяется как величина наилучшего приближения множества K посредством n-мерных линейных подпространств в X:

$$d_n(K, X) := \inf_{\substack{Q_n \subset X \\ \dim Q_n \leqslant n}} \sup_{x \in K} \inf_{y \in Q_n} \rho(x, y).$$

Поперечник d_n — фундаментальная аппроксимативная характеристика и, более того, — важное и актуальное понятие функционального анализа в целом. Последовательность n-поперечников $(d_n(K, X))_{n=1}^{\infty}$ характеризует "массивность" тела K в нормированном пространстве, поведение этой последовательности связано с ε-энтропией и другими характеристиками тела.

Классическая задача о колмогоровских поперечниках классов Соболева одной переменной W_p^r в L_q активно изучалась в 1960-х — 1980-х гг. советскими и некоторыми западными математиками. После случая $p = q = 2$, рассмотренного Колмогоровым, порядки этих поперечников были вычислены усилиями С.Б. Стечкина, Г.Г. Лоренца, В.М. Тихомирова, Р.С. Исмагилова и других математиков. Б.С. Кашин в 1977 году завершил решение этой задачи в случае $rp > 1$, когда пространство Соболева компактно вложено в C, построив для этого случайным методом n-мерные подпространства, хорошо приближающие в равномерной норме евклидов шар B_2^N большой размерности N. Один из основных результатов диссертации — нахождение порядков поперечников в единственном остававшемся открытым для целой гладкости случае $r = p = 1$, $2 < q < \infty$. Подробная история этого вопроса изложена на стр. 18; см. также книги [119, 127] и обзор [80]. Здесь лишь отметим, что важный вклад в нахождение поперечников внесли также Е.Д. Глускин, В.Е. Майоров, Ю.И. Маковоз,

А.П. Буслаев.

В связи с поперечниками классов Соболева изучалась задача о поперечниках конечномерных ℓ_p^N-шаров в ℓ_q^N. Точные значения были найдены при $p = 1$, $q = 2$ (см. стр. 12) и при $p \geqslant q$ (А. Пич, М.И. Стесин). Результат Кашина 1977 года о поперечнике шара, дающий правильный порядок этой величины при $n \asymp N$, был уточнён А.Ю. Гарнаевым и Глускиным (1984); ими было доказано порядковое равенство, справедливое при всех $n < N$:

$$d_n(B_2^N, \ell_\infty^N) \asymp \min\{1, n^{-1/2}\log^{1/2}(2N/n)\}. \tag{0.1}$$

Ряд окончательных результатов получен Глускиным (1983). Наиболее важные оценки приведены в §1.1.

Позже основной интерес сместился к классам функций многих переменных, в том числе классам смешанной гладкости; их исследовали К.И. Бабенко, В.Н. Темляков, Э.М. Галеев, Динь Зунг, А.А. Васильева, Ян Выбирал, Т. Ульрих и другие (см. [98]). Важные статьи и монографии по теории поперечников принадлежат также М.Ш. Бирману и М.З. Соломяку, Г.Г. Магарил-Ильяеву, Н.П. Корнейчуку, Ю.Н. Субботину, А. Пинкусу. Упомянем и направление, связанное с поперечниками аналитических функций (Г. Видом, Ч. Мичелли, К.Ю. Осипенко, Ю.А. Фарков, ...).

Двойственный к колмогоровскому поперечнику — гельфандовский — измеряет минимальный диаметр сечения тела подпространствами заданной коразмерности. Эта величина важна в теории восстановления.

Поперечники играют важную роль в геометрии банаховых пространств. Результат Кашина о поперечнике шара в терминах поперечника по Гельфанду влечёт существование подпространств $\mathbb{R}^N$ размерности $N(1-\delta)$, где $\delta > 0$ произвольно мало, в которых евклидова и ℓ_1-нормы пропорциональны: $c(\delta)N^{1/2}|x| \leqslant \|x\|_1 \leqslant N^{1/2}|x|$. Одновременно, Т. Фигелем, Дж. Линденштрауссом и В.Д. Мильманом был получен в определённом смысле двойственный результат о подпространствах ℓ_1^N размерности εN, где $\varepsilon > 0$ достаточно мало, с почти евклидовой нормой: $a(1-\varepsilon)N^{1/2}|x| \leqslant \|x\|_1 \leqslant a(1+\varepsilon)N^{1/2}|x|$. Эти результаты получили дальнейшее развитие в работах С. Шарека, Н. Томчак-Егерман, А. Пажора, М. Пизье, Р. Вершинина и других специалистов по асимптотическому геометрическому анализу. См. об этом книгу [87] и недавнюю работу [58].

В теории функциональных пространств изучаются s-числа операторов, определение которых в общем виде предложил Пич [125]. Частными случаями

s-чисел являются колмогоровские и гельфандовские числа: для оператора компактного вложения $\mathrm{Id}\colon X \to Y$ они равны соответствующим поперечникам шара $K = B(X)$ в пространстве Y. Другой пример связан с классической задачей о приближении компактных операторов конечномерными: аппроксимативные s-числа и соответствующие им линейные поперечники. Эти и другие s-числа служат важным инструментом изучения операторов и, в частности, вложений функциональных пространств. Существенный вклад в теорию s-чисел внесли также Б. Карл, Х. Трибель, Д.Е. Эдмундс.

Подчеркнём важную (особенно в последние годы) связь поперечников с компьютерными науками и смежными разделами математики. Задача о поперечнике конечного множества векторов в равномерной метрике эквивалентна приближению составленной из этих векторов матрицы в чебышёвской (равномерной) норме матрицами заданного ранга:

$$d_n(\{A^1,\dots,A^N\},\ell_\infty^M) = \min_{\operatorname{rank} B\leqslant n} \|A-B\|_C, \quad \text{где } \|X\|_C := \max_{\substack{1\leqslant i\leqslant M\\ 1\leqslant j\leqslant N}} |X_{i,j}|. \qquad (0.2)$$

Малоранговое приближение возникает в различных контекстах. Оно используется в численных методах для ускорения алгоритмов, в машинном обучении — для регуляризации. Е.Е. Тыртышников и его ученики разрабатывают численные алгоритмы малоранговой аппроксимации, см. [31]. Здесь проявляется связь теории приближений (нижних оценок) и теории сложности; приведём два примера. Минимальный ранг, нужный для приближения заданной булевой матрицы с фиксированной погрешностью (ε-ранг) играет важную роль в коммуникационной сложности; это хорошо представлено в обзорах [117, 115]. Отметим работы по этой теме А.А. Разборова [62] и Н. Алона с соавторами [86]. Близкие вопросы рассмотрены в работе [113]. Другой вариант задачи о малоранговом приближении использует метрику Хэмминга (количество отличий) вместо равномерной метрики. Л. Вэлиант [147] показал, что достаточно "жёсткие", т.е. плохо приближаемые (в метрике Хэмминга) матрицы дают пример (линейного) отображения с суперлинейной оценкой схемной сложности. Важная задача о построении конструктивного семейства таких матриц до сих пор не решена, хотя в недавнее время появился ряд прорывных работ по этой тематике [83, 99]. Некоторые методы и конструкции из теории сложности удалось использовать в этой диссертации. Ещё одна важная связь с поперечниками была обнаружена в 2000-х годах в теории сжатых измерений (Compressed Sensing), которая бы-

ла разработана Д. Донохо и Э. Кандесом, Т. Тао, Дж. Ромбергом, и другими математиками [101]. Ключевую роль в этой теории играют подпространства, хорошо приближающие евклидов шар в равномерной метрике: (0.1). Этим подпространствам соответствуют матрицы, позволяющие эффективно восстанавливать разреженные векторы высокой размерности с помощью малого числа линейных измерений (см. [44]). Наконец, задача о вычислении колмогоровского поперечника и поиске экстремальных подпространств есть частный случай задачи минимизации на многообразии Грассмана; численные алгоритмы такой минимизации изучались [100, 89], в том числе российскими математиками (В.М. Бухштабер [10], М.В. Балашов [6]).

Тихомиров в известном обзоре теории приближений [80] предложил разделить развитие этой теории на три этапа. В основе первого, чебышёвского этапа — приближение индивидуальных функций посредством полиномов и рациональных дробей. Второй этап — приближение классов функций (аналитических, дифференцируемых, липшицевых и т.п.) классическими методами (суммы Фурье, интерполяция, наилучшие полиномы, ...). На третьем этапе был поставлен вопрос о сравнении разных средств аппроксимации и отыскании оптимальных. Колмогоровский поперечник является одним из важнейших понятий "третьего этапа". Многие задачи о поперечниках ещё не решены и исследования по этой тематике продолжают оставаться актуальными. Полученные в диссертации результаты вносят вклад в развитие теории поперечников.

Нахождение точных значений поперечников — сложная задача; сделать это удаётся лишь в редких случаях. Чаще речь идёт о порядках убывания d_n, т.е. их нахождении с точностью до мультипликативной константы. Задача, таким образом, разделяется на две: верхние оценки — т.е. построение хорошей аппроксимации, и нижние оценки — доказательство, что "слишком хорошая" аппроксимация невозможна. Верхние оценки и используемые при их доказательстве методы полезны с точки зрения приложений; нижние оценки помогают обосновать оптимальность существующих методов, они необходимы для нахождения окончательного ответа в задаче. В диссертации делается упор на нижние оценки; как правило, там, где были получены новые порядки убывания поперечников, была улучшена именно нижняя оценка. Неформально, будем называть N-мерное множество "жёстким", если его нельзя хорошо приблизить подпространствами размерности существенно меньше N. Пример экстремаль-

но жёсткого множества даёт теорема Тихомирова о поперечнике шара: любое N-мерное сечение единичного шара в нормированном пространстве X нельзя приблизить в X с погрешностью меньше 1 никаким n-мерным подпространством, если $n < N$. Многие нижние оценки поперечников, полученные в диссертации, можно интерпретировать как некоторого рода "жёсткость".

Характерной чертой диссертации является последовательное использование теоретико–вероятностного подхода к поперечникам. Отметим, что в классической теории функций теоретико–вероятностный подход внедрили А.Н. Колмогоров, А.Я. Хинчин, В. Орлич, Р. Пэли, А. Зигмунд и успешно применяли Ж.П. Кахан, М. Талагран, С.В. Бочкарёв, Ж. Пизье, Б.С. Кашин, С.В. Асташкин и другие.

Цели и задачи исследования. Целью работы является развитие методов получения нижних оценок колмогоровских поперечников и близких аппроксимативных характеристик, а также решение ряда конкретных вопросов о поперечниках. В работе рассматриваются следующие задачи.

- Получить нижние оценки усреднённых поперечников случайных векторов в достаточно общей ситуации; с их помощью найти достаточные условия жёсткости (плохой приближаемости) конечных систем функций и конечномерных множеств.

- Вычислить порядки поперечников классов функций "критической" гладкости.

- Исследовать свойства поперечников классов дифференцируемых функций, специфичные для случая высокой гладкости.

- Найти порядки погрешности полилинейной аппроксимации некоторых классов функций нескольких переменных.

- Найти порядки поперечников шаров в смешанных нормах в новых случаях, ответить на некоторые вопросы о поперечниках конечномерных множеств.

Методы исследования. Применяются теоретико–вероятностные методы (неравенства больших уклонений, безусловные распределения), методы

функционального анализа (операторы, двойственность, интерполяция, тип пространства), гармонического анализа (классические ядра, теория Литлвуда–Пэли, всплески), выпуклой геометрии (оценки объёмов, энтропия многомерных тел, балансировка векторов), комбинаторики (комбинаторная размерность), методы из алгебры и некоторые результаты вещественной алгебраической геометрии.

Научная новизна. Результаты диссертации являются новыми. Основные из них состоят в следующем.

1. Получены нижние оценки усреднённых поперечников случайных векторов.

2. Получены достаточные условия жёсткости (слабая лакунарность, безусловность, независимость) для конечных систем функций. Установлено, что система Уолша не является жёсткой в пространстве L_q, $q < 2$.

3. Доказано, что безусловные координатно-однородные (инвариантные относительно транзитивного действия на множестве координат) множества являются жёсткими в ℓ_q^N, $q \leqslant 2$.

4. Найдены порядки колмогоровских поперечников класса Соболева W_1^1 в L_q, $2 < q < \infty$. Тем самым завершено решение классической задачи о порядках поперечников классов Соболева целой гладкости на отрезке. Найдены порядки поперечников классов Бесова в аналогичном случае.

5. Получены оценки относительных поперечников классов W_p^r в L_p для равномерной и интегральной метрик. В частности, показано, что невозможна аппроксимация этих классов n-мерными подпространствами с равномерно ограниченными нормами r-х производных приближающих функций и погрешностью, не более чем в константу раз большей n-поперечника.

6. Найдено асимптотическое поведение чебышёвских сплайнов, возникающих в задаче о поперечнике класса Соболева в равномерной метрике на отрезке.

7. Доказано, что классы равномерно ограниченных тригонометрических полиномов со спектром в параллелепипедах плохо приближаются комбина-

циями полилинейных функций. В качестве следствия получены новые порядки погрешности полилинейных приближений классов смешанной гладкости.

8. Доказано, что декартово произведение октаэдров плохо приближается в смешанной $(2,1)$-метрике. В качестве следствия получены порядки линейных поперечников классов Гёльдера–Никольского в определённых случаях. Найден критерий жёсткости для поперечников шаров в смешанных нормах.

9. Изучена функция распределения элементов матриц малого ранга, хорошо приближающих единичную матрицу.

Теоретическая и практическая ценность. Работа имеет теоретический характер. Полученные результаты могут найти применение в теории приближений, в функциональном и гармоническом анализе. Разделы диссертации могут составить содержание специальных курсов для студентов высших учебных заведений и аспирантов, обучающихся по специальности "Математика".

Апробация работы. Результаты диссертации докладывались на конференциях: Саратовская зимняя математическая школа (2016, 2018, 2024); Воронежская зимняя математическая школа (2017, 2023, 2025); Summer Symposium in Real Analysis XLII (Санкт-Петербург, 2018); High-dimensional approximation and discretization (Москва, 2018); Approximation, sampling and compression in data science (Cambridge, Великобритания, 2019); Approximation and Data Analysis (Нижний Новгород, 2019); Вторая конференция Математических центров России (Москва, 2022); Applied Functional Analysis (Oaxaca, Мексика, online, 2022); Школа-конференция С.Б. Стечкина по теории функций (Кыштым, 2023, 2024); Приближение, разложения и компьютерные науки (Сочи, 2023); Вероятностные методы в анализе и теория аппроксимации (Санкт–Петербург, 2024).

Были сделаны доклады по теме диссертации на научных семинарах: по теории приближений в МГУ, рук. проф. И.Г. Царьков, А.С. Кочуров, А.Р. Алимов, А.А. Васильева (2023, 2024), по теории приближений в МИАН, рук. проф. С.А. Теляковский (неоднократно, 2015–2019 гг.); по теории функций действительного переменного в МГУ, рук. акад. РАН Б.С. Кашин, акад. РАН С.В. Конягин, проф. П.А. Бородин, проф. М.И. Дьяченко (2018, 2021, 2022, 2025); по

теории функций многих действительных переменных и ее приложениям к задачам математической физики (Семинар Никольского) в МИАН, рук. чл.-корр. РАН О.В. Бесов (2016, 2025); по теории операторов и теории функций в ПОМИ, рук. акад. РАН С.В. Кисляков (2025); по геометрической теории приближений в МГУ, рук. проф. П.А. Бородин (2022); "Математика и ее приложения" — общеинститутский семинар МИАН (2023); "Бесконечномерный анализ и стохастика" в МГУ, рук. проф. В.И. Богачев, проф. С.В. Шапошников, Н.А. Толмачев, Е.Д. Косов (2024); объединённый семинар отделов теории приближения функций и теории аппроксимаций и приложений в ИММ УрО РАН (2024).

Публикации. Результаты, представленные в диссертации, опубликованы в 11 работах [155]– [165]. Статьи опубликованы в журналах, входящих в соответствующий перечень ВАК.

Личный вклад. Результаты, выносимые на защиту и составляющие содержание диссертации, — в том числе опубликованные в совместных статьях — получены автором самостоятельно.

Структура работы. Диссертация состоит из введения, пяти глав, заключения и списка литературы из 165 наименований. Главы делятся на параграфы, те, в свою очередь — на подпараграфы (пункты). Общий объём диссертации — 189 страниц.

Во введении мы приведём все доказанные в работе теоремы и некоторые их следствия. Для удобства в основном тексте глав формулировки теорем повторены (со всеми техническими деталями).

Глава 1 является центральной для диссертационной работы, методы и некоторые доказанные в этой главе утверждения используются в других главах.

Перейдём к обзору содержания глав.

Первая глава начинается с **§1.1** "Предварительные сведения", в котором даются полные определения используемых в диссертации понятий (в том числе тех, что фигурируют во введении), а также формулируются используемые при доказательствах утверждения.

В **§1.2** даётся определение усреднённого колмогоровского поперечника.

В классическом колмогоровском поперечнике ошибка аппроксимации понимается в "худшем случае" — берётся максимум погрешности по элементам приближаемого множества. Естественный и полезный метод оценки поперечников снизу состоит в замене максимума на среднее по некоторой мере. В вероятностных терминах — на математическое ожидание. Так мы приходим к следующему определению.

Усреднённым n-поперечником случайного элемента ξ в пространстве X называется величина

$$d_n^{\mathrm{avg}}(\xi, X) := \inf_{\dim Q_n \leqslant n} \mathsf{E}\rho(\xi, Q_n)_X \tag{0.3}$$

(где $\rho(x, M)$ это кратчайшее расстояние от точки до множества). Сказанное выше можно выразить в виде неравенства

$$d_n(K, X) \geqslant d_n^{\mathrm{avg}}(\xi, X), \quad \text{если } \mathsf{P}(\xi \in K) = 1. \tag{0.4}$$

Будут приведены некоторые базовые свойства этого поперечника.

Впервые определение усреднённого поперечника было дано в работе С.М. Воронина и Н.Т. Темиргалиева [13], они рассмотрели случай меры Банаха в L_2. Ряд результатов о поперечниках классов гладких функций в пространствах Винера, а также о поперечниках для гауссовой меры в ℓ_q^N, были получены Буслаевым [9], Майоровым [52, 53, 54], а также Я. Крёйцигом [94]. Ещё в конце 1950-х годов аппроксимацию в функциональных пространствах с мерой и усреднение ошибки по этой мере изучал А.В. Сульдин [72]. Подход "ошибки в среднем" систематически развивался в теории "Information Based Complexity" (IBC) такими математиками, как Дж. Трауб, Дж. Васильковский, Г. Вожняковский [141], Э. Новак и другими. Основной мотивацией в IBC и других упомянутых исследованиях было то, что в задаче аппроксимации может быть известно, какие приближаемые элементы встречаются чаще других (или важнее других), на них есть вероятностное распределение; в этом случае естественно вычислять ошибку в среднем, а не ошибку в наихудшем случае, как в классическом поперечнике.

Для нас мотивацией является неравенство (0.4) (хотя усреднённый поперечник интересен и сам по себе). Многие нижние оценки поперечников $d_n(K, X)$, полученные в диссертации, можно уложить в следующую схему: строится случайный вектор ξ, принимающий значения в K (обычно, в небольшом подмножестве K — например, часто работает равномерное распределение на множестве

крайних точек), оценивается усреднённый поперечник $d_n^{\mathrm{avg}}(\xi, X)$ и применяется неравенство (0.4). Многие идеи мы заимствовали из работ Глускина, который успешно применял метод усреднений для получения нижних оценок.

Обозначение (0.3) является новым, по этому поводу см. §1.2. Нам потребуется ещё несколько определений. Если $K = \{x_1, \dots, x_N\}$ — конечное множество, то через $d_n^{\mathrm{avg}}(K, X)$ обозначим усреднённый поперечник случайно и равномерно выбранной точки из K (т.е. минимизируется среднее арифметическое расстояний от этих точек). Наконец, если вместо усреднения берётся p-усреднение, $(\mathsf{E}\rho(\xi, Q_n)_X^p)^{1/p}$, то мы пишем $d_n^{\mathrm{avg}}(\xi, X)_p$.

Усреднённый поперечник не следует путать с поперечником в смысле средней размерности (φ-средним поперечником), введённым Г.Г. Магарил-Ильяевым, чтобы охватить случай классов в $L_q(\mathbb{R})$.

Основная часть главы начинается в **§1.3** с несложного, но общего результата.

Теорема 1.1. *Пусть $x_1, \dots, x_N$ — система элементов нормированного пространства X, система $x_1^*, \dots, x_N^* \in X^*$ является сопряжённой к $(x_k)_1^N$, и для некоторых $1 \leqslant s \leqslant 2$ и $B > 0$ выполнено неравенство:*

$$\| \sum_{k=1}^{N} a_k x_k^* \|_{X^*} \leqslant B (\sum_{k=1}^{N} |a_k|^s)^{1/s}, \quad \forall a_1, \dots, a_N \in \mathbb{R}. \tag{0.5}$$

Тогда при $1 \leqslant p \leqslant 2$ справедливо неравенство:

$$d_n^{\mathrm{avg}}(\{x_1, \dots, x_N\}, X)_p \geqslant B^{-1} \cdot \begin{cases} (1 - n/N)^{1/p}, & \text{если } s = 2,\ n < N, \\ c \min\{1, n^{-1/2} N^{1/s'}\}, & \text{если } s < 2,\ n \leqslant N/4. \end{cases}$$

Хорошо известно, что для любой ортонормированной системы $\varphi_1, \dots, \varphi_N$ в евклидовом пространстве E верно равенство

$$d_n(\{\varphi_1, \dots, \varphi_N\}, E) = (1 - n/N)^{1/2} \tag{0.6}$$

(обычно его записывают в виде (1.11)). Оценка снизу в (0.6) следует из работы Колмогорова, А.А. Петрова и Ю.М. Смирнова [46], оценка сверху была доказана А.И. Мальцевым [56]; в терминах поперечников равенство записал и применил Стечкин [67]. Приведём известное рассуждение, дающее оценку снизу — оно короткое и важное для нас, поскольку основано на усреднении. Без

ограничения общности считаем, что $E = \operatorname{span}\{\varphi_k\}_{k=1}^N$, так как всегда можно спроектировать на эту линейную оболочку. Для любого n-мерного подпространства $Q_n \subset E$ имеем (где $v_1, \dots, v_n$ это какой-то его ортонормированный базис, а proj — ортопроектор):

$$\sum_{k=1}^N \rho(\varphi_k, Q_n)_E^2 = N - \sum_{k=1}^N |\operatorname{proj}_{Q_n} \varphi_k|^2 = N - \sum_{j=1}^n \sum_{k=1}^N \langle \varphi_k, v_j \rangle^2 = N - n.$$

Таким образом, любая ортонормированная система функций $f_1, \dots, f_N$ является жёсткой в L_2 в том смысле, что n-поперечник $\{f_k\}$ отделён от нуля, если $n \leqslant N(1-\varepsilon)$. В работе [67] это обстоятельство применялось для оценок снизу поперечников функциональных классов в L_2 (по выражению Стечкина, он помещал внутрь класса "ежа" — систему $\{f_k\}$). Мы изучим, что происходит в других пространствах L_q.

Как мы увидим дальше, в L_q при $q < 2$ ортонормированная система иногда может быть хорошо приближена пространствами размерности $o(N)$. Теорема 1.1 позволяет получить достаточные условия жёсткости и в этом случае. Неравенство (0.5) при $s = 2$ превращается в хорошо известную p-лакунарность, или S_p-условие, для $p = q'$:

$$\| \sum_{k=1}^N a_k f_k \|_{q'} \leqslant B (\sum_{k=1}^N |a_k|^2)^{1/2}, \quad \forall a_1, \dots, a_N \in \mathbb{R}.$$

Следовательно, ортонормированные системы с этим свойством являются жёсткими в L_q. Глубокий результат Ж. Бургейна [92] утверждает, что из любой равномерно ограниченной ортонормированной системы можно выделить q'-лакунарную подсистему размера $N^{2/q'}$. Следовательно, при n, существенно меньших $N^{2/q'}$, хорошая аппроксимация равномерно ограниченных О.Н.С. невозможна.

Эта техника работает при $q \in (1, 2)$. Случай $q = 1$ сложнее и требует более ограничительных условий.

Теорема 1.2. *Пусть* $1 \leqslant q \leqslant 2$, *случайные величины* $\xi_1, \dots, \xi_N \in L_q(\Omega)$ *и* $\|\xi_i\|_q \geqslant 1$ *при* $i = 1, \dots, N$. *Предположим, что случайный вектор* $\xi := (\xi_1, \dots, \xi_N)$ *удовлетворяет хотя бы одному из условий:*

- $\xi_1, \dots, \xi_N$ *независимы и обладают нулевым средним;*

- *распределение вектора ξ безусловно (не меняется при изменении знака у любой координаты);*

- *ξ_i дизъюнктны: $\xi_i\xi_j \equiv 0$, $i \neq j$.*

Тогда при $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$, справедливо неравенство

$$d_n^{\mathrm{avg}}(\{\xi_1,\dots,\xi_N\}, L_q(\Omega))_q = N^{-1/q} d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant c(q,\varepsilon). \tag{0.7}$$

Таким образом, каждое из трёх условий (независимость, безусловность, дизъюнктность) обеспечивает жёсткость системы в L_q, $1 \leqslant q \leqslant 2$, включая случай $q = 1$. Простейшим примером независимой системы является система Радемахера.

Известно, что лакунарные системы $\{\varphi(k_j x)\}$ ведут себя похоже на независимые случайные величины, см., например, [20]. Свойство жёсткости не является исключением — в утверждении 1.13 мы доказываем жёсткость лакунарных систем.

Обратим внимание, что в теореме оцениваются два равных (с точностью до нормировки) поперечника. В левой части равенства в (0.7) усреднение происходит по конечному множеству в L_q, аналогично теореме 1.1; в правой части усреднение происходит по распределению случайного вектора в $\mathbb{R}^N$. Равенство этих двух поперечников — простой, но полезный факт, он позволяет получать следствия для поперечников множеств в $\mathbb{R}^N$. Приведём такое следствие.

Теорема 1.3. *Пусть группа H транзитивно действует на множестве N координат, и K — безусловное множество в $\mathbb{R}^N$, инвариантное относительно действия группы H, то есть, для любого элемента $h \in H$,*

$$(x_1,\dots,x_N) \in K \quad \Rightarrow \quad (\pm x_{h(1)},\dots,\pm x_{h(N)}) \in K.$$

Тогда при $1 \leqslant q \leqslant 2$ для любого $\varepsilon \in (0,1)$ справедливо неравенство

$$d_n(K, \ell_q^N) \geqslant c(q,\varepsilon) d_0(K, \ell_q^N) = c(q,\varepsilon) \sup_{x\in K} \|x\|_q, \quad \text{при } n \leqslant N(1-\varepsilon).$$

Неформально говоря, это означает, что безусловное координатно-однородное множество K является жёстким в ℓ_q^N. В качестве K можно взять любое множество, инвариантное относительно смен знаков и любых перестановок координат (или только циклических перестановок). Приведём

ещё ряд примеров: $V_k^N := B_\infty^N \cap kB_1^N$ ($k \in \mathbb{N}$) — в этом случае "жёсткость" доказал Глускин [27]; пересечение любого семейства шаров $r_\alpha B_{p_\alpha}^N$ ([14, 149]); шары смешанных норм и их пересечения ([12, 150]); см. также §5.1 .

Отметим, что при n, очень близких к N, поперечники таких множеств всё же могут быть малы — в этом отличие от случая теоремы Тихомирова о сечениях шара. Б. Грин [105] и А. Сах, М. Соуни, Я. Чжао в [131] рассмотрели поперечники орбиты Gx фиксированного единичного вектора под действием конечной подгруппы G ортогональной группы $\mathrm{O}(N)$ и доказали, что $d_{N-k}(Gx, \ell_2^N) \lesssim \log^{-1/2}(N/k)$ для любого такого множества при $1 \leqslant k \lesssim N/\log^C(3N)$.

При $q > 2$ аналог теоремы 1.2 не имеет места, см. утверждение 1.9. Пример поперечника $d_n(B_1^N, \ell_q^N)$ показывает (см. (1.15)), что и теорема 1.3 не верна в этом случае.

В основные результаты главы не вошли и другие утверждения, доказанные в этом параграфе. Кратко упомянем некоторые из них: оценка поперечников класса функций, определяемого условием $\sum k^\alpha |\langle f, f_k\rangle| \leqslant 1$ (следствие 1.1); жёсткость изотропных лог-вогнутых векторов (утверждение 1.6, п.2), оценки в L_q, $q > 2$ (утверждение 1.10 и его следствие), жёсткость независимых функций в метрике L_0, отвечающей за сходимость по мере (утверждения 1.11, 1.12).

В **§1.4** получены оценки сверху на поперечники вида $d_n(\{f_1, \dots, f_N\}, L_q)$ для двух классических систем функций — системы Уолша и тригонометрической системы.

Теорема 1.4. *Пусть $w_0, w_1, w_2, \dots$ — система Уолша в нумерации Пэли. Для любого $q \in [1, 2)$ существует $\delta = \delta(q) > 0$ такое, что справедливо неравенство*

$$d_n(\{w_0, \dots, w_{N-1}\}, L_q[0,1]) \leqslant C_q N^{-\delta}, \quad \text{при } n \geqslant N^{1-\delta}.$$

Кроме того,

$$d_n(\{w_0, \dots, w_{N-1}\}, L_q[0,1]) \geqslant c \min\{1, n^{-1/2} N^{1/q'}\}, \quad \text{при } n \leqslant N/4.$$

Мы видим, что некоторые равномерно ограниченные ортонормированные системы могут быть хорошо приближены пространствами малой размерности во всех L_q, $q < 2$. Это показывает содержательность результатов предыдущего параграфа.

Теорема 1.4 получена с помощью конструкции, возникшей в теории сложности в связи с близкой задачей о *жёсткости матриц*. Функция жёсткости

матрицы A определяется как расстояние Хэмминга от A до множества матриц заданного ранга:

$$\operatorname{Rig}(A,n) := \min_{\operatorname{rank} B \leqslant n} \#\{(i,j)\colon A_{i,j} \neq B_{i,j}\}.$$

Эта величина была введена Л. Вэлиантом [147] в 1977 г. как способ получения нижних оценок на схемную сложность — скажем, неравенство $\operatorname{Rig}(A_N, \varepsilon N) \geqslant N^{1+\varepsilon}$ с некоторым фиксированным $\varepsilon > 0$ влечёт суперлинейную оценку схемной сложности отображения $x \mapsto A_N x$, см. обзор [117]. Функция жёсткости выражается через усреднённый поперечник для метрики Хэмминга в $\mathbb{R}^M$ (для единообразия мы нормируем её единицей), аналогично (0.2):

$$\operatorname{Rig}(A,n) = MN d_n^{\mathrm{avg}}(\{A^1,\dots,A^N\}, L_{\mathrm{H}}^M).$$

Недавно был получен достаточно неожиданный результат [83]: Дж. Алман и Р. Виллиамс доказали, что семейство матриц Уолша–Адамара не является жёстким — их ранг может быть значительно уменьшен путём изменения небольшого количества элементов. Теорема 1.4 использует эту конструкцию.

Далее в параграфе рассматриваются метрика L_0 и функциональный аналог метрики Хэмминга: $\|f-g\|_{L_{\mathrm{H}}} := \mu\{x\colon f(x) \neq g(x)\}$. Отметим, что метрика L_{H} возникает в теории функций в контексте теорем об "исправлении". Так, классическая теорема Лузина утверждает, что C плотно в пространстве всех измеримых функций в метрике L_{H}, а теорема Меньшова — что плотно множество функций с равномерно сходящимся рядом Фурье. Другие примеры см. в работах Кашина [42], М.Г. Григоряна и В.Г. Кротова [29] и по ссылкам в этих работах.

Те же соображения, что в теореме 1.4, позволяют установить интересный эффект: для приближения N функций систем Уолша или тригонометрической системы в метрике L_0 достаточно пространств очень малой размерности (меньше любой степени N), см. следствие 1.5 и следствие 1.6.

Естественно задать вопрос: возможна ли аппроксимация как в теореме 1.4, для тригонометрической системы, хотя бы в пространстве L_1? К сожалению, удалось построить хорошую аппроксимацию лишь при $q < 1$. В этом случае возможно приближение гармоники другими гармониками:

Теорема 0.A ([32]). *Для любого $0 < q < 1$ равномерно по M и $N \geqslant M$,*

$$\inf_{\{c_k\}} \|\cos Mx - \sum_{|k|\leqslant N,\, k\neq M} c_k e(kx)\|_q \asymp (N-M+1)^{1-1/q}.$$

В утверждении 1.17 мы показываем, что при любом $q \in (0,1)$ можно хорошо приблизить N подряд идущих гармоник пространством, порождённым некоторыми $o(N)$ гармониками.

Глава 2. В этой главе изучаются поперечники классов дифференцируемых функций в случае граничной, "критической" гладкости.

Для $r \in \mathbb{N}$ и $1 \leqslant p \leqslant \infty$ через $W_p^r[a,b]$ обозначим класс Соболева, состоящий из функций $f\colon [a,b] \to \mathbb{R}$, абсолютно непрерывных вместе с производными $f', \dots, f^{(r-1)}$, для которых $\|f^{(r)}\|_p \leqslant 1$. Классы 2π-периодических функций $W_p^r(\mathbb{T})$ определяются аналогично.

Колмогоров в работе 1936 года не только ввёл понятие поперечника, но и вычислил точные значения поперечников классов $W_2^r[0,1]$ и $W_2^r(\mathbb{T})$ в метрике L_2. Оказалось, что

$$d_{2n-1}(W_2^r(\mathbb{T}), L_2) = d_{2n}(W_2^r(\mathbb{T}), L_2) = n^{-r}$$

и экстремальным пространством является пространство $\mathcal{T}_{n-1}$ тригонометрических полиномов порядка меньше n. В случае отрезка ответ был дан в терминах спектра специального дифференциального оператора.

Нетрудно видеть, что порядки убывания поперечников в периодическом и непериодическом случае совпадают; в этом контексте мы можем записывать класс просто как W_p^r.

Тем самым, возникла задача: найти порядки поперечников $d_n(W_p^r, L_q)$ при $r \in \mathbb{N}$, $1 \leqslant p, q \leqslant \infty$.

В **§2.1** доказывается следующая теорема.

Теорема 2.1. *Для любого $q \in (2, \infty)$ справедливо порядковое равенство*

$$d_n(W_1^1, L_q) \asymp n^{-1/2} \log n.$$

(Запись $A \lesssim B$ означает $A = O(B)$; мы пишем $A \asymp B$, если $A \lesssim B$ и $A \gtrsim B$.)

Этот результат завершает решение упомянутой задачи. Ответ в ней полу-

чился таким:

$$d_n(W_p^r, L_q) \asymp \begin{cases} n^{-r}, & \text{если } p \geqslant q \text{ или } 2 < p < q, \\ n^{-r-1/2+1/p}, & \text{если } p \leqslant 2 < q,\ rp > 1, \\ n^{-r-1/q+1/p}, & \text{если } p < q \leqslant 2, \\ n^{-1/2}\log n, & \text{если } r = p = 1,\ 2 < q < \infty. \end{cases}$$

Поперечник класса W_1^1 в L_∞, очевидно, равен бесконечности.

Поперечник можно оценить сверху погрешностью приближения тригонометрическими полиномами: $\rho(W_p^r(\mathbb{T}), \mathcal{T}_n)_q \asymp n^{-r+(1/p-1/q)_+}$ при всех p, q. Это порядковое равенство давно известно специалистам; для оценок используются классические методы Фурье, Валле Пуссена, Фавара; см. [79, §3.2]. (Такая оценка поперечника не всегда точна, см. далее.)

Задача о поперечниках классов Соболева обладает большой историей. После случая $p = q = 2$, рассмотренного Колмогоровым, появилась работа У. Рудина [130], где, по сути, были найдены порядки поперечников $d_n(W_1^1, L_2)$ и $d_n(W_\infty^1, L_2)$. Стечкин [67] нашёл порядки для $p = \infty$, $2 \leqslant q \leqslant \infty$ (и всех r). Лоренц [118] получил точную по порядку нижнюю оценку для $p = \infty$, $q = 1$; отсюда следуют порядки для всего "верхнего треугольника" $p \geqslant q$ (название связано с традицией рисовать множество параметров (p, q) как единичный квадрат с осями $(1/p, 1/q)$, см. рис. 1). Независимо, порядки (а также некоторые точные значения) в верхнем треугольнике были найдены Тихомировым [77, 79] и Маковозом [55]. М.З. Соломяк и Тихомиров [68] разобрали (в двойственной формулировке) случай $p = 1$, $q \leqslant 2$. Исмагилов [33] получил точные оценки снизу для $1 \leqslant p < q = 2$ и позже [34] при $1 \leqslant p < q \leqslant 2$. В итоге оказалось, что при $q \leqslant \max\{p, 2\}$ тригонометрические полиномы оптимальны; открытым оставался случай "трапеции" $q > \max\{p, 2\}$.

Исмагилов в работе [34] обнаружил, что для трапеции тригонометрические полиномы не обязательно оптимальны, в частности, $d_n(W_1^2, L_\infty) \lesssim n^{-6/5}\log n$. Первый ответ в трапеции (для этого же поперечника) нашёл Глускин [23]. Прорыв в этой задаче был совершён Кашиным в известной работе [39] 1977 года, он нашёл порядки для почти всех оставшихся случаев, включая принципиальный случай $p = 2$, $q = \infty$.

Оставался открытым только один случай: $r = p = 1$, $2 < q < \infty$. Название "критическая гладкость" связано с тем, что при $rp > 1$ пространство Соболева

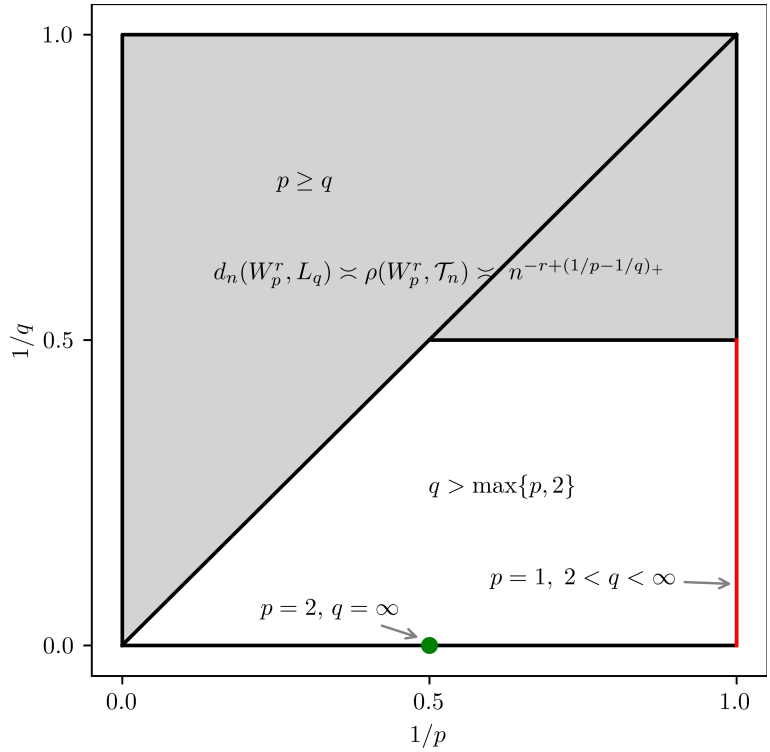


Рис. 1: Параметры (p, q) в задаче о поперечниках $d_n(W_p^r, L_q)$.

$\mathcal{W}_p^r[a, b]$ компактно вложено в $C[a, b]$, а при $r = p = 1$ это вложение нарушается. Е.Д. Куланин [50] получил следующие оценки:

$$n^{-1/2} \log^{1/2-\varepsilon} n \lesssim d_n(W_1^1, L_q) \lesssim n^{-1/2} \log n, \tag{0.8}$$

для любого фиксированного $\varepsilon > 0$; оказалось, что порядок поперечников содержит логарифмический множитель. Э.С. Белинский [7] (см. также [145, Ch.10]) доказал, что верхняя оценка точна, если рассматривать только подпространства, порождённые произвольными n гармониками: $d_n^T(W_1^1, L_q) \asymp n^{-1/2} \log n$. (Определение тригонометрического поперечника d_n^T см. в §1.4.) Позже, в работе [163] удалось убрать ε из (0.8), но логарифмический зазор оставался. Наконец, в работе автора [165] была доказана теорема 2.1 и тем самым найден порядок поперечника. Подчеркнём, что здесь $r \in \mathbb{N}$; о случае дробной гладкости см. заключение.

В конце параграфа приведены некоторые конечномерные следствия теоремы 2.1, например, (2.11).

Параграф **§2.2** посвящён близким к W_1^1 классам Бесова. Пусть заданы па-

раметры $r > 0$, $p, \theta \in [1, \infty]$. Рассмотрим в $L_p[0,1]$ полунорму (где $s = 1 + \lfloor r \rfloor$):

$$|f|_{B^r_{p,\theta}} := \begin{cases} \left(\int_0^\infty (t^{-r}\omega_s(f,t)_p)^\theta \frac{dt}{t}\right)^{1/\theta}, & 1 \leqslant \theta < \infty, \\ \sup_{t>0} t^{-r}\omega_s(f,t)_p, & \theta = \infty, \end{cases}$$

где $\omega_s(f,t)_p$ — обычный модуль гладкости в L_p. Класс Бесова $B^r_{p,\theta} := B^r_{p,\theta}[0,1]$ состоит из функций, для которых $\|f\|_p + |f|_{B^r_{p,\theta}} \leqslant 1$.

Хорошо известно, что имеют место вложения пространств

$$\mathcal{B}^1_{1,1}[0,1] \subset \mathcal{W}^1_1[0,1] \subset \mathcal{B}^1_{1,\infty}[0,1]. \tag{0.9}$$

Порядки поперечников классов Бесова были получены в работах Динь Зунга [30] (верхние оценки), А.С. Романюка [63, 64], Галеева [18] во всех случаях $1 < p, q < \infty$, кроме $q < p < 2$ и $p \geqslant 2 > q$, $\theta < 2$ (см. [18]). Также есть результаты о граничных случаях $p, q \in \{1, \infty\}$, см., например, [65].

В связи с классом W^1_1 и соотношением (0.9) был рассмотрен случай $r = p = 1$ для классов Бесова.

Теорема 2.2. *При $2 < q < \infty$ и $1 \leqslant \theta \leqslant \infty$ справедливо порядковое равенство*

$$d_n(B^1_{1,\theta}, L_q) \asymp \begin{cases} n^{-1/2}\log^{1/2} n, & 1 \leqslant \theta \leqslant 2, \\ n^{-1/2}\log^{1-1/\theta} n, & 2 \leqslant \theta \leqslant \infty. \end{cases}$$

Глава 3. В случае гладкости $r > 1$ порядки поперечников классов Соболева W^r_p давно известны. В этой главе рассматриваются две задачи, связанные с более тонкими аппроксимативными свойствами этих классов.

§3.1 посвящён относительным поперечникам в интегральной и равномерной метриках. Относительный поперечник был введён В.Н. Коноваловым [47]; это вариант колмогоровского поперечника, в котором множество приближается не всем подпространством, а его пересечением с другим множеством. Формально, для $W, V \subset X$ определим n-поперечник W относительно V равенством

$$d_n(W, V, X) := \inf_{\substack{Q_n \subset X \\ \dim Q_n \leqslant n}} \rho(W, Q_n \cap V)_X.$$

Нас будут интересовать поперечники $d_n(W^r_p, MW^r_p, L_p)$. Параметр $M \geqslant 1$ ограничивает аппроксимацию теми функциями $g \in Q_n$, у которых $\|g^{(r)}\|_p \leqslant M$. Мы докажем две теоремы.

Теорема 3.1. *Пусть $r, n \in \mathbb{N}$, $r \geqslant 3$. Для всех $\varepsilon \in [n^{-2}, 1]$ справедливо неравенство:*

$$d_n(W_\infty^r(\mathbb{T}), (1+\varepsilon)W_\infty^r(\mathbb{T}), L_\infty(\mathbb{T})) \geqslant c(r)n^{-r}\varepsilon^{1-r/2}.$$

Теорема 3.2. *Пусть $n, r \in \mathbb{N}$, $r \geqslant 3$, $M \geqslant 2$; $p = 1$ или $p = \infty$.*

1. Для $\delta_1 := \exp(-c_1 M)$ и всех $n \geqslant \delta_1^{-1}$ имеет место неравенство

$$d_{2n}(W_p^r(\mathbb{T}), MW_p^r(\mathbb{T}), L_p(\mathbb{T})) \geqslant \frac{1}{8}\left(\frac{1+\delta_1}{n}\right)^r.$$

2. Для $\delta_2 := \exp(-c_2 M)$ и всех $n \geqslant C\delta_2^{-1}$ справедливо неравенство

$$d_{2n-1}(W_p^r(\mathbb{T}), MW_p^r(\mathbb{T}), L_p(\mathbb{T})) \leqslant \mathcal{K}_r\left(\frac{1+C\delta_2}{n}\right)^r.$$

($\mathcal{K}_r$ — константа Фавара, см. §1.1.)

Эти два результата, в основном, завершают решение задачи об относительных поперечниках в случаях $p = 1$ и $p = \infty$. Опишем получившуюся картину и одновременно историю вопроса. Начнём для определённости с $p = \infty$. Обозначим

$$d_n := d_n(W_\infty^r(\mathbb{T}), L_\infty(\mathbb{T})), \quad K_n := d_n(W_\infty^r(\mathbb{T}), MW_\infty^r(\mathbb{T}), L_\infty(\mathbb{T})).$$

(Ясно, что d_n и K_n зависят ещё от параметра r, а K_n зависит ещё и от M.) Поперечник d_n был найден Тихомировым [77, 78]:

$$d_{2n-1} = d_{2n} = \mathcal{K}_r n^{-r}. \tag{0.10}$$

При $r = 1, 2$ поперечники совпадают: $K_n = d_n$ (В.Н. Малозёмов, Ю.Н. Субботин; см. [69]). Пусть $r \geqslant 3$; будем двигаться в сторону увеличения M. Случай $M = 1$, с которого и началось исследование относительных поперечников, был изучен Коноваловым в работе [47]. Им был обнаружен интересный эффект порядкового отличия относительного ($K_n \asymp n^{-2}$) и колмогоровского ($d_n \asymp n^{-r}$) поперечников:

- Если $M = 1$, то $c_1(r)n^{-2} \leqslant K_n \leqslant c_2(r)n^{-2}$.

(То же верно при $M \leqslant 1 + 1/n^2$.) Позже простое и изящное доказательство результата Коновалова было дано Тихомировым [140].

"Фазовый переход" между случаями $M = 1$ и $M > 1$ описывается следующим образом.

- Если $M = 1 + \varepsilon$, где $1/n^2 \leqslant \varepsilon \leqslant 1$, то

$$c_1(r)n^{-r}\varepsilon^{1-r/2} \leqslant K_n \leqslant c_2(r)n^{-r}\varepsilon^{1-r/2}. \tag{0.11}$$

Верхняя оценка в (0.11) была получена В.Ф. Бабенко в [4] с помощью приближения сплайнами. Нижняя оценка составляет теорему 3.1. Из верхней оценки вытекает, что при фиксированном $M > 1$ порядки колмогоровских и относительных поперечников совпадают:

- При фиксированном $M > 1$ имеем $K_n \asymp n^{-r}$.

(Тем самым Бабенко ответил на вопрос Стечкина.)

Ю. Н. Субботин и С. А. Теляковский в работе [69] изучили вопрос о том, когда относительные и колмогоровские поперечники в точности совпадают.

- При достаточно большом M поперечники равны:

$$d_n = K_n, \quad \text{если } M = \tfrac{4}{\pi^2}\ln\min(n, r) + C_0. \tag{0.12}$$

Это потребовало весьма тонких оценок норм производных ядер Фавара. Оставался открытым вопрос, можно ли в качестве M взять абсолютную постоянную. Другими словами, возможно ли приближение класса W^r_∞ линейными подпространствами заданной размерности, при котором 1) погрешность приближения не более чем в фиксированное число раз хуже, чем колмогоровский поперечник и 2) нормы r-х производных приближающих функций равномерно ограничены? (При любом фиксированном $p \in (1, +\infty)$ обычный проектор Фурье даёт такое приближение в силу теоремы Рисса.) Отрицательный ответ на этот вопрос следует из п.1 теоремы 3.2. Действительно, оценка поперечника снизу гарантирует следующее.

- Если $M = c\ln r$, где c достаточно мало, то при достаточно больших r: $\varliminf\limits_{n\to\infty} K_n/d_n \geqslant \exp(r^{1/2})$.

В случае L_1 всё аналогично. Значение колмогоровского поперечника такое же, как в равномерной метрике. Для чётных n это доказали независимо Маковоз и Субботин, а при нечётных n — В.И. Рубан (см. [69]). Равенство поперечников для $r = 1, 2$ и отличие порядков при $M = 1$ установил Бабенко [3].

Фазовый переход (0.11) выглядит так же; верхняя оценка доказана в [5], нижняя оценка получена Н.В. Парфинович [123] (отметим, что теорема 3.1 доказана автором независимо и другим методом). Утверждение о равенстве поперечников и теорема 3.2 выглядят одинаково и доказываются одновременно для $p = 1$ и $p = \infty$.

Отметим также работу Коновалова [48].

Ключевым элементом доказательства теоремы 3.2 является теорема Ал.А. Привалова (см. теорему 3.A) о норме проектора $C(\mathbb{T}) \to \mathcal{T}_n$ на подпространстве заданной коразмерности. Мы переносим результат Привалова с равномерной метрики на интегральную (см. утверждение 3.1).

Следующий параграф, **§3.2**, посвящён чебышёвским сплайнам. Тихомиров в работе [78] выразил колмогоровские поперечники класса W^r_∞ на отрезке в равномерной метрике через норму так называемых чебышёвских сплайнов:

$$d_{n+r}(W^r_\infty[-1,1], L_\infty[-1,1]) = \|x_{n,r}\|_\infty, \quad n \in \mathbb{Z}_{\geqslant 0}. \tag{0.13}$$

Сплайн $x_{n,r}$ определяется как наименее уклоняющаяся от нуля на $[-1,1]$ функция, у которой r-я производная равна по модулю единице и меняет знак в n точках внутри $[-1,1]$. Ясно, что при $n = 0$ получается последовательность многочленов Чебышёва: $x_{0,r} = T_r/(r!2^{r-1})$; отсюда название *чебышёвский сплайн*. Как устроена последовательность $\{x_{n,r}\}_{r\geqslant 1}$ при фиксированном $n > 0$? Во сколько раз (асимптотически, при $r \to \infty$) уменьшится норма по сравнению с $x_{0,r}$, если разрешить n "изломов", и где они будут располагаться? Мы ответим на эти вопросы дальше.

Для формулировки теоремы потребуется следующее определение. Пусть $\{\psi_j\}_{j=1}^M$ — система функций в $L(0,\infty)$. Набор точек $\{x_i\}_{i=1}^K$, $0 < x_1 < x_2 < \ldots < x_K < \infty$, называется *каноническим набором* для $\{\psi_j\}$, если выполняются соотношения (в которых полагаем $x_0 := 0$, $x_{K+1} := \infty$):

$$\sum_{k=0}^{K} (-1)^k \int_{x_k}^{x_{k+1}} \psi_j(x)\,dx = 0, \quad j = 1, \ldots, M.$$

Сплайн $x_{n,r}$ симметричен относительно нуля и его узлы тоже; при четном $n = 2n'$ он имеет n' положительных узлов $\{\tau_j\}_{j=1}^{n'}$ и n' отрицательных $\{-\tau_j\}_{j=1}^{n'}$; при нечетном $n = 2n' + 1$ имеем n' положительных узлов, n' отрицательных и узел $\tau = 0$. (Эти и другие свойства сплайна изложены в §3.2.)

Теорема 3.3. *Пусть $n \geqslant 1$ фиксировано, $n' := \lfloor n/2 \rfloor$. При $n \geqslant 2$ положительные узлы $\{\tau_j\}$ сплайнов $x_{n,r}$ удовлетворяют асимптотике при $r \to \infty$:*

$$\tau_k(x_{n,r}) = \varkappa_k r^{-1/2}(1 + o(1)), \quad k = 1, \ldots, n',$$

где числа $0 < \varkappa_1 < \ldots < \varkappa_{n'} < \infty$ образуют канонический набор для системы $\{x^{n-2q}e^{-x^2}\}_{q=1}^{n'}$ в $L(0,\infty)$. Имеет место асимптотика при $r \to \infty$:

$$\frac{\|x_{n,r}\|_\infty}{\|x_{0,r}\|_\infty} = A_n r^{-n/2}(1 + o(1)),$$

где числа $A_n > 0$ определяются следующим образом: $A_1 = 1/\sqrt{\pi}$,

$$A_n = (-1)^{n'-1}\frac{2}{\sqrt{\pi}}\sum_{k=1}^{n'}(-1)^{k-1}\varkappa_k^{n-1}e^{-\varkappa_k^2}, \quad n \geqslant 2.$$

Что это означает для поперечников? При фиксированном r известна асимптотика

$$d_n(W_\infty^r[-1,1], L_\infty) = \left(\frac{2}{\pi}\right)^r \mathcal{K}_r n^{-r}(1 + o(1)), \quad n \to \infty.$$

Это соответствует периодическому случаю, в котором значение поперечника вычисляется явно, см. (0.10). Другими словами, при $n \to \infty$ и фиксированном r "концевые эффекты" пропадают, и рассматриваемая нами задача оказывается эквивалентна периодической. Однако в общем случае при произвольных n и r, формула (0.13) не дает исчерпывающего представления о поведении величины $d_n(W^r) := d_n(W_\infty^r, L_\infty)$. Рассмотрим, например, следующий вопрос. Ясно, что $d_n(W^r) = \infty$ при $n < r$ и $d_r(W^r) = \|x_{0,r}\| = (r!2^{r-1})^{-1}$. Что можно сказать о $d_{r+1}(W^r)$ при больших r? Этот вопрос задавался в книге [79, §2.5.4]. Из сказанного выше вытекает

Следствие. *Для фиксированного $n \geqslant 1$ имеем*

$$\frac{d_{n+r}(W_\infty^r, L_\infty)}{d_r(W_\infty^r, L_\infty)} = A_n r^{-n/2}(1 + o(1)), \quad \textit{при } r \to \infty.$$

Проделанные нами вычисления являются шагом к явным оценкам величины поперечника (0.13) при произвольных N и r; по-видимому, получение хороших оценок в общем случае представляет собой трудоемкую задачу.

Сплайны $x_{n,r}$ используются для построения совершенных сплайнов на полупрямой, которые также обладают важными экстремальными свойствами, см. недавнюю работу А.Р. Акопяна, В.В. Арестова и В.Г. Тимофеева [1].

Глава 4 посвящена некоторым оценкам погрешности полилинейных приближений. Рассмотрим функции d переменных, $f\colon I^d \to \mathbb{R}$, где для определённости $I = [a, b]$ или $I = \mathbb{T}$. Полилинейные приближения — это приближения вида

$$f(x_1, \dots, x_d) \approx \sum_{k=1}^{n} u^{1,k}(x_1) u^{2,k}(x_2) \cdots u^{d,k}(x_d) \tag{0.14}$$

или, по современной терминологии, n-членные приближения по словарю полилинейных функций. В дискретном варианте речь идёт о тензорах $\mathbb{R}^{m_1 \times \cdots \times m_d} = \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2} \otimes \cdots \otimes \mathbb{R}^{m_d}$ и приближении тензорами ранга n. О вычислительных аспектах в тензорах и различных вариантах понятия ранга см. книгу [106].

Обозначим через $\Theta_n(f)_p$ величину наилучшего полилинейного приближения вида (0.14) в метрике L_p. Для класса функций положим $\Theta_n(\mathcal{F}, L_p) := \sup_{f\in\mathcal{F}} \Theta_n(f, L_p)$.

При $d = 2$ получается билинейное приближение. Классическая теорема Шмидта выражает величину $\Theta_n(f)_2$ через сингулярные числа интегрального оператора с ядром $f(x_1, x_2)$. Отметим ещё связь билинейного приближения с поперечниками. Мы уже видели, что поперечник конечных множеств в ℓ_∞^N равен погрешности аппроксимации матрицами заданного ранга: (0.2). На языке функций это означает, что поперечник семейства $\mathcal{F} := \{f(\cdot, y)\colon y \in I\}$, параметризованного переменной y и построенного по функции $f\colon I^2 \to \mathbb{R}$, равен погрешности билинейной аппроксимации f:

$$d_n(\mathcal{F}, L_\infty(I)) = \Theta_n(f)_\infty.$$

В **§4.1, §4.2** мы показываем, что большинство $m_1 \times \cdots \times m_d$ тензоров с элементами ± 1 являются “жёсткими” в том смысле, что они сильно отличаются от тензоров вида $\operatorname{sign} T$ с рангом существенно меньше максимально возможного: $\operatorname{rank} T \lesssim \prod m_j / \max m_j$. Точную формулировку см. в утверждении 4.2. В доказательстве используется алгебро–геометрический результат Шапиро–Уоррена [152] о количестве векторов, составленных из знаков многочленов от нескольких переменных. Отметим, что матричные методы часто не работают для тензоров (при $d > 2$); например, нет хорошего аналога сингулярного разложения (SVD). Оказалось, что подход Шапиро–Уоррена, применённый в случае матриц ($d = 2$) в работе [85], работает и для тензоров. Этот подход применим и для других тензорных форматов, например, для тензорного поезда ([122]).

Используя жёсткость тензоров и подходящую дискретизацию, в **§4.3** мы получаем нижние оценки на погрешность полилинейных приближений функциональных классов.

Обозначим через $\mathcal{T}_{\mathbf{m}}$, где $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$, множество вещественных тригонометрических полиномов с со спектром в $\prod_{j=1}^d [-m_j, m_j]$. Пусть также B_∞ — единичный шар $L_\infty(\mathbb{T}^d)$.

Теорема 4.1. *Для любых $d \in \mathbb{N}$, $d \geqslant 2$, $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$ имеем*

$$\Theta_n(\mathcal{T}_{\mathbf{m}} \cap B_\infty, L_1(\mathbb{T}^d)) \geqslant c_2(d), \quad \text{при } n \leqslant c_1(d) \frac{\prod_{j=1}^d m_j}{\sum_{j=1}^d m_j}. \tag{0.15}$$

Мы применим этот результат к классам функций смешанной гладкости $W_p^{\mathbf{r}}(\mathbb{T}^d)$, $\mathbf{r} = (r_1, \dots, r_d)$, $r_j > 0$. Для целых гладкостей такой класс можно определить условием

$$\left\| \frac{\partial^{r_1}}{\partial x_1^{r_1}} \cdots \frac{\partial^{r_d}}{\partial x_d^{r_d}} f(x_1, \dots, x_d) \right\|_p \leqslant 1.$$

В общем случае определение см. в §4.3. По поводу классов смешанной гладкости см. монографию [73] и книги [98, 139]. Нас интересует изотропный случай $\mathbf{r} = r\mathbf{1} = (r, r, \dots, r)$.

Полилинейные приближения классов $W_p^{\mathbf{r}}$ изучались В.Н. Темляковым в работах [74, 76] ($d = 2$) и [75] ($d \geqslant 2$), а также Темляковым и Д. Базархановым [88] ($d > 2$). Другие тензорные форматы рассматривались в [133].

В работе [75] была получена оценка сверху:

$$\Theta_n(W_2^{\mathbf{r}}, L_2(\mathbb{T}^d)) \lesssim n^{-\frac{rd}{d-1}}.$$

Используя теорему 4.1, мы получаем точную по порядку нижнюю оценку

$$\Theta_n(W_\infty^{\mathbf{r}}, L_1(\mathbb{T}^d)) \gtrsim n^{-\frac{rd}{d-1}}.$$

Объединяя эти неравенства, получим

Следствие 4.1. *Для любых $d \in \mathbb{N}$, $d \geqslant 2$, $r > 0$, $\mathbf{r} := (r, \dots, r)$, $2 \leqslant p \leqslant \infty$, $1 \leqslant q \leqslant 2$ справедливо порядковое равенство*

$$\Theta_n(W_p^{\mathbf{r}}, L_q(\mathbb{T}^d)) \asymp n^{-\frac{rd}{d-1}}.$$

Отметим, что случай $d = 2$ хорошо изучен, теорема 4.1 и порядки в следствии 4.1 при $d = 2$ известны, см. [76]. Основную новизну результатов этой главы составляет случай $d > 2$.

В случае $d = 2$ в [76] был получен более общий результат о приближении класса $\mathcal{T}_{\mathbf{m}} \cap B_\infty$ суммами функций вида $u(x_1, \dots, x_s)v(x_{s+1}, \dots, x_d)$. В конце параграфа доказывается утверждение 4.7 (предложенное Темляковым), которое является обобщением этого результата и теоремы 4.1.

Глава 5 посвящена нескольким избранным задачам о поперечниках конечномерных множеств. Параграфы этой главы независимы друг от друга.

§5.1 посвящён поперечникам шаров в смешанной норме. Хорошо известно, что оценки поперечников функциональных классов часто сводятся к оценкам поперечников специальных конечномерных множеств. Так, при дискретизации классов Соболева W_p^r возникают ℓ_p^N-шары. Галеев [15] исследовал поперечники некоторых классов функций многих переменных и для них получил нижние оценки через поперечники шаров в смешанных нормах.

Дадим определение смешанной нормы в интересующем нас случае. Пусть $N = sb$ и координаты пространства $\mathbb{R}^N$ разбиты на b блоков размера s. Для $x \in \mathbb{R}^N$ через $x[j]$ обозначим вектор, образованный координатами из j-го блока. Через $\ell_{q_1,q_2}^{s,b}$ обозначим пространство $\mathbb{R}^N$ с нормой

$$\|x\| := \Big\| (\|x[1]\|_{\ell_{q_1}^s}, \dots, \|x[b]\|_{\ell_{q_1}^s}) \Big\|_{\ell_{q_2}^b}.$$

Единичный шар этого пространства обозначаем $B_{q_1,q_2}^{s,b}$.

В работе [15] Галеев доказал точное по порядку неравенство:

$$d_{N/2}(B_{1,\infty}^{s,b}, \ell_{2,q}^{s,b}) \geqslant c_q b^{1/q}, \quad 1 < q < \infty.$$

Он поставил вопрос о случае $q = 1$; А.Д. Изаак [35] получил оценку

$$d_{N/2}(B_{1,\infty}^{s,b}, \ell_{2,1}^{s,b}) \geqslant cb\frac{\sqrt{\log\log b}}{\log b}. \tag{0.16}$$

При этом, как нетрудно видеть, $d_0(B_{1,\infty}^{s,b}, \ell_{2,1}^{s,b}) = b$.

Нами получен точный по порядку результат:

Теорема 5.1. *Для всех $s, b \in \mathbb{N}$, $\varepsilon \in (0,1)$, $N := sb$, имеет место неравенство*

$$d_n(B_{1,\infty}^{s,b}, \ell_{2,1}^{s,b}) \geqslant c(\varepsilon)b, \quad \textit{при } n \leqslant N(1-\varepsilon).$$

Из теоремы 5.1 получаются порядковые равенства для линейных поперечников классов Гёльдера–Никольского. Необходимые определения см. в §5.1.

Следствие 5.2. *Пусть $d \in \mathbb{N}$, $r > 0$, $\mathbf{r} = (r, r, \dots, r)$. Справедливы порядковые равенства:*

$$\lambda_n(H_p^{\mathbf{r}}(\mathbb{T}^d), L_q) \asymp \begin{cases} \left(\frac{\log^{d-1} n}{n}\right)^{r-1/2+1/q} \log^{(d-1)/q} n, & \frac{1}{p} + \frac{1}{q} < 1,\ p \leqslant 2,\ r > 1 - \frac{1}{q}, \\ \left(\frac{\log^{d-1} n}{n}\right)^{r-1/p+1/q} \log^{(d-1)/q} n, & 2 \leqslant p < q,\ r > \frac{1}{p} - \frac{1}{q}. \end{cases}$$

Приведём теперь известные результаты о поперечниках шаров в смешанных нормах для других наборов параметров (p_1, p_2, q_1, q_2). Большинство результатов относятся к случаю "жёсткости"; точнее говоря, доказывались соотношения вида $d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \asymp d_0(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2})$ при $n \leqslant N/2$. Галеев [16]: (p_1, ∞, q_1, q_2), где $p_1 \in \{1, \infty\}$; $1 < q_2 \leqslant \infty$; $q_1 = 2$ или $1 < q_1 \leqslant \min\{q_2, 2\}$. Позже, в работе [18] он добавил случай $(\infty, 1, q, q)$, $1 < q \leqslant 2$. А.Д. Изаак [36] упоминает случай $(p, p, 2, 1)$, $p \in \{1\} \cup [2, \infty]$. А.А. Васильева [150] разобрала случай $p_1 \geqslant q_1$, $p_2 \geqslant q_2$.

Порядки поперечников известны и в некоторых других случаях. Васильева [148, 150] получила решение в случае $2 \leqslant q_1, q_2 < \infty$. С. Дирксен и Т. Ульрих [96] рассмотрели поперечники по Гельфанду в смешанных нормах, интересуясь в основном полунормированным случаем $(0 < p \leqslant 1)$; по двойственности из их работы получаются порядки для а) $(2, p, 2, \infty)$, $p \geqslant 2$; b) (p, p, ∞, p), $p \geqslant 2$.

Эти оценки поперечников $d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2})$ применялись в изучении аппроксимативных свойств разных функциональных классов: Гёльдера–Никольского (см. [15, 17, 36, 158]), классов Бесова смешанной гладкости ([18, 96]), весовых классов Бесова [148]. Отметим, что смешанные нормы возникают в поперечниках классов Шаттена, см. [109, 128]. Обзор результатов по поперечникам шаров в смешанных нормах можно также найти в статьях [19, 12].

Известно, что имеет место "дихотомия": при $q > \max\{p, 2\}$ шары B_p^N хорошо приближаются маломерными пространствами (см. (1.16)): $d_0(B_p^N, \ell_q^N) = 1$, но

$$d_n(B_p^N, \ell_q^N) \leqslant C(p, q) N^{-\delta}, \quad \text{если } n \geqslant N^{1-\delta},\ \delta = \delta(p, q) > 0. \tag{0.17}$$

Однако при $q \leqslant \max\{p, 2\}$ шары B_p^N являются жёсткими в ℓ_q^N:

$$d_n(B_p^N, \ell_q^N) \geqslant c(\varepsilon) d_0(B_p^N, \ell_q^N), \quad \text{при } n \leqslant N(1 - \varepsilon). \tag{0.18}$$

Это следует из (1.12) при $p \geqslant q$ и из (1.11) при $p < q \leqslant 2$.

В следующей теореме мы даём аналогичный критерий в случае смешанных норм.

Теорема 5.2. *Пусть* $1 \leqslant p_1, p_2, q_1, q_2 \leqslant \infty$; $s, b \in \mathbb{N}$ *и* $N := sb$.

A) Предположим, что набор (p_1, p_2, q_1, q_2) *удовлетворяет трём условиям: (i)* $q_1 \leqslant \max\{p_1, 2\}$; *(ii)* $q_2 \leqslant \max\{p_2, 2\}$; *(iii) он не попадает в "исключительный случай"*

$$q_1 < \min\{p_1, q_2\} \text{ и } p_2 < q_2 \leqslant 2.$$

Тогда для любого $\varepsilon > 0$ *справедливо неравенство*

$$d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \geqslant c(q_2, \varepsilon) d_0(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}), \quad \text{при } n \leqslant N(1-\varepsilon).$$

B) Для любого набора (p_1, p_2, q_1, q_2), *не удовлетворяющего A), существует* $\gamma > 0$ *такое, что для достаточно больших* $t := \min(s, b)$ *имеем*

$$d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \leqslant t^{-\gamma} d_0(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}), \quad \text{если } n \geqslant Nt^{-\gamma}.$$

В **§5.2** рассматривается вопрос о структуре матриц, поэлементно приближающих единичную матрицу. Погрешность приближения единичной матрицы размера $N \times N$ матрицами ранга не выше заданного n равна колмогоровскому поперечнику октаэдра:

$$\min_{\operatorname{rank} A \leqslant n} \max_{1 \leqslant i,j \leqslant N} |A_{i,j} - \delta_{i,j}| = d_n(B_1^N, \ell_\infty^N). \tag{0.19}$$

Этот поперечник изучался в ряде работ: [37, 26, 111] и др. (см. также §1.1). О различных приложениях оценок этой величины в комбинаторике см. [84].

Вероятностный метод ($A_{i,j} = n^{-1}\langle x^i, x^j\rangle$ со случайно выбранными $x^1, \dots, x^N \in \{-1, 1\}^n$, см. работу Кашина [37]) даёт оценку $d_n(B_1^N, \ell_\infty^N) \leqslant 2\sqrt{\ln(N)/n}$, из которой следует, что возможна хорошая аппроксимация единичной матрицы при $n \asymp \log N$. С другой стороны, как известно, для нетривиальной аппроксимации требуется логарифмический ранг. Скажем, если величина (0.19) не превосходит $1/3$, то N столбцов приближающей матрицы лежат в n-мерном подпространстве ℓ_∞^N в шаре радиуса $5/6$; попарные расстояния между ними $\geqslant 1/3$. Подсчёт объёмов даёт $N \leqslant 6^n$.

Кашин задал следующий вопрос: верно ли, что для матриц ранга $n \asymp \log N$, приближающих единичную, положительная доля элементов будет по модулю больше $cn^{-1/2}$?

С матрицей $A \in \mathbb{R}^{N\times N}$ свяжем функцию распределения

$$F_A^*(\gamma) := N^{-2}|\{(i,j)\colon |A_{i,j}| > \gamma\}|.$$

Теорема 5.3. *Пусть A — матрица размера $N \times N$, ранга n такая, что*

$$\max_{1\leqslant i,j\leqslant N} |A_{i,j} - \delta_{i,j}| \leqslant 1/3.$$

Тогда справедливо неравенство

$$F_A^*(\gamma) \geqslant c\frac{\log N}{n\log(2+\frac{n}{\log N})}, \quad \text{где } \gamma := c\max\{n^{-3/2}\log N,\, n^{-1}\}.$$

Следствие 5.3. *В условиях теоремы, если $n \leqslant K\log N$ для некоторого K, то не менее $c(K)N^2$ элементов матрицы A по модулю больше $c(K)n^{-1/2}$.*

Это следствие даёт положительный ответ на вопрос Кашина. Вероятностный метод показывает, что порядок $n^{-1/2}$ нельзя увеличить.

§5.3. Постановка задачи этого параграфа является новой. Пусть K — выпуклое центрально-симметричное тело в $\mathbb{R}^N$. Ясно, что приблизить всё тело K линейным подпространством — это то же самое, что приблизить крайние точки K. Однако множество крайних точек может быть большим и даже бесконечным. Если мы интересуемся порядком поперечников, то можно построить конечное множество точек $W = \{x_1, \dots, x_m\} \subset K$ такое, что

$$d_n(W, X) \geqslant c\, d_n(K, X), \quad \text{при } n \lesssim N. \tag{0.20}$$

Неформально будем называть такое W множеством "свидетелей" (в англоязычной литературе в похожих контекстах употребляется слово "witness"). Другими словами, речь идёт об аппроксимации тела K многогранником $\operatorname{conv} W \subset K$ в смысле близости поперечников (0.20). Ясно, что в качестве свидетелей можно взять какое-либо множество точек $x_1, \dots, x_m \in K$, для которых

$$K \subset C\operatorname{conv}\{\pm x_1, \dots, \pm x_m\}.$$

Однако и в этом случае может потребоваться много точек. Скажем, для евклидова шара $|W| = m \geqslant \exp(cN)$ — это вытекает из следующего результата Глускина об оценке объёма. Пусть $x_1, \dots, x_m \in B_2^N$, тогда ([28, Теорема 2])

$$\left(\frac{\operatorname{Vol}\operatorname{conv}\{\pm x_1, \dots, \pm x_m\}}{\operatorname{Vol} B_2^N}\right)^{1/N} \lesssim \min\{1, N^{-1/2}\log^{1/2}(2+m/N)\}.$$

Оказалось, что во многих случаях свидетелей нужно не очень много. Этот феномен ещё не вполне изучен, и первым шагом является следующая

Теорема 5.4. *При всех* $1 \leqslant p, q \leqslant \infty$, *за исключением случая* $1 < p < 2$, $q = \infty$, *для любого* $N \in \mathbb{N}$ *существует множество* $W \subset B_p^N$ *мощности* $|W| \leqslant CN \log N$ *такое, что*

$$d_n(W, \ell_q^N) \geqslant c(p,q) d_n(B_p^N, \ell_q^N) \quad \textit{для всех } 1 \leqslant n \leqslant cN.$$

(При $1 < p < 2$, $q = \infty$ трудность в том, что порядки поперечников не известны.)

Примечания

Глава 1. Основная работа автора по усреднённым поперечникам — [156]; там содержится ряд результатов первой главы, в том числе важное утверждение 1.7, достаточность p-лакунарности для жёсткости и случай метрики L_0. Теорема 1.1 не публиковалась ранее. Теоремы 1.2 (условия (I), (U)) и 1.3 опубликованы в [157]; доказательство опиралось (при $q = 1$) на утверждение 1.7 и (при $q > 1$) на утверждение 1.6, п.1. К.С. Рютиным в этой работе получены оценки при $q > 1$ для случайных векторов, удовлетворяющих несколько более слабому, чем безусловность, условию. Теорема 1.4 (верхняя оценка) получена раньше в [155]; там же доказана жёсткость системы Радемахера в L_1.

Глава 2. Теорема 2.1 о классе W_1^1 опубликована в [165]. Теорема 2.2 о классах Бесова — в [159] (с константами $c(q)$ вместо $c(q, \theta)$ — но независимость констант от θ при дискретизации требует дополнительных обоснований).

Глава 3. Относительные поперечники изучались в работах [162] и [161]: теорема 3.2 получена в первой работе, а 3.1 — во второй. Теорема 3.3 о чебышёвских сплайнах доказана в [160].

Глава 4. Основные результаты этой главы получены в [155]; утверждение 4.3 доказано в [156].

Глава 5. Теорема 5.1 о произведении октаэдров доказана (для $\varepsilon = 1/2$) в работе [158]; критерий жёсткости — теорема 5.2 — содержится (в более слабом виде) в виде замечания в [157]. Результат о жёсткости матриц — теорема 5.3 — составляет содержание работы [164]. Теорема 5.4 о “свидетелях” не опубликована.

Глава 1

Усреднённый колмогоровский поперечник

1.1 Предварительные сведения

Соглашения о константах. Через $c, c_1, C(q), \dots$ обозначаются некоторые положительные числа — "константы". Во всех формулировках утверждений зависимость от параметров указывается явно, например: $c_q, C(q, \varepsilon)$; таким образом, c, c_1, C и т.п. — абсолютные положительные постоянные.

Если зависимость от параметров ясна из контекста, мы часто пишем $F \lesssim G$: это является синонимом записи $F = O(G)$ и означает, что $F \leqslant CG$ для некоторого C. В такой ситуации параметры не указываются. Отношение $F \asymp G$ означает, что $F \lesssim G$ и $F \gtrsim G$.

Обычно мы пишем $\log x$, не указывая основание логарифма; будем считать, что основание раз и навсегда зафиксировано. Там, где важны численные константы, пишем явно: $\log_2 x$ или $\ln x = \log_e x$.

Функциональный анализ. Мы используем стандартную терминологию и обозначения из функционального анализа: span — линейная оболочка, conv — выпуклая оболочка, extr — множество крайних точек, $\overline{K}$ — замыкание. Vol_m обозначает m-мерный объём.

Пусть $(\mathcal{X}, \mu)$ — пространство с сигма-конечной мерой. Как обычно отождествляем функции (или случайные величины), отличающиеся на множестве нулевой меры.

Основными для нас будут классические L_p-пространства:

$$\|f\|_p := \|f\|_{L_p} := \left(\int_{\mathcal{X}} |f|^p \, d\mu\right)^{1/p}, \quad p < \infty,$$

и $\|f\|_\infty := \operatorname{ess\,sup} |f(x)|$. В диссертации рассматривается нормированный случай $1 \leqslant p \leqslant \infty$, за исключением §1.4, где возникает также $p \in (0, 1)$.

Пространство L_p допускает нижнюю p-оценку:

$$\left\| \sum_{k=1}^m |f_k| \right\|_p^p \geqslant \sum_{k=1}^m \|f_k\|_p^p, \quad 1 \leqslant p < \infty, \tag{1.1}$$

для любого набора функций $f_k \in L_p$. (Это легко увидеть: используем числовое неравенство $(a+b)^p \geqslant a^p + b^p$ для $a, b \geqslant 0$.)

При $p \in [2, \infty)$ пространство L_p имеет тип 2 (см., например, [116, Vol.II, §1e]), т.е.

$$\mathsf{E} \| \sum_{k=1}^m \varepsilon_k f_k \|_p \leqslant C_p (\sum_{k=1}^m \|f_k\|_p^2)^{1/2}, \quad 2 \leqslant p < \infty,$$

где математическое ожидание берётся по случайным (независимым и равномерно распределённым) знакам $\varepsilon_1, \ldots, \varepsilon_m \in \{-1, 1\}$.

Для $1 \leqslant p \leqslant \infty$ через p' обозначаем сопряжённый показатель, $1/p + 1/p' = 1$.

Пусть $N \in \mathbb{N}$. Через ℓ_p^N обозначается пространство $\mathbb{R}^N$ с нормой, соответствующей считающей мере на $\{1, \ldots, N\}$:

$$\|x\|_p := \|x\|_{\ell_p^N} := (\sum_{k=1}^N |x_k|^p)^{1/p}, \quad 1 \leqslant p < \infty,$$

$$\|x\|_\infty := \|x\|_{\ell_\infty^N} := \max_{1 \leqslant k \leqslant N} |x_k|,$$

и B_p^N — единичный шар в этом пространстве. Смысл обозначения $\|\cdot\|_p$ будет ясен из контекста: это либо $\|\cdot\|_{L_p}$-норма, если аргумент — функция, либо $\|\cdot\|_{\ell_p^N}$-норма, если аргумент — конечномерный вектор. В евклидовом случае для нормы применяется обозначение $|\cdot|$. Мы часто используем неравенство

$$\|x\|_p \leqslant N^{(1/p - 1/q)_+} \|x\|_q, \quad x \in \mathbb{R}^N.$$

Иногда будет удобнее работать с пространствами L_p^N с нормой $\|x\|_{L_p^N} := N^{-1/p} \|x\|_{\ell_p^N}$.

Пусть $\mu(X) < \infty$; рассмотрим пространство $L_0(\mathcal{X}, \mu)$, состоящее из всех измеримых функций. Сходимость по мере в L_0 может быть задана различными метриками. Нам будет удобно работать с метрикой Ки Фаня:

$$\|f - g\|_{L_0} := \inf\{\varepsilon \geqslant 0 \colon \mu\{x \colon |f(x) - g(x)| > \varepsilon\} \leqslant \varepsilon\}. \tag{1.2}$$

Минимум в (1.2) всегда достигается; это и другие свойства метрики Ки Фаня см., например, в [97, §9.2].

В §1.4 мы также используем аналог метрики Хэмминга

$$\|f - g\|_{L_\text{Н}} := \mu\{x \colon f(x) \neq g(x)\}.$$

Метрика L_0 является самой слабой из рассматриваемых: $\|f\|_{L_\text{Н}} \geqslant \|f\|_{L_0}$ и

$$\|f\|_p \geqslant \|f\|_{L_0}^{1+1/p}. \tag{1.3}$$

В дискретном случае мы рассматриваем только L_0^N и $L_\text{Н}^N$ (не $\ell_\text{Н}^N$, ℓ_0^N):

$$\|x\|_{L_0^N} = \min\{\varepsilon \geqslant 0 \colon \#\{i \colon |x_i| > \varepsilon\} \leqslant \varepsilon N\}, \tag{1.4}$$

$$\|x\|_{L_\text{Н}^N} = N^{-1}\#\{i \colon x_i \neq 0\}.$$

Пространства гладких функций. Пространство Соболева $\mathcal{W}_p^r[a, b]$ состоит из функций $f \colon [a, b] \to \mathbb{R}$, абсолютно непрерывных вместе с производными $f', \ldots, f^{(r-1)}$, для которых конечна полунорма $|f|_{W_p^r} := \|f^{(r)}\|_p$. Норма в этом пространстве: $\|f\|_p + |f|_{W_p^r}$ (впрочем, нам она не понадобится).

Класс Соболева $W_p^r[a, b]$ состоит из функций пространства $\mathcal{W}_p^r[a, b]$, для которых $\|f^{(r)}\|_p \leqslant 1$. Отметим, что $W_p^r[a, b]$ содержит подпространство многочленов степени меньше r. Технически удобнее определять класс именно через полунорму, а не норму; при таком подходе можно находить точные значения поперечников.

Аналогично определяются пространства 2π-периодических функций $\mathcal{W}_p^r(\mathbb{T})$ и соответствующие классы $W_p^r(\mathbb{T})$.

Пространства Соболева тесно связаны с пространствами Бесова.

Пусть заданы параметры $r > 0$, $p, \theta \in [1, \infty]$ и конечный или бесконечный промежуток $I \subset \mathbb{R}$ (мы будем использовать только случаи прямой $I = \mathbb{R}$ и

отрезка $I = [a, b]$). Пространство Бесова $\mathcal{B}^r_{p,\theta}(I)$ состоит из функций $f \in L_p(I)$ с конечной полунормой (где $s := 1 + \lfloor r \rfloor$)

$$|f|_{B^r_{p,\theta}(I)} := \begin{cases} \left(\int_0^\infty (t^{-r}\omega_s(f,t)_p)^\theta \frac{dt}{t}\right)^{1/\theta}, & 1 \leqslant \theta < \infty, \\ \sup_{t>0} t^{-r}\omega_s(f,t)_p, & \theta = \infty. \end{cases}$$

Здесь $\omega_s(f,t)_p := \sup_{0<h\leqslant t} \|\Delta^s_h f\|_{L_p(I\cap(I-sh))}$ — стандартный модуль гладкости. Норма в пространстве Бесова: $\|f\|_{L_p(I)} + |f|_{B^r_{p,\theta}(I)}$.

Класс Бесова $B^r_{p,\theta}(I)$ — это единичный шар соответствующего пространства.

Гармонический анализ. Мы будем рассматривать 2π-периодические функции на прямой, или, эквивалентно, функции на окружности $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$. Норма в $L_p(\mathbb{T})$ берётся без веса, то есть как интеграл по периоду:

$$\|f\|^p_{L_p(\mathbb{T})} = \int_{\mathbb{T}} |f(x)|^p\,dx = \int_0^{2\pi} |f(x)|^p\,dx = \int_{-\pi}^{\pi} |f(x)|^p\,dx.$$

Мы используем тригонометрическую систему $\{e^{ikx}\}_{k\in\mathbb{Z}}$ и коэффициенты Фурье

$$\hat{f}(k) := \frac{1}{2\pi}\int_{\mathbb{T}} f(x)e^{-ikx}\,dx.$$

Заметим однако, что в большинстве случаев в диссертации рассматриваются вещественные функции и вещественные пространства.

Через $\mathcal{T}_m$ обозначается пространство вещественных тригонометрических полиномов порядка не выше m:

$$\mathcal{T}_m := \{T(x) = \sum_{k=-m}^{m} c_k e^{ikx},\ c_k \in \mathbb{C},\ c_{-k} = \overline{c_k},\ k = -m, \dots, m\}.$$

Используется неравенство Бернштейна:

$$\|t'\|_p \leqslant m\|t\|_p, \quad t \in \mathcal{T}_m, \quad 1 \leqslant p \leqslant \infty. \tag{1.5}$$

и неравенство Фавара–Ахиезера–Крейна:

$$\rho(f, \mathcal{T}_{m-1})_p \leqslant \mathcal{K}_r m^{-r}\|f^{(r)}\|_p, \quad f \in \mathcal{W}^r_p(\mathbb{T}), \quad 1 \leqslant p \leqslant \infty, \tag{1.6}$$

где $\mathcal{K}_r$ — константы Фавара:

$$\mathcal{K}_r := \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{(-1)^{k(r+1)}}{(2k+1)^{r+1}}, \quad \frac{\pi}{8} \leqslant \mathcal{K}_r \leqslant \frac{\pi}{2}.$$

Нам будут полезны классические ядра Дирихле, Фейера и Валле Пуссена. Напомним их определения и свойства. Ядро Дирихле

$$D_m(x) := \sum_{|k|\leqslant m} e^{ikx} = \frac{\sin(m+1/2)x}{\sin(x/2)}. \tag{1.7}$$

Ядро Фейера $K_m := (D_0 + D_1 + \ldots + D_{m-1})/m$. Хорошо известна оценка

$$|K_m(x)| \leqslant \min\{m, \frac{\pi^2}{mx^2}\}, \quad |x| \leqslant \pi. \tag{1.8}$$

Функция K_m неотрицательна и $\frac{1}{2\pi}\|K_m\|_1 = 1$. Ядро Валле Пуссена

$$V_{m,l} := \frac{1}{l-m}\sum_{k=m}^{l-1} D_k = \frac{1}{l-m}(lK_l - mK_m).$$

Ясно, что $V_{m,l} \in \mathcal{T}_{l-1}$ и $\widehat{V}_{m,l}(k) = 1$ при $|k| \leqslant m$. При всех $p \geqslant 1$ справедлива оценка

$$\|V_{m,l} * f\|_p \leqslant v_{m,l}\|f\|_p, \quad \text{где } v_{m,l} := \frac{1}{2\pi}\int_{\mathbb{T}} |V_{m,l}(x)|\,dx.$$

Из выражения ядра Валле Пуссена через ядро Фейера и из результата Никольского [59] следует известная оценка норм:

$$v_{m,l} \leqslant \min\left\{\frac{l+m}{l-m}, \frac{4}{\pi^2}\ln\frac{l+m}{l-m} + C\right\}. \tag{1.9}$$

В частности, $v_{m,2m} \leqslant 3$.

Поперечник по Колмогорову. Пусть X — линейное метрическое пространство с метрикой ρ_X. Расстояние от точки $x \in X$ до множества $M \subset X$ задаётся естественным образом:

$$\rho(x, M)_X := \inf_{y\in M} \rho_X(x, y).$$

Определим также уклонение от множества $K \subset X$ до множества M:

$$\rho(K, M)_X := \sup_{x\in K} \rho(x, M)_X.$$

Колмогоровский n-поперечник ($n \in \mathbb{Z}_+$) множества $K \subset X$ это наименьшее уклонение от K до всевозможных линейных подпространств в X размерности не выше n:

$$d_n(K, X) := \inf_{\substack{Q_n\subset X \\ \dim Q_n\leqslant n}} \rho(K, Q_n)_X.$$

Обычно X является нормированным пространством с расстоянием $\rho_X(x,y) = \|x-y\|_X$; однако в оригинальной работе Колмогорова этого не требуется; мы так же будем рассматривать поперечники и в более общей ситуации.

Множество K всегда можно считать выпуклым и симметричным относительно нуля; для любого $K \subset X$ имеем

$$d_n(K,X) = d_n(\overline{\operatorname{conv}}(K \cup (-K)), X). \tag{1.10}$$

Для компактных множеств в нормированных пространствах, в силу теоремы Крейна–Мильмана, $d_n(K,X) = d_n(\operatorname{extr} K, X)$.

Нередко будет возникать 0-поперечник. В нормированном случае он равен просто $d_0(K,X) = \sup_{x\in K}\|x\|$; для центрально-симметричных множеств эта величина равна половине диаметра.

Приведём основные оценки поперечников $d_n(B_p^N, \ell_q^N)$; они будут использоваться в нескольких главах. Если область изменения n не указана, считаем, что $0 \leqslant n < N$.

1. Октаэдр в евклидовой норме (см. (0.6)):

$$d_n(B_1^N, \ell_2^N) = (1-n/N)^{1/2}. \tag{1.11}$$

Отсюда сразу следует, что при $p < q \leqslant 2$ хорошей аппроксимации также нет:

$$1 \geqslant d_n(B_p^N, \ell_q^N) \geqslant d_n(B_1^N, \ell_2^N) \geqslant \varepsilon^{1/2}, \quad \text{если } n \leqslant N(1-\varepsilon).$$

2. "Верхний треугольник": $p \geqslant q$. Пич [125], Стесин [66]:

$$d_n(B_p^N, \ell_q^N) = (N-n)^{1/q-1/p}. \tag{1.12}$$

Здесь тоже нет хорошей аппроксимации, при $n \leqslant N(1-\varepsilon)$ поперечник имеет тот же порядок, что и $d_0(B_p^N, \ell_q^N) = \sup_{x\in B_p^N}\|x\|_q = N^{1/q-1/p}$.

3. Октаэдр в ℓ_∞^N. Из (1.11) сразу следует оценка при $n \leqslant N/2$:

$$d_n(B_1^N, \ell_\infty^N) \geqslant d_n(B_1^{2n}, \ell_\infty^{2n}) \geqslant (2n)^{-1/2} d_n(B_1^{2n}, \ell_2^{2n}) = (4n)^{-1/2}.$$

Эту оценку снизу уточнил Глускин [26]:

$$d_n(B_1^N, \ell_\infty^N) \geqslant c\sqrt{\frac{\log N}{n\log(2+n/\log N)}}, \quad \log N \leqslant n \leqslant N/2. \tag{1.13}$$

Кашин [37] получил оценку сверху

$$d_n(B_1^N, \ell_\infty^N) \leqslant 2\sqrt{\frac{\ln N}{n}}.$$

При больших n работает неравенство Хеллига [111]:

$$d_n(B_1^N, \ell_\infty^N) \leqslant Cn^{-1/2}\frac{\log N}{\log n}. \tag{1.14}$$

Из него следует, что в зоне $N^\gamma \leqslant n \leqslant N/2$ поперечник имеет порядок $\asymp n^{-1/2}$ (этот факт был установлен ещё в [38]).

4. Октаэдр в ℓ_q^N, $2 < q < \infty$. Точные по порядку оценки были получены Кашиным [40]:

$$\frac{1}{4}\min\{1, N^{1/q}n^{-1/2}\} \leqslant d_n(B_1^N, \ell_q^N) \leqslant c_q \min\{1, N^{1/q}n^{-1/2}\}, \quad n \leqslant N/2. \tag{1.15}$$

5. Случай $\max\{2, p\} < q < \infty$. Порядки нашёл Глускин [24, 25]. Выпишем их при $n \leqslant N/2$ для упрощения формулы:

$$\Phi(N, n, p, q) := \begin{cases} \min\{1, N^{1/q}n^{-1/2}\}^\omega, \ \omega := \frac{1/p-1/q}{1/2-1/q}, & 2 \leqslant p < q < \infty, \\ \max\{N^{1/q-1/p}, \min\{1, N^{1/q}n^{-1/2}\}\}, & 1 \leqslant p < 2 < q < \infty, \end{cases}$$

$$c(p,q)\Phi(N, n, p, q) \leqslant d_n(B_p^N, \ell_q^N) \leqslant C(p,q)\Phi(N, n, p, q), \quad n \leqslant N/2. \tag{1.16}$$

6. Случай $2 \leqslant p < \infty$, $q = \infty$. Кашин [39], Гарнаев–Глускин [22]:

$$\frac{1}{4}\min\left\{1, \left(c_1\frac{\log(2N/n)}{n}\right)^{1/p}\right\} \leqslant d_n(B_p^N, \ell_\infty^N) \leqslant \min\left\{1, \left(c_2\frac{\log(2N/n)}{n}\right)^{1/p}\right\}. \tag{1.17}$$

7. При $1 \leqslant p < 2$, $q = \infty$ порядки неизвестны.

Двойственность в нормированных пространствах даёт важные инструменты для получения оценок снизу на расстояние и поперечники. Представленные ниже факты содержатся, например, в книге Корнейчука [49, Глава 2].

Если $M \subset X$ и $x \in X$, то

$$\rho(x, M)_X \geqslant \sup_{\xi \in X^*, \|\xi\| \leqslant 1} \left(\langle \xi, x \rangle - \sup_{y \in M} \langle \xi, y \rangle \right). \tag{1.18}$$

Действительно, для любого $\xi \in X^*$, $\|\xi\| \leqslant 1$ и любого $y \in M$ имеем

$$\|x - y\| \geqslant \langle \xi, x - y\rangle = \langle \xi, x\rangle - \langle \xi, y\rangle \geqslant \langle \xi, x\rangle - \sup_{u\in M}\langle \xi, u\rangle.$$

Для выпуклых множеств этот метод точен, достигается равенство:

$$\rho(x, M)_X = \sup_{\xi\in X^*, \|\xi\|\leqslant 1} \left(\langle \xi, x\rangle - \sup_{y\in M}\langle \xi, y\rangle\right). \tag{1.19}$$

В частности, для подпространств $Q \subset X$:

$$\rho(x, Q)_X = \sup\{\langle \xi, x\rangle \colon \xi \in X^*, \|\xi\| \leqslant 1,\ \xi \perp Q\}, \tag{1.20}$$

где запись $\xi \perp Q$ означает, что $\langle \xi, y\rangle = 0$ для всех $y \in Q$.

В случае пространства L_∞ удобны неравенства для расстояния в сопряжённом пространстве. Как установил Никольский, для любого конечномерного подпространства $V \subset X^*$ и $\xi \in X^*$ справедлива формула:

$$\rho(\xi, V)_{X^*} = \sup\{\langle \xi, x\rangle \colon x \in X,\ \|x\| \leqslant 1,\ x \perp V\}, \quad \dim V < \infty. \tag{1.21}$$

Поперечник по Колмогорову выражается через двойственный поперечник по Гельфанду. Последний определяется так:

$$d^n(K, X) := \inf_{\substack{Q^n\subset X\\ \operatorname{codim} Q^n\leqslant n}} \sup_{x\in K\cap Q^n} \|x\|_X.$$

Если $X = (\mathbb{R}^N, \|\cdot\|)$ и $Y = (\mathbb{R}^N, \|\cdot\|)$ — нормированные пространства с единичными шарами B_X и B_Y, соответственно, то из (1.20) следует равенство

$$d_n(B_X, Y) = d^n(B_{Y^*}, X^*). \tag{1.22}$$

Отметим, что шаром X^* является поляра B_X.

Вероятность. Мы используем стандартные обозначения из теории вероятностей: P обозначет вероятность; E — математическое ожидание; Law — распределение; $\mathbf{1}_A$ или $\mathbf{1}\{\ldots\}$ — индикатор события. Используется также условное математическое ожидание $\mathsf{E}(\xi|A) = \mathsf{E}(\xi\mathbf{1}_A)/\mathsf{P}(A)$ и условное распределение $\operatorname{Law}(\xi|A)$ (соответствующее ограничению ξ на A и вероятностному пространству (Ω', P'), $\Omega' := A$, $\mathsf{P}'(B) := \mathsf{P}(B)/\mathsf{P}(A)$). Для L_p-нормы мы используем оба обозначения: $\|\zeta\|_{L_p} = (\mathsf{E}|\zeta|^p)^{1/p}$.

Через $\operatorname{Unif} K$ обозначается равномерное распределение на K; соответственно, $\xi \sim \operatorname{Unif} K$ — это случайный элемент, равномерно распределённый на K.

Нам потребуются стандартные оценки вероятностей больших уклонений (см. [91, Ch.2] или [151, Ch.2]).

Пусть случайные величины $X_1, \dots, X_k$ независимы и $\mathsf{E}X_i = 0$ при всех i.

Неравенство Хёффдинга. Если $a_i \leqslant X_i \leqslant b_i$ п.н., то

$$\mathsf{P}(|\sum_{i=1}^{k} X_i| \geqslant t) \leqslant 2\exp(-2t^2/\sigma^2), \quad \text{где } \sigma^2 := \textstyle\sum_{i=1}^{k}(b_i - a_i)^2. \tag{1.23}$$

Иногда полезно одностороннее неравенство:

$$\mathsf{P}(\sum_{i=1}^{k} X_i \geqslant t) \leqslant \exp(-2t^2/\sigma^2). \tag{1.24}$$

Неравенство Бернштейна. Если $|X_i| \leqslant M$ п.н., то

$$\mathsf{P}(|\sum_{i=1}^{k} X_i| \geqslant t) \leqslant 2\exp\left(-\frac{t^2/2}{\sigma^2 + Mt/3}\right), \quad \sigma^2 := \sum_{i=1}^{k} \mathsf{E}X_i^2. \tag{1.25}$$

Это неравенство хорошо работает в следующей ситуации. Пусть Λ — случайное подмножество в $[k] := \{1, \dots, k\}$, и каждая точка $i \in [k]$ попадает в Λ независимо от других. Мы можем применить (1.25) к случайным величинам $X_i := \mathbf{1}\{i \in \Lambda\} - \tau_i$, где $\tau_i := \mathsf{P}(i \in \Lambda)$. Пусть $\mathsf{E}|\Lambda| \leqslant t$. Пользуясь неравенством (1.25) и тем, что $\sigma^2 = \sum_{i=1}^{k} \tau_i(1 - \tau_i) \leqslant t$, мы получаем оценку

$$\mathsf{P}(|\#\Lambda - \mathsf{E}(\#\Lambda)| \geqslant \varepsilon t) \leqslant 2\exp(-c(\varepsilon)t), \quad t \geqslant \mathsf{E}|\Lambda|,\ \varepsilon > 0. \tag{1.26}$$

Нам понадобится известная оценка хвоста гауссовского распределения:

$$\frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2}\, dx \leqslant \frac{1}{2} e^{-t^2/2}, \quad t \geqslant 0. \tag{1.27}$$

Пусть $h(t)$ это разность левой и правой частей неравенства (1.27); проверим, что $h(t) \leqslant 0$. Имеем $h(0) = 0$, $h'(t) = (t/2 - 1/\sqrt{2\pi})\exp(-t^2/2)$, значит, $h(t) < 0$ на $(0, \sqrt{2/\pi}]$. Далее, $h(+\infty) = 0$ и $h(t)$ возрастает при $t > \sqrt{2/\pi}$; следовательно, $h(t) < 0$ при $t > 0$.

Метрическая энтропия. Пусть M — множество в метрическом пространстве (X, ρ), $\varepsilon > 0$. Через $N_\varepsilon(M, X)$ обозначается минимальный размер ε-сети

для M в X, т.е. минимальное количество точек $x_1, \dots, x_m \in X$, таких что $\rho(M, \{x_1, \dots, x_m\})_X \leqslant \varepsilon$. Иногда мы пишем $N_\varepsilon(M, \rho)$. Через $P_\varepsilon(M, X)$ обозначается максимальное количество точек $y_1, \dots, y_m \in M$ с попарными расстояниями больше ε.

Пусть X — N-мерное нормированное пространство, B_X — его единичный шар, $K \subset X$. Мы используем хорошо известные неравенства (см., например, [151, Ch. 4]:

$$\frac{\operatorname{Vol}(K)}{\operatorname{Vol}(\varepsilon B_X)} \leqslant N_\varepsilon(K, X) \leqslant P_\varepsilon(B_X, X) \leqslant \frac{\operatorname{Vol}(K + (\varepsilon/2)B_X)}{\operatorname{Vol}((\varepsilon/2)B_X)}, \quad \varepsilon > 0. \tag{1.28}$$

В случае единичного шара, $K = B_X$, неравенства выглядят так:

$$(1/\varepsilon)^N \leqslant N_\varepsilon(B_X, X) \leqslant P_\varepsilon(B_X, X) \leqslant (1 + 2/\varepsilon)^N, \quad \varepsilon > 0. \tag{1.29}$$

Нам потребуется оценка объёма евклидова шара:

$$\operatorname{Vol}(B_2^N) = \frac{\pi^{N/2}}{\Gamma(N/2 + 1)} = \left(\frac{2\pi e + o(1)}{N}\right)^{N/2}. \tag{1.30}$$

Комбинаторная размерность. Напомним определение *комбинаторной размерности* класса $\mathcal{F}$ функций $f\colon I \to \mathbb{R}$. Скажем, что конечное множество $J \subset I$ является t-раздробленным ($t > 0$) классом $\mathcal{F}$, если существует функция $h\colon J \to \mathbb{R}$ такая, что для любого выбора знаков $\sigma\colon J \to \{-1, 1\}$ найдётся $f_\sigma \in \mathcal{F}$ с условием

$$\min_{x \in J} \sigma(x)(f_\sigma(x) - h(x)) \geqslant t. \tag{1.31}$$

Комбинаторная размерность $\operatorname{vc}(\mathcal{F}, t)$ — это максимальная мощность множества, t-раздробленного классом $\mathcal{F}$. Данное понятие обобщает широко известную размерность Вапника–Червоненкиса (VC-размерность); см. [151, §8.3].

Если класс $\mathcal{F}$ выпуклый и центрально–симметричный, то можно считать $h(x) \equiv 0$:

$$\min_{x \in J} \sigma(x) f_\sigma(x) \geqslant t. \tag{1.32}$$

Действительно, в этом случае мы можем заменить $\{f_\sigma\}$ на $\{\frac{1}{2}(f_\sigma - f_{-\sigma})\}$.

Мы будем использовать следующий простой (и известный) факт:

$$\text{если} \ \ \operatorname{vc}(\mathcal{F}, t) > n, \quad \text{то} \ \ d_n(\mathcal{F}, \ell_\infty(I)) \geqslant t. \tag{1.33}$$

Действительно, пусть t-раздроблено множество J из $n+1$ точки, но есть аппроксимация функциями из n-мерного подпространства: $\mathcal{F} \ni f \approx g$, $g \in Q_n$, $\|f-g\|_\infty < t$. Существует нетривиальный линейный функционал $\Lambda(f) := \sum_{x\in J} \lambda(x) f(x)$, такой что $\Lambda(g) \equiv 0$ для $g \in Q_n$. Предположим, $\Lambda(h) \geqslant 0$, где h из (1.31). Выберем функцию f, для которой $\operatorname{sign}(\lambda(x))(f(x)-h(x)) \geqslant t$ для всех $x \in J$: $\lambda(x) \neq 0$. Но тогда для приближающей функции g имеем $\operatorname{sign}(\lambda(x))(g(x)-h(x)) > 0$, и, значит, $\Lambda(g) = \Lambda(g-h) + \Lambda(h) > 0$. Противоречие. Случай $\Lambda(h) \leqslant 0$ аналогичен.

Также мы используем классическую VC оценку для классов $\mathcal{F}\colon I \to \{-1,1\}^k$ (т.е., по существу, семейств подмножеств):

$$|\mathcal{F}| > \binom{k}{0} + \binom{k}{1} + \ldots + \binom{k}{d} \quad \Longrightarrow \quad \operatorname{vc}(\mathcal{F},1) > d. \tag{1.34}$$

Другими словами, при выполнении левого неравенства в (1.34) проекция класса на некоторое множество J, $|J| > d$, даст весь куб $\{-1,1\}^J$.

Для оценки биномиальных коэффициентов нам понадобятся следующие неравенства:

$$\binom{k}{0} + \ldots + \binom{k}{m} \leqslant (ek/m)^m, \quad 0 \leqslant m \leqslant k. \tag{1.35}$$

(отметим, что функция $(ek/x)^x$ возрастает при $x \in (0,k]$) и

$$\binom{k}{0} + \ldots + \binom{k}{m} \leqslant 2^{kh_2(m/k)}, \quad 0 \leqslant m \leqslant k/2, \tag{1.36}$$

где $h_2(u) := -u\log_2 u - (1-u)\log_2(1-u)$. Их можно вывести из неравенства Чернова–Хёфдинга (см. [91, Ch.2]) или доказать непосредственно. Пусть $t := m/k$, тогда (1.35) следует из неравенства

$$t^m \sum_{j=0}^{m} \binom{k}{j} \leqslant \sum_{j=0}^{m} \binom{k}{j} t^j = (1+t)^k = (1+m/k)^k \leqslant e^m,$$

а неравенство (1.36) следует из того, что при $t \leqslant 1/2$:

$$1 = (1-t+t)^k \geqslant \sum_{j=0}^{m} \binom{k}{j} (1-t)^{k-j} t^j \geqslant \sum_{j=0}^{m} \binom{k}{j} \cdot (1-t)^{k-m} t^m.$$

Мы используем также, что $h_2(1/2-\varepsilon) \leqslant 1 - c\varepsilon^2$, $\varepsilon \in (0,1/2)$.

Разные обозначения.

- $\mathbb{N} := \{1, 2, 3, \ldots\}$, $\mathbb{Z}_+ := \{0, 1, 2, \ldots\}$;
- $|M|$ или $\#M$ — количество элементов множества;
- $\lfloor x \rfloor$ — целая часть числа (снизу);
- $\lceil x \rceil$ — целая часть числа сверху (минимальное $k \in \mathbb{Z}$ такое, что $k \geqslant x$);
- $x_+ := \max\{x, 0\}$;
- $[m] := \{1, 2, \ldots, m\}$;
- $\operatorname{sign} t := \begin{cases} 1, & t > 0, \\ -1, & t < 0, \\ 0, & t = 0; \end{cases}$
- символ Кронекера: $\delta_{i,j} = 1$ при $i = j$, и $\delta_{i,j} = 0$ иначе;
- $\mathbb{Z}_N$ — циклическая группа $\mathbb{Z}/N\mathbb{Z}$;
- S_N — группа перестановок множества $[N]$;
- $\mathcal{P}_d$ — пространство вещественных многочленов одной переменной степени не выше d;
- чебышёвская (равномерная) норма матрицы $A \in \mathbb{R}^{M \times N}$:

$$\|A\|_C := \max_{\substack{1 \leqslant i \leqslant M \\ 1 \leqslant j \leqslant N}} |A_{i,j}|,$$

- конечная разность $\Delta_h f(x) := f(x + h) - f(x)$ и её степени

$$\Delta_h^l f(x) = \sum_{k=0}^{l} (-1)^k \binom{l}{k} f(x + (l - k)h);$$

- proj_Q — ортопроектор на подпространство $Q \subset E$ в евклидовом пространстве;
- $\mathbf{1}_A$ — характеристическая функция (индикатор) множества;

- $V_k^N := B_\infty^N \cap kB_1^N$ (для $k \in \mathbb{N}$, $k \leqslant N$); легко видеть, что

$$\operatorname{extr} V_k^N = \{\pm e_{i_1} \pm \ldots \pm e_{i_k} \colon 1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant N\}.$$

Обычно N задаёт размерность приближаемого объекта, n — размерность приближающего подпространства, q — метрику, в которой измеряется погрешность.

1.2 Определение и свойства d_n^{avg}

Определение. Пусть X — линейное пространство с метрикой. Напомним, что борелевской сигма-алгеброй называется минимальная сигма-алгебра, содержащая все открытые подмножества X. Далее мы рассматриваем борелевские меры (т.е. меры, заданные на борелевских множествах) и борелевские случайные векторы (случайные элементы, измеримые относительно борелевской сигма-алгебры).

Пусть μ — борелевская мера на X. Для $p \in [1, \infty)$ определим p-усреднённый поперечник по Колмогорову меры μ в X следующим образом:

$$d_n^{\mathrm{avg}}(\mu, X)_p := \inf_{\substack{Q_n \subset X \\ \dim Q_n \leqslant n}} \Big(\int_X \rho(x, Q_n)_X^p \, d\mu(x)\Big)^{1/p}.$$

При $p = 1$ пишем просто $d_n^{\mathrm{avg}}(\mu, X)$. Случаи $p < 1$ и $p = \infty$ мы не рассматриваем.

Всюду далее мы будем предполагать, что усреднённый поперечник рассматривается для вероятностных мер. Эквивалентным образом мы можем определить поперечники для борелевского случайного вектора $\xi \colon \Omega \to X$, заданного на некотором вероятностном пространстве Ω:

$$d_n^{\mathrm{avg}}(\xi, X)_p := \inf_{\substack{Q_n \subset X \\ \dim Q_n \leqslant n}} (\mathsf{E}\rho(\xi, Q_n)_X^p)^{1/p}.$$

Примем также следующее соглашение: под усреднённым поперечником конечного множества точек будем понимать поперечник случайно и равновероятно выбранной точки этого множества:

$$d_n^{\mathrm{avg}}(\{x_1, \ldots, x_N\}, X)_p := d_n^{\mathrm{avg}}(\xi, X)_p, \quad \text{где } \xi \sim \mathrm{Unif}\{x_1, \ldots, x_N\}.$$

По определению,

$$d_n^{\mathrm{avg}}(\{x_1,\dots,x_N\},X)_p = \inf_{Q_n\subset X}\left(\frac{1}{N}\sum_{k=1}^{N}\rho(x_k,Q_n)_X^p\right)^{1/p}.$$

В литературе для усреднённых поперечников используются различные другие обозначения. Приведём аргументы в пользу новых обозначений $d_n^{\mathrm{avg}}(\mu,X)_p$ и $d_n^{\mathrm{avg}}(\xi,X)_p$. Ряд авторов указывают в обозначении поперечника и меру, и множество, по которым ведётся усреднение — нам это менее удобно; кроме того, вместо указания множества всегда можно ограничить меру на это множество или рассмотреть условный случайный вектор с распределением $\mathrm{Law}(\xi|\xi\in K)$. Указание на усреднение (average) делается обычно с помощью буквы "a": $d_n^{(a)}(...)$; но так может возникнуть путаница с абсолютным поперечником (см. [121]), поэтому был выбран вариант avg. Кроме того, avg используется в Information-based complexity [141]. Наконец, порядок аргументов в $d_n^{\mathrm{avg}}(\xi,X)$ выбран так, чтобы соответствовать классическому поперечнику — на первом месте объект, поперечник которого считается.

Базовые свойства. Усреднённый поперечник даёт оценку снизу обычного поперечника:

$$d_n(K,X)\geqslant d_n^{\mathrm{avg}}(\xi,X)_p,\quad \text{если } \mathsf{P}(\xi\in K)=1.$$

Можно несколько уточнить эту оценку:

$$d_n(K,X)\geqslant \frac{d_n^{\mathrm{avg}}(\xi,X)_p}{(\mathsf{E}\|\xi\|_K^p)^{1/p}},\quad \text{где } \|x\|_K:=\inf\{r>0\colon x/r\in K\}.$$

Следующее неравенство вытекает из выпуклости L_p-нормы:

$$d_{n_1+n_2}^{\mathrm{avg}}(\xi_1+\xi_2,X)_p\leqslant d_{n_1}^{\mathrm{avg}}(\xi_1,X)_p+d_{n_2}^{\mathrm{avg}}(\xi_2,X)_p,\quad 1\leqslant p<\infty.$$

Ясно, что поперечник монотонен по параметру p:

$$d_n^{\mathrm{avg}}(\xi,X)_{p_1}\leqslant d_n^{\mathrm{avg}}(\xi,X)_{p_2},\quad \text{если } p_1<p_2. \tag{1.37}$$

Пусть ξ,η — случайные элементы в X, и распределение $\mu:=\mathrm{Law}\,\xi$ абсолютно непрерывно относительно распределения $\nu:=\mathrm{Law}\,\eta$. Возьмём производную

Радона–Никодима $\varphi := d\mu/d\nu$. По определению, для любого подпространства $Q \subset X$ имеем

$$\int_X \rho^p(x, Q)\, d\mu = \int_X \varphi(x)\rho^p(x, Q)\, d\nu,$$

откуда

$$d_n^{\mathrm{avg}}(\xi, X)_p = d_n^{\mathrm{avg}}(\varphi(\eta)^{1/p}\eta, X)_p.$$

Таким образом мы можем "перегонять" вероятность в скалярный множитель.

Для любого линейного непрерывного оператора $T\colon X \to Y$ и случайного элемента ξ в X имеем простое неравенство:

$$d_n^{\mathrm{avg}}(T\xi, Y)_p \leqslant \|T\colon X \to Y\| \cdot d_n^{\mathrm{avg}}(\xi, X)_p. \tag{1.38}$$

Нам также понадобится

Утверждение 1.1. *Пусть $q \in [1, \infty)$ и $\xi_1, \ldots, \xi_N\colon \Omega \to \mathbb{R}$ случайные величины такие, что $\mathsf{E}|\xi_i|^q < \infty$ при всех i. Для всех $n \in \mathbb{Z}_+$ имеет место равенство*

$$d_n^{\mathrm{avg}}(\{\xi_1, \ldots, \xi_N\}, L_q(\Omega))_q = N^{-1/q} d_n^{\mathrm{avg}}((\xi_1, \ldots, \xi_N), \ell_q^N)_q. \tag{1.39}$$

В левой части равенства усреднение проводится по конечному множеству "точек" $\xi_1, \ldots, \xi_N$ в пространстве $L_q(\Omega)$; в правой части равенства — поперечник случайного вектора $\xi = (\xi_1, \ldots, \xi_N)$ в $\mathbb{R}^N$. В случае конечного множества Ω равенство (1.39) сразу следует из равенства размерностей пространства строк и пространства столбцов матрицы. Для полноты изложения мы приведём доказательство в общем случае.

Доказательство. Рассмотрим погрешность аппроксимации случайных величин $\xi_k \approx \eta_k$:

$$\sum_{k=1}^N \|\xi_k - \eta_k\|_{L_q(\Omega)}^q = \sum_{k=1}^N \mathsf{E}|\xi_k - \eta_k|^q = \mathsf{E}\|\xi - \eta\|_{\ell_q^N}^q. \tag{1.40}$$

Заметим, что случайные величины $\eta_1, \ldots, \eta_N$ лежат в подпространстве $L_q(\Omega)$ размерности не выше n тогда и только тогда, когда найдётся подпространство $Q \subset \mathbb{R}^N$, $\dim Q \leqslant n$, для которого

$$\eta = (\eta_1, \ldots, \eta_N) \in Q \quad \text{почти наверное.}$$

Следовательно,

$$\mathsf{E}\|\xi - \eta\|_{\ell_q^N}^q \geqslant \mathsf{E}\rho(\xi, Q)_{\ell_q^N}^q \geqslant d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q^q.$$

Это обеспечивает "$\geqslant$" в (1.39).

Для доказательства обратного неравенства нужно построить η по (почти) оптимальному подпространству $Q \subset \mathbb{R}^N$. Для любого $\varepsilon > 0$ найдётся борелевское (например, кусочно–постоянное) отображение $\pi\colon \mathbb{R}^N \to Q$ такое, что

$$\|x - \pi(x)\|_{\ell_q^N} \leqslant \rho(x, Q)_{\ell_q^N} + \varepsilon, \quad x \in \mathbb{R}^N.$$

Тогда $\eta := \pi(\xi)$ является (борелевским) случайным вектором. Кроме того, $\eta \in Q$ п.н., поэтому $\eta_1, \ldots, \eta_N$ порождают пространство размерности не выше n в $L_q(\Omega)$. Подставим $x = \xi$ в неравенство выше, возьмём q-среднее и применим равенство (1.40), завершая доказательство. □

Поперечники множества базисных векторов. Как мы уже говорили во введении (см. (0.6)), для стандартного базиса $e_1, \ldots, e_N$ справедливо равенство

$$d_n(\{e_1, \ldots, e_N\}, \ell_2^N) = (1 - n/N)^{1/2}, \tag{1.41}$$

и для любого n-мерного подпространства $Q_n \subset \mathbb{R}^N$ имеем

$$\sum_{k=1}^{N} \rho(e_k, Q_n)_2^2 = N - n.$$

Отсюда

$$d_n^{\mathrm{avg}}(\{e_1, \ldots, e_N\}, \ell_2^N)_2 = (1 - n/N)^{1/2}.$$

Зафиксируем ещё несколько простых наблюдений об усреднённых поперечниках базисных векторов.

Утверждение 1.2. *Пусть $N, n \in \mathbb{N}$, $n < N$.*

1. При $1 \leqslant p < \infty$ справедливо равенство

$$d_n^{\mathrm{avg}}(\{e_1, \ldots, e_N\}, \ell_2^N)_p = (1 - n/N)^{1/p^*}, \quad \text{где } p^* := \min\{p, 2\}. \tag{1.42}$$

2. Для любых $r_1 \geqslant r_2 \geqslant \ldots \geqslant r_N \geqslant 0$ имеем

$$d_n^{\mathrm{avg}}(\{r_1 e_1, \ldots, r_N e_N\}, \ell_2^N)_2 = N^{-1/2} \left(\sum_{k=n+1}^{N} r_k^2 \right)^{1/2}. \tag{1.43}$$

3. Для любых $1 \leqslant q \leqslant \infty$ и $n \leqslant N/2$ имеем

$$d_n^{\mathrm{avg}}(\{e_1, \ldots, e_N\}, \ell_q^N) \geqslant \frac{1}{2} d_n(B_1^{\lceil N/2 \rceil}, \ell_q^{\lceil N/2 \rceil}). \tag{1.44}$$

Доказательство. 1. При $p = 2$ равенство уже установлено. При $p > 2$ в силу монотонности (1.37), имеет место та же оценка снизу. Оценка сверху при $p \geqslant 2$ вытекает из (1.41). Пусть $1 \leqslant p < 2$. В силу неравенств $\rho(e_k, Q_n)_2 \leqslant 1$, получаем

$$\sum_{k=1}^{N} \rho(e_k, Q_n)_2^p \geqslant N - n,$$

откуда $d_n^{\mathrm{avg}}(\dots)_p \geqslant (1-n/N)^{1/p}$. Оценка сверху достигается на подпространстве $Q_n := \operatorname{span}\{e_1, \dots, e_n\}$.

2. Запишем сумму квадратов расстояний:

$$\sum_{k=1}^{N} \rho(r_k e_k, Q_n)^2 = \sum_{k=1}^{N} \theta_k r_k^2, \quad \text{где } \theta_k := \rho(e_k, Q_n)^2.$$

Коэффициенты (θ_k) удовлетворяют соотношениям $0 \leqslant \theta_k \leqslant 1$, $\sum \theta_k = N - n$. Ясно, что сумма минимизируется, когда $\theta_k = 0$, $k < N - n$ и $\theta_k = 1$, $k \geqslant N - n$. Отсюда вытекает оценка снизу. Оценка сверху достигается на $Q_n := \operatorname{span}\{e_1, \dots, e_n\}$.

3. Обозначим $d_n := d_n(B_1^{\lceil N/2 \rceil}, \ell_q^{\lceil N/2 \rceil})$. Возьмём произвольное n-мерное подпространство $Q_n \subset \mathbb{R}^N$. Рассмотрим множество "хороших" координат $\Lambda := \{k\colon \rho(e_k, Q_n)_q < d_n\}$. Если $|\Lambda| > N/2$, то

$$d_n > d_n(\{e_k\colon k \in \Lambda\}, \ell_q^N) \geqslant d_n(\{e_1, \dots, e_{\lceil N/2 \rceil}\}, \ell_q^{\lceil N/2 \rceil}) = d_n,$$

что даёт противоречие. Значит, $|\Lambda| \leqslant N/2$ и есть не менее $N/2$ "плохих" координат, откуда

$$\frac{1}{N} \sum_{k=1}^{N} \rho(e_k, Q_n)_q \geqslant \frac{1}{2} d_n.$$

□

Усредненный поперечник в гильбертовом пространстве. Приведём формулу для 2-усреднённого поперечника случайного вектора в гильбертовом пространстве. Эта формула является непосредственным обобщением теоремы Эккарта–Юнга о приближении матриц в норме Фробениуса и, по существу, переформулировкой теоремы Исмагилова [33]. Однако используемая нами терминология представляется весьма удобной и естественной, и такая формулировка может быть полезна. Доказательство приводится для полноты изложения и использует приведённые в этом параграфе свойства усреднённого поперечника.

Пусть H — сепарабельное гильбертово пространство (возможно, конечномерное), ξ — случайный вектор в H и $\mathsf{E}|\xi|^2 < \infty$. Рассмотрим корреляционный оператор A этого случайного вектора, определяемый равенством $\langle Au, v\rangle = \mathsf{E}\langle \xi, u\rangle\langle \xi, v\rangle$. Известно, что A — самосопряжённый неотрицательный компактный оператор (см., например, [11, Гл.3, §2]). Его собственные значения (λ_n) характеризуются тем, что существует ортонормированный базис $\{\varphi_n\}$ (собственный для оператора A), в котором выполнены равенства

$$\mathsf{E}\langle \xi, \varphi_k\rangle^2 = \lambda_k, \quad \mathsf{E}\langle \xi, \varphi_k\rangle\langle \xi, \varphi_l\rangle = 0 \text{ при } k \neq l. \tag{1.45}$$

Условимся записывать собственные значения в невозрастающем порядке: $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant 0$.

Утверждение 1.3. *Имеет место равенство* $d_n^{\mathrm{avg}}(\xi, H)_2 = (\sum\limits_{k>n} \lambda_k)^{1/2}$.

Доказательство. Зафиксируем базис (1.45); положим $\xi_k := \langle \xi, \varphi_k\rangle$. Оценим поперечник снизу. Зафиксируем $N > n$ и положим $\overline{\xi} := (\xi_1, \ldots, \xi_N)$. Поскольку $|x|^2 = \sum\langle x, \varphi_k\rangle^2$, имеем

$$d_n^{\mathrm{avg}}(\xi, H)_2 \geqslant d_n^{\mathrm{avg}}(\overline{\xi}, \ell_2^N)_2.$$

Воспользуемся утверждением 1.1, тем что ξ_k ортогональны в $L_2(\Omega)$ и $\|\xi_k\|_2 = \lambda_k^{1/2}$, и равенством (1.43):

$$d_n^{\mathrm{avg}}(\overline{\xi}, \ell_2^N)_2 = N^{1/2} d_n^{\mathrm{avg}}(\{\xi_1, \ldots, \xi_N\}, L_2(\Omega))_2 =$$
$$= N^{1/2} d_n^{\mathrm{avg}}(\{\lambda_1^{1/2} e_1, \ldots, \lambda_N^{1/2} e_N\}, \ell_2^N)_2 = (\sum_{k=n+1}^{N} \lambda_k)^{1/2}.$$

Устремим N к бесконечности и получим оценку снизу. Оценка сверху достигается на подпространстве $Q_n := \mathrm{span}\{\varphi_1, \ldots, \varphi_n\}$. □

Чтобы применить утверждение 1.3, достаточно найти какой-либо ортонормированный базис, в котором выполнено соотношение ортогональности

$$\mathsf{E}\langle \xi, \varphi_k\rangle\langle \xi, \varphi_l\rangle = 0, \quad l \neq k,$$

и взять в качестве (λ_k) невозрастающую перестановку последовательности чисел $\{\mathsf{E}\langle \xi, \varphi_k\rangle^2\}$.

Нам пригодится следующий частный случай.

Утверждение 1.4. *Пусть* $\mathbf{f}$ *— случайный элемент* $L_2(\mathbb{T})$ *и* $\mathsf{E}\hat{\mathbf{f}}(k)\overline{\hat{\mathbf{f}}(m)} = 0$ *при* $k \neq m$. *Тогда* $d_n^{\mathrm{avg}}(\mathbf{f}, L_2(\mathbb{T}))_2 = (2\pi \sum_{k>n} \lambda_k)^{1/2}$, *где* $\lambda_1 \geqslant \lambda_2 \geqslant \ldots$ *это невозрастающая перестановка чисел* $\{\mathsf{E}|\hat{\mathbf{f}}(k)|^2\}_{k\in\mathbb{Z}}$.

Доказательство. Дадим чисто вещественное доказательство. Пусть $\mathbf{a}_0 = a_0(\mathbf{f})$, $\mathbf{a}_k = a_k(\mathbf{f})$ и $\mathbf{b}_k = b_k(\mathbf{f})$, $k \in \mathbb{N}$ — это коэффициенты Фурье функции f по тригонометрической системе $\{1, \cos kx, \sin kx \colon k \in \mathbb{N}\}$. Докажем ортогональность этих случайных величин. Ясно, что $\mathbf{a}_0 = \hat{\mathbf{f}}(0)$, $\mathbf{a}_k = \hat{\mathbf{f}}(k) + \hat{\mathbf{f}}(-k)$ и $\mathbf{b}_k = i(\hat{\mathbf{f}}(k) - \hat{\mathbf{f}}(-k))$; обратно, $\hat{\mathbf{f}}(k) = (\mathbf{a}_k - i\mathbf{b}_k)/2$ для $k \in \mathbb{N}$. Кроме того, так как функция вещественна, то $\hat{\mathbf{f}}(-k) = \overline{\hat{\mathbf{f}}(k)}$. Из представления для $\mathbf{a}_k, \mathbf{b}_k$ ясно, что числа $\mathsf{E}\mathbf{a}_k\mathbf{a}_m$, $\mathsf{E}\mathbf{a}_k\mathbf{b}_m$ и $\mathsf{E}\mathbf{b}_k\mathbf{b}_m$ равны нулю при $k \neq m$. Кроме того,

$$0 = \mathsf{E}\hat{\mathbf{f}}(k)\overline{\hat{\mathbf{f}}(-k)} = \mathsf{E}\hat{\mathbf{f}}(k)^2 = \frac{1}{4}(\mathsf{E}\mathbf{a}_k^2 - \mathsf{E}\mathbf{b}_k^2 - 2i\mathsf{E}\mathbf{a}_k\mathbf{b}_k),$$

откуда $\mathsf{E}\mathbf{a}_k\mathbf{b}_k = 0$ и $\mathsf{E}\mathbf{a}_k^2 = \mathsf{E}\mathbf{b}_k^2$. Наконец, $\mathsf{E}|\hat{\mathbf{f}}(k)|^2 = \frac{1}{4}\mathsf{E}(\mathbf{a}_k^2 + \mathbf{b}_k^2)$, и мы находим $\mathsf{E}\mathbf{a}_k^2 = \mathsf{E}\mathbf{b}_k^2 = 2\mathsf{E}|\hat{\mathbf{f}}(k)|^2$. Остаётся нормировать тригонометрическую систему в $L_2[0, 2\pi]$: $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}}\cos kx$, $\frac{1}{\sqrt{\pi}}\sin kx$ и воспользоваться утверждением 1.3. □

1.3 Нижние оценки поперечников в L_q и ℓ_q^N

Напомним, что системы векторов в нормированных пространствах $\{x_i\} \subset X$ и $\{x_j^*\} \subset X^*$ называются сопряжёнными, если $\langle x_i, x_j^* \rangle = \delta_{i,j}$.

Теорема 1.1. *Пусть* $x_1, \ldots, x_N$ *— система элементов нормированного пространства* X, *система* $x_1^*, \ldots, x_N^* \in X^*$ *является сопряжённой к* $(x_k)_1^N$, *и для некоторых* $1 \leqslant s \leqslant 2$ *и* $B > 0$ *выполнено неравенство:*

$$\|\sum_{k=1}^N a_k x_k^*\|_{X^*} \leqslant B(\sum_{k=1}^N |a_k|^s)^{1/s}, \quad \forall a_1, \ldots, a_N \in \mathbb{R}. \tag{1.46}$$

Тогда при $1 \leqslant p \leqslant 2$ *справедливо неравенство:*

$$d_n^{\mathrm{avg}}(\{x_1, \ldots, x_N\}, X)_p \geqslant B^{-1} \cdot \begin{cases} (1 - n/N)^{1/p}, & \text{если } s = 2,\ n < N, \\ c\,\min\{1, n^{-1/2}N^{1/s'}\}, & \text{если } s < 2,\ n \leqslant N/4. \end{cases}$$

Доказательство. Рассмотрим операторы "синтеза" $S\colon \mathbb{R}^N \to X^*$ и "анализа" $A\colon X \to \mathbb{R}^N$:

$$S\colon (a_1, \dots, a_N) \mapsto \sum_{k=1}^{N} a_k x_k^*, \quad A\colon x \mapsto (\langle x, x_1^* \rangle, \dots, \langle x, x_N^* \rangle).$$

Нетрудно проверить, что $A^* = S$. Из неравенства (1.38) следует, что для любого случайного элемента ζ в X:

$$d_n^{\mathrm{avg}}(\zeta, X)_p \geqslant \frac{d_n^{\mathrm{avg}}(A\zeta, \ell_{s'}^N)_p}{\|A\colon X \to \ell_{s'}^N\|}.$$

Заменим норму оператора на норму сопряжённого: $\|S\colon \ell_s^N \to X^*\|$. Условие (1.46) означает, что $\|S\| \leqslant B$. В силу сопряжённости систем (x_k) и (x_k^*), имеем $Ax_k = e_k$. Следовательно, если $\zeta \sim \mathrm{Unif}\{x_1, \dots, x_N\}$, то $A\zeta \sim \mathrm{Unif}\{e_1, \dots, e_N\}$, и мы получаем следующую оценку:

$$d_n^{\mathrm{avg}}(\{x_1, \dots, x_N\}, X)_p \geqslant B^{-1} d_n^{\mathrm{avg}}(\{e_1, \dots, e_N\}, \ell_{s'}^N)_p.$$

В случае $s = 2$ воспользуемся равенством (1.42).

В случае $s < 2$ и $p = 1$ оценим усреднённый поперечник через обычный, используя (1.44). Далее применим известную оценку (1.15). При $p > 1$ та же оценка имеет место, в силу монотонности $d_n^{\mathrm{avg}}(\cdots)_p$ по p. □

Применим общий результат к функциональным пространствам $X = L_q$ (мера предполагается сигма-конечной). Пусть $f_1, \dots, f_N \in L_q$ и $g_1, \dots, g_N \in L_{q'}$ — сопряжённые системы. Тогда условие (1.46) запишется в следующем виде:

$$\|\sum_{k=1}^{N} a_k g_k\|_{L_{q'}} \leqslant B(\sum_{k=1}^{N} |a_k|^s)^{1/s}, \quad \forall a_1, \dots, a_N \in \mathbb{R}. \tag{1.47}$$

Ортонормированные системы в L_q, $1 < q < 2$. Рассмотрим случай ортонормированных систем функций $f_1, \dots, f_N$ в пространстве L_q при $q \in (1, 2)$. Как мы увидим далее, в случае $q < 2$ для жёсткости недостаточно ортонормированности. В силу теоремы 1.1, $s = 2$, достаточно q'-лакунарности или, в других терминах, условия $S_{q'}$:

$$\|\sum_{k=1}^{N} a_k f_k\|_{q'} \leqslant B(\sum_{k=1}^{N} a_k^2)^{1/2}. \tag{1.48}$$

Это условие влечёт жёсткость системы в следующем смысле:

$$d_n(\{f_1,\dots,f_N\},L_q)\gtrsim \max_{1\leqslant k\leqslant N}\|f_k\|_q,\quad \text{при } n\leqslant N(1-\varepsilon). \tag{1.49}$$

Как мы уже говорили во введении, теорема Бургейна позволяет выделять из равномерно ограниченной ортонормированной системы лакунарные подсистемы и получать оценки снизу на n, при которых n-поперечник мал. Более прямой путь к оценкам снизу даёт условие (1.47) при $s<2$.

Утверждение 1.5. *Пусть система функций $f_1,\dots,f_N\in L_2$ ортонормирована и ограничена в L_r, $2<r\leqslant\infty$:*

$$\|f_k\|_r\leqslant M,\quad k=1,\dots,N.$$

Тогда для всех $q\in(r',2)$ и $1\leqslant n\leqslant N/4$ имеем оценку

$$d_n(\{f_1,\dots,f_N\},L_q)\geqslant cM^{1-2/s}\min\{1,n^{-1/2}N^{1/s'}\},\quad s:=\frac{2(1/2-1/r)}{1/q-1/r}.$$

Отметим, что $s\in(1,2)$, поскольку $q\in(r',2)$.

Доказательство. Рассмотрим оператор $S\colon(a_1,\dots,a_N)\mapsto\sum_{k=1}^N a_kf_k$. Мы применим теорему 1.1 при подходящем $s<2$; число B в условии (1.47) равно $\|S\colon\ell_s^N\to L_{q'}\|$ и нужно лишь оценить эту норму.

В силу ортонормированности, $\|S\colon\ell_2^N\to L_2\|=1$ и в силу ограниченности $\|S\colon\ell_1^N\to L_r\|\leqslant M$. Применим интерполяцию между $(2,2)$ и $(1,r)$: пусть $1/q'=\theta/2+(1-\theta)/r$, откуда $\theta=(1/q'-1/r)/(1/2-1/r)$; $1/s=\theta/2+(1-\theta)/1$, откуда находим $s=2/(2-\theta)$ и оцениваем норму: $\|S\colon\ell_s^N\to L_{q'}\|\leqslant M^{1-\theta}=M^{2/s-1}$. □

Мы видим, в частности, что ортонормированная ограниченная в L_r система $f_1,\dots,f_N$ не может быть хорошо приближена с точностью $o(1)$ в L_q, $q\in(r',2)$, подпространством размерности N^θ, где $\theta=2/s'=(1/q'-1/r)/(1/2-1/r)$.

Следствие 1.1. *Пусть система функций $f_1,f_2,\dots$ ортонормирована и равномерно ограничена: $\|f_k\|_\infty\leqslant M$. Тогда при $\alpha>0$ для множества*

$$K:=\{f\in L_1\colon\ \sum_{k=1}^\infty k^\alpha|\langle f,f_k\rangle|\leqslant 1\}$$

при любом $1<q\leqslant 2$ справедлива оценка

$$d_n(K,L_q)\geqslant c(M,q,\alpha)n^{-\beta},\quad \beta:=\min\{\alpha q'/2,\alpha+1/2-1/q'\}.$$

Доказательство. Для любого N имеем $K\supset N^{-\alpha}\{f_1,\dots,f_N\}$. Применим утверждение 1.5 с $r=\infty$ для $N\asymp n$ и $N\asymp n^{q'/2}$. □

Случайные векторы в ℓ_q^N, $1 < q < \infty$. Дадим теперь несколько следствий теоремы 1.1 в терминах случайных векторов. Напомним ряд определений.

Слабым моментом случайного вектора η в $\mathbb{R}^N$ порядка p называется величина

$$\sigma_p(\eta) := \sup_{a\in\mathbb{R}^N,\, |a|=1} (\mathsf{E}|\langle \eta, a\rangle|^p)^{1/p}.$$

Мера μ в $\mathbb{R}^N$ называется лог-вогнутой, если $\mu((1-\lambda)A + \lambda B) \geqslant \mu(A)^{1-\lambda}\mu(B)^{\lambda}$ для любых компактов $A, B \subset \mathbb{R}^N$ и $\lambda \in (0,1)$. Случайный вектор ξ в $\mathbb{R}^N$ называется лог-вогнутым, если его распределение лог-вогнуто.

Случайный вектор $\xi = (\xi_1, \ldots, \xi_N)$ называется изотропным, если $\mathsf{E}\xi_i^2 = 1$, $\mathsf{E}\xi_i\xi_j = 0$ для всех $i \neq j$.

Утверждение 1.6. *Пусть $1 < q < \infty$, $\xi = (\xi_1, \ldots, \xi_N)$ — случайный вектор в $\mathbb{R}^N$ и $n \leqslant N(1-\varepsilon)$, где $\varepsilon \in (0,1)$.*

1. Если случайный вектор η в $\mathbb{R}^N$ таков, что $\mathsf{E}\xi_i\eta_j = \delta_{i,j}$, $1 \leqslant i, j \leqslant N$, то

$$d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant c(q,\varepsilon)\frac{N^{1/q}}{\sigma_{q'}(\eta)}.$$

2. Если ξ изотропен и лог-вогнут, то $d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant c(q,\varepsilon)N^{1/q}$.

3. Для любого выпуклого тела $K \subset \mathbb{R}^N$ такого, что при некотором $\alpha > 0$,

$$\frac{1}{\operatorname{Vol} K}\int_K x_i x_j\, dx = \alpha^2\delta_{i,j}, \quad i,j = 1,\ldots,N,$$

имеем $d_n(K, \ell_q^N) \geqslant c(q,\varepsilon)\alpha N^{1/q}$.

Доказательство. 1. Применим утверждение 1.1 и теорему 1.1 для $s = 2$. Условие (1.47) теоремы выполнено с $B := \sigma_{q'}(\eta)$, откуда получаем (где $q^* := \min\{q, 2\}$):

$$N^{-1/q} d_n^{\mathrm{avg}}(\xi, \ell_q^N) = d_n^{\mathrm{avg}}(\{\xi_1,\ldots,\xi_N\}, L_q(\Omega))_q \geqslant$$
$$\geqslant d_n^{\mathrm{avg}}(\{\xi_1,\ldots,\xi_N\}, L_q(\Omega))_{q^*} \geqslant \sigma_{q'}(\eta)^{-1}(1-n/N)^{1/q^*} \geqslant \sigma_{q'}(\eta)^{-1}\varepsilon^{1/q^*}.$$

2. Известно (см., например, [93, Th.2.4.6]), что для любого лог-вогнутого случайного вектора ξ и любых коэффициентов (a_k):

$$(\mathsf{E}|\sum_{k=1}^N a_k\xi_k|^{p_2})^{1/p_2} \leqslant C\,\frac{p_2}{p_1}\,(\mathsf{E}|\sum_{k=1}^N a_k\xi_k|^{p_1})^{1/p_1}, \quad 1 \leqslant p_1 < p_2 < \infty.$$

Следовательно, слабый момент вектора ξ оценивается: $\sigma_{q'}(\xi) \leqslant Cq'/2$. Остаётся применить п.1 для $\eta := \xi$.

3. Если K выпукло, то случайный равномерно распределённый вектор ξ в K будет лог-вогнутым (см., например, [93, §2.1]). Кроме того, для вектора ξ имеем $\mathsf{E}\xi_i\xi_j = \alpha^2\delta_{i,j}$, поэтому ξ/α изотропен и

$$d_n(K, \ell_q^N) \geqslant d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q = \alpha d_n^{\mathrm{avg}}(\xi/\alpha, \ell_q^N)_q \geqslant c(q,\varepsilon)\alpha N^{1/q}.$$

□

Следствие 1.2. *Для любого ортогонального оператора $U \in \mathrm{O}(N)$ справедлива оценка $d_n(UB_\infty^N, \ell_q^N) \geqslant c(q,\varepsilon)N^{1/q}$, при $1 < q < \infty$, $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$.*

Случай L_1. Метрика L_1 более слабая, и для жёсткости нужны более ограничительные условия. Удаётся доказать что-либо при предположениях, так или иначе близких к независимости.

В этом подпараграфе мы используем функцию $\mathrm{sg}(x) := 1$ для $x \geqslant 0$ и $\mathrm{sg}(x) := -1$ для $x < 0$. (Отличие от стандартной функции sign только в том, что $\mathrm{sign}\, 0 = 0$.)

Утверждение 1.7. *Пусть $\xi = (\xi_1, \ldots, \xi_N)$ — случайный вектор в $\mathbb{R}^N$, $\mathsf{E}|\xi_i| \geqslant 1$ для всех $i = 1, \ldots, N$. Предположим также, что для всех i и для всех знаков $\sigma \in \{-1, 1\}^N$ имеем*

$$\mathsf{E}(\xi_i | \operatorname{sg} \xi_j = \sigma_j \ \forall j \neq i) = 0. \tag{1.50}$$

Тогда для $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$ справедливо неравенство:

$$d_n^{\mathrm{avg}}(\xi, \ell_1^N) \geqslant c(\varepsilon)N. \tag{1.51}$$

(Мы требуем выполнения равенства (1.50) для тех σ, i, для которых соответствующее событие имеет ненулевую вероятность.)

В силу двойственности (1.20), расстояние в ℓ_1^N от вектора $x \in \mathbb{R}^N$ до подпространства $Q_n \subset \mathbb{R}^N$ может быть записано в виде:

$$\rho(x, Q_n)_1 = \sup_{z \in K} \langle x, z \rangle, \quad K := B_\infty^N \cap Q_n^\perp. \tag{1.52}$$

Поэтому центральные сечения куба B_∞^N играют здесь важную роль.

Лемма 1.1. *Пусть K — это центральное m-мерное сечение куба B_∞^N и $m \geqslant \varepsilon N$. Тогда* $\operatorname{vc}(K, c_1(\varepsilon)) \geqslant c_2(\varepsilon)N$.

Доказательство леммы опирается на два важных результата.

Теорема 1.A ([146])**.** *Любое m-мерное центральное сечение куба единичного объёма имеет объём не меньше единицы:* $\operatorname{Vol}_m(\frac{1}{2}B_\infty^N \cap L_m) \geqslant 1$.

Теорема 1.B ([120])**.** *Пусть $\mathcal{F}$ — некоторый класс функций, заданных на множестве I и ограниченных по модулю единицей. Для любой вероятностной меры μ на I,*

$$N_t(\mathcal{F}, L_2(\mu)) \leqslant (2/t)^{C\operatorname{vc}(\mathcal{F}, ct)}, \quad 0 < t < 1, \tag{1.53}$$

где C и c — положительные абсолютные постоянные.

Мы применяем эту теорему в случае $|I| < \infty$ и равномерной меры на I.

Доказательство леммы 1.1. Обозначим $K = L_m \cap B_\infty^N$. Из теоремы 1.A мы получаем оценку $\operatorname{Vol}_m(K) \geqslant 2^m$. Используя стандартную оценки энтропии через объёмы (1.28) и оценку объёма шара (1.30), получим

$$N_r(K, \ell_2^N) = N_r(K, \ell_2^N \cap L_m) \geqslant \frac{\operatorname{Vol}_m(K)}{\operatorname{Vol}_m(rB_2^m)} \geqslant \frac{2^m}{(crm^{-1/2})^m}.$$

Положим $r := c^{-1}m^{1/2}$, тогда $N_r(K, \ell_2^N) \geqslant 2^m$. Осталось применить (1.53) ко множеству K (отметим, что по построению $\|z\|_\infty \leqslant 1$ для $z \in K$) в пространстве L_2^N и $t := N^{-1/2}r \asymp \varepsilon^{1/2}$:

$$\operatorname{vc}(K, ct) \geqslant c_1 \log N_t(K, L_2^N) / \log(2/t) \geqslant c(\varepsilon)N.$$

□

Теперь мы готовы доказать утверждение 1.7.

Доказательство. Возьмём n-мерное подпространство $Q_n \subset \mathbb{R}^N$; в силу формулы (1.52) имеем

$$\mathsf{E}\rho(\xi, Q_n)_1 = \mathsf{E}\sup_{z \in K}\langle \xi, z\rangle, \quad K := B_\infty^N \cap Q_n^\perp.$$

Лемма 1.1 даёт множество координат $\Lambda \subset \{1, \dots, N\}$, $|\Lambda| \geqslant c_2(\varepsilon)N$, и число $v \geqslant c_1(\varepsilon)$ такие, что для любого выбора знаков $\sigma\colon \Lambda \to \{-1, 1\}$ найдётся вектор $z^\sigma \in K$,

$$\min_{i \in \Lambda} \sigma_i z_i^\sigma \geqslant v. \tag{1.54}$$

(Напомним, что в силу выпуклости и центрально-симметричности K, мы можем считать, что векторы $\{z^\sigma\}$ осциллируют вокруг нуля, см. (1.32).)

Из (1.54) мы далее выведем оценку

$$\mathsf{E}\max_\sigma\langle\xi, z^\sigma\rangle \geqslant v\sum_{i\in\Lambda}\mathsf{E}|\xi_i|. \tag{1.55}$$

Зафиксируем выбор знаков $\sigma^*\colon \Lambda\to\{-1,1\}$, пусть E^* обозначает условное математическое ожидание $\mathsf{E}^*(\cdot) = \mathsf{E}(\cdot|\operatorname{sg}\xi_i = \sigma_i^*,\ i\in\Lambda)$. Из условия (1.50) следует, что $\mathsf{E}^*(\xi_i) = 0$ при $i\notin\Lambda$. Далее имеем

$$\mathsf{E}^*\max_\sigma\langle\xi, z^\sigma\rangle \geqslant \mathsf{E}^*\langle\xi, z^{\sigma^*}\rangle = \mathsf{E}^*\left(\sum_{i\in\Lambda}\xi_i z_i^{\sigma^*} + \sum_{i\notin\Lambda}\xi_i z_i^{\sigma^*}\right) \geqslant v\sum_{i\in\Lambda}\mathsf{E}^*|\xi_i|.$$

Усредним по σ^* и получим (1.55). Из этого неравенства вытекает (1.51). Утверждение доказано. □

Рассмотрим три условия на случайный вектор ξ в $\mathbb{R}^N$:

(I) $\xi_1,\dots,\xi_N$ независимы и в среднем равны нулю: $\mathsf{E}\xi_1 = \dots = \mathsf{E}\xi_N = 0$;

(U) распределение вектора ξ безусловно, т.е. для любого выбора знаков

$$\operatorname{Law}(\xi_1,\dots,\xi_N) = \operatorname{Law}(\pm\xi_1,\dots,\pm\xi_N);$$

(D) случайные величины ξ_i дизъюнктны, т.е. $\xi_i\xi_j = 0$ п.н. для любых $1\leqslant i < j\leqslant N$.

Теорема 1.2. *Пусть $1\leqslant q\leqslant 2$, случайные величины $\xi_1,\dots,\xi_N\in L_q(\Omega)$ и $\|\xi_i\|_q\geqslant 1$ при $i = 1,\dots,N$. Предположим, что случайный вектор $\xi := (\xi_1,\dots,\xi_N)$ удовлетворяет хотя бы одному из условий (I), (U), (D). Тогда при $n\leqslant N(1-\varepsilon)$, $\varepsilon\in(0,1)$, справедливо неравенство*

$$d_n^{\mathrm{avg}}(\{\xi_1,\dots,\xi_N\}, L_q(\Omega))_q = N^{-1/q}d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant c(q,\varepsilon).$$

Доказательство. Равенство поперечников следует из утверждения 1.1.

Рассмотрим сначала случаи (I) и (U). При $q = 1$ теорема следует из Утверждения 1.7, поскольку в обоих случаях выполнено условие (1.50).

Пусть $1 < q \leqslant 2$. Используя теорему 1.1, мы докажем, что при при $n < N$ справедливо неравенство:

$$d_n^{\mathrm{avg}}(\{\xi_1, \dots, \xi_N\}, L_q)_q \geqslant c_q(1 - n/N)^{1/q}. \tag{1.56}$$

Без ограничения общности можем считать, что $\|\xi_i\|_{L_q} = 1$ для всех i.

Положим $\phi(t) := |t|^{q-1} \operatorname{sign} t$. Ясно, что $\mathsf{E}\xi_i\phi(\xi_i) = 1$. Рассмотрим случайные величины $\eta_i := \phi(\xi_i) - \mathsf{E}\phi(\xi_i)$. Нетрудно видеть, что системы (ξ_i) и (η_i) сопряжены: например, если распределение ξ безусловно, то $\mathrm{Law}(\xi_i, \xi_j) = \mathrm{Law}(-\xi_i, \xi_j)$ и поэтому

$$\mathsf{E}\xi_i\eta_j = \mathsf{E}\xi_i\phi(\xi_j) = \mathsf{E}(-\xi_i)\phi(\xi_j) = 0.$$

Покажем, что система (η_i) обладает тем же свойством (I) или (U), что и (ξ_i). Действительно, независимость ξ_i влечет независимость η_i, и $\mathsf{E}\eta_i = 0$. Для безусловных ξ в силу нечетности функции ϕ выполнено $\mathsf{E}\phi(\xi_i) = 0$ и распределение вектора $(\phi(\xi_i))_{i=1}^N$ также будет безусловным.

Заметим, что

$$\|\eta_i\|_{L_{q'}} \leqslant \|\phi(\xi_i)\|_{L_{q'}} + \|\phi(\xi_i)\|_{L_1} \leqslant 2\|\phi(\xi_i)\|_{L_{q'}} = 2\|\xi_i\|_{L_q}^{q/q'} = 2.$$

Система (η_i) является безусловной по норме в любом пространстве L_p, $1 \leqslant p \leqslant \infty$, а именно

$$\max_{\sigma\in\{-1,1\}^N} \|\sum_{k=1}^N \sigma_k a_k \eta_k\|_p \leqslant C \min_{\sigma\in\{-1,1\}^N} \|\sum_{k=1}^N \sigma_k a_k \eta_k\|_p$$

для любых коэффициентов (a_k): в случае (U) это так по определению; для (I) безусловность по норме также известна — см., например, [2, §3].

Воспользуемся безусловностью для $p = q'$ и тем, что пространство $L_{q'}$ обладает типом 2 (так как $2 \leqslant q' < \infty$):

$$\|\sum_{k=1}^N a_k\eta_k\|_{L_{q'}} \leqslant C\mathsf{E}\|\sum_{k=1}^N \pm a_k\eta_k\|_{L_{q'}} \leqslant C(q)(\sum_{k=1}^N a_k^2\|\eta_k\|_{L_{q'}}^2)^{1/2} \leqslant C_1(q)(\sum_{k=1}^N a_k^2)^{1/2}.$$

Следовательно, выполнено условие (1.47) для $s = 2$, и мы можем применить теорему 1.1, которая даёт неравенство (1.56).

В случае (D) из теоремы 1.1 при $1 \leqslant q \leqslant 2$ можно вывести оценку

$$d_n^{\mathrm{avg}}(\{\xi_1, \dots, \xi_N\}, L_q)_q \geqslant (1 - n/N)^{1/q}.$$

Действительно, положим $\eta_i := \phi(\xi_i)$; это сопряжённая система и она также обладает свойством (D). Следовательно, при всех $q \in [1,2]$ имеем:

$$\|\sum_{k=1}^N a_k\eta_k\|_{q'} = \left(\sum_{k=1}^N \|a_k\eta_k\|_{q'}^{q'}\right)^{1/q'} \leqslant \left(\sum_{k=1}^N |a_k|^{q'}\right)^{1/q'} \max_k \|\eta_k\|_{q'} \leqslant$$

$$\leqslant \left(\sum_{k=1}^N |a_k|^2\right)^{1/2}$$

и выполнено условие (1.47). □

Приведём некоторые следствия доказанной теоремы.

Пусть H — подгруппа группы S_N перестановок множества $\{1,\dots,N\}$. Тогда H действует на множестве $\{1,\dots,N\}$. Действие называется транзитивным, если для любых $i,j \in \{1,\dots,N\}$ найдётся элемент $h \in H$, такой что $h(i) = j$.

Действие H на $\{1,\dots,N\}$ индуцирует действие этой группы на N-мерные векторы:

$$hx := (x_{h^{-1}(1)}, x_{h^{-1}(2)}, \dots, x_{h^{-1}(N)}), \quad h \in H,\ x \in \mathbb{R}^N.$$

Тем самым, получилось (левое) действие: $(h_1h_2)x = h_1(h_2x)$. Группа $\mathbb{Z}_2^N$ действует на $\mathbb{R}^N$ изменениями знаков:

$$\tau x := ((-1)^{\tau_1}x_1, \dots, (-1)^{\tau_N}x_N), \quad \tau \in \mathbb{Z}_2^N,\ x \in \mathbb{R}^N.$$

(Отметим, что действие $(\tau,h)\colon x \mapsto \tau\, h\, x$ является действием полупрямого произведения $\mathbb{Z}_2^N \rtimes H$; чтобы не входить в алгебраические детали, мы предпочитаем рассматривать действия обеих групп отдельно.)

Пусть $\boldsymbol{\tau}$, $\mathbf{h}$ обозначают независимые и равномерно распределенные элементы $\mathbb{Z}_2^N$ и H, соответственно.

Утверждение 1.8. *Пусть $1 \leqslant q \leqslant 2$, $N \in \mathbb{N}$, и группа H транзитивно действует на множестве $\{1,\dots,N\}$. Для любого $x \in \mathbb{R}^N$ и $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$, выполнено:*

$$d_n^{\mathrm{avg}}(\boldsymbol{\tau}\mathbf{h}x, \ell_q^N)_q \geqslant c(q,\varepsilon)\|x\|_q. \tag{1.57}$$

Доказательство. Рассмотрим случайный вектор $\xi := \boldsymbol{\tau}\mathbf{h}x$. Возьмём любые две координаты i,j. В силу транзитивности действия H, существует $h_0 \in H$, такой что $h_0(j) = i$. Тогда

$$\mathsf{E}|\xi_i|^q = \mathsf{E}|(\mathbf{h}x)_i|^q = \mathsf{E}|((h_0\mathbf{h})x)_i|^q = \mathsf{E}|(h_0(\mathbf{h}x))_i|^q = \mathsf{E}|(\mathbf{h}x)_j|^q = \mathsf{E}|\xi_j|^q.$$

Следовательно, $\mathsf{E}|\xi_i|^q = N^{-1}\mathsf{E}\|\xi\|_q^q = N^{-1}\|x\|_q^q$. Кроме того, ξ является безусловным. Остаётся применить теорему 1.2. $\square$

Теорема 1.3. *Пусть группа H транзитивно действует на множестве $\{1,\dots,N\}$. Пусть K — безусловное множество в $\mathbb{R}^N$, инвариантное относительно действия группы H, то есть, для любого элемента $h \in H$,*

$$(x_1,\dots,x_N) \in K \quad \Rightarrow \quad (\pm x_{h(1)},\dots,\pm x_{h(N)}) \in K.$$

Тогда при $1 \leqslant q \leqslant 2$ для любого $\varepsilon \in (0,1)$ справедливо неравенство

$$d_n(K,\ell_q^N) \geqslant c(q,\varepsilon)\sup_{x\in K}\|x\|_q, \quad \text{при } n \leqslant N(1-\varepsilon).$$

Доказательство. Если K содержит некоторый вектор x, то

$$d_n(K,\ell_q^N) \geqslant d_n^{\mathrm{avg}}(\boldsymbol{\tau}\mathbf{h}x,\ell_q^N)_q \geqslant c(q,\varepsilon)\|x\|_q.$$

Для завершения доказательства возьмём супремум по $x \in K$. $\square$

Следствие 1.3. *Если множество $K \subset \mathbb{R}^N$ инвариантно относительно изменений знаков координат и циклических сдвигов координат, то при $1 \leqslant q \leqslant 2$ и $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$, справедливо неравенство $d_n(K,\ell_q^N) \geqslant c(q,\varepsilon)\sup_{x\in K}\|x\|_q$.*

Случай L_q, $q > 2$. Ортонормированная ограниченная в L_r, $r > 2$, система функций является жёсткой в L_2 и, следовательно, во всех L_q, $2 \leqslant q \leqslant r$ (в смысле (1.49)). Однако при $q > r$ это уже не обязательно так, поскольку L_q-нормы функций могут быть велики по сравнению с L_r-нормами. Мы построим пример случайных величин со свойствами (I) и (U), показывающий, что аналог теоремы 1.2 не имеет места при $q > 2$.

Утверждение 1.9. *Для любого $q \in (2,\infty)$ существует $\delta > 0$ такое, что для достаточно больших N найдётся последовательность $\xi_1,\dots,\xi_N$ независимых одинаково распределённых случайных величин, нормированных условиями $\mathsf{E}\xi_i = 0$, $\mathsf{E}|\xi_i|^q = 1$, для которых*

$$d_{N^{1-\delta}}(\{\xi_1,\dots,\xi_N\},L_q) \leqslant C(q)N^{-\delta}.$$

Доказательство. Положим $n := \lfloor N^{1-\beta} \rfloor$, где β выберём позже, и пусть Q_n^* — экстремальное n-мерное подпространство для октаэдра в ℓ_∞^N; в силу оценки поперечника (1.14), имеем

$$\rho(B_1^N, Q_n^*)_\infty = d_n(B_1^N, \ell_\infty^N) \leqslant C(\beta) n^{-1/2}.$$

Мы приближенно "эмулируем" вершины октаэдра B_1^N случайными векторами. Возьмём малое $\varepsilon > 0$, и пусть ξ_1 это случайная величина со следующим распределеним:

$$\mathsf{P}(\xi_1 = 0) = 1 - \varepsilon,\ \mathsf{P}(\xi_1 = M) = \mathsf{P}(\xi_1 = -M) = \varepsilon/2,$$

где M определено из условия $\mathsf{E}|\xi_1|^q = 1$; т.е., $\varepsilon M^q = 1$. Рассмотрим независимые копии $\xi_2, \ldots, \xi_N$ величины ξ_1 и объединим их в вектор $\xi = (\xi_1, \ldots, \xi_N)$. Мы построим приближение этого вектора случайным вектором η, лежащим в Q_n^* (следовательно, $\eta_1, \ldots, \eta_N$ лежат в n-мерном подпространстве L_q).

Рассмотрим три события: $\mathcal{A}_0 = \{\xi = 0\}$, $\mathcal{A}_1$ — ровно одна координата вектора ξ не равна нулю; и $\mathcal{A}_2$ — по крайней мере две координаты ξ не равны нулю. Положим $\eta(\omega) := 0$ для $\omega \in \mathcal{A}_0 \cup \mathcal{A}_2$ и пусть $\eta(\omega)$ это наилучшее ℓ_∞-приближение из Q_n^* вектора $\xi(\omega)$ в случае $\omega \in \mathcal{A}_1$. Отметим, что в последнем случае $\|\xi - \eta\|_\infty \leqslant M\rho(B_1^N, Q_n^*)_\infty$.

В силу симметрии, достаточно оценить $\mathsf{E}|\xi_1 - \eta_1|^q$:

$$\mathsf{E}|\xi_1 - \eta_1|^q \leqslant \mathsf{P}(\mathcal{A}_1) M^q \rho(B_1^N, Q_n^*)_\infty^q + \mathsf{P}(\mathcal{A}_2) M^q.$$

Первое слагаемое оценивается:

$$N\varepsilon M^q C(\beta)^q n^{-q/2} \lesssim N^{1-(1-\beta)q/2} \lesssim N^{-\delta},$$

если β достаточно мало. Второе слагаемое не превосходит $N^2\varepsilon^2 M^q = N^2\varepsilon$ и его можно сделать сколь угодно малым, устремив ε к нулю. □

Нам будет полезно следующее утверждение, обобщение метода Глускина [24].

Утверждение 1.10. *Пусть $2 < q < \infty$ и ξ — случайный вектор в $\mathbb{R}^N$ такой, что при некотором $\alpha > 0$ вектор $\alpha\xi$ изотропен. Тогда*

$$d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant \frac{1}{2} \frac{\mathsf{E}|\xi|^2}{(\mathsf{E}\|\xi\|_{q'}^{q'})^{1/q'}} \quad \textit{при } n \leqslant \frac{1}{16} N^{2/q} \frac{\mathsf{E}|\xi|^2}{(\mathsf{E}\|\xi\|_q^q)^{2/q}}.$$

Доказательство. В силу однородности можем считать, что $\alpha = 1$, т.е. ξ изотропен. В этом случае $\mathsf{E}|\xi|^2 = N$ и доказываемое неравенство можно записать так:

$$d_n^{\mathrm{avg}}(\xi, \ell_q^N)_q \geqslant \frac{N}{2}(\mathsf{E}\|\xi\|_{q'}^{q'})^{-1/q'} \quad \text{при } n \leqslant \frac{1}{16}N^{1+2/q}(\mathsf{E}\|\xi\|_q^q)^{-2/q}.$$

Пусть вектор ξ приближен случайным вектором η из n-мерного подпространства. Используем двойственность ℓ_q^N и $\ell_{q'}^N$:

$$\mathsf{E}\,\|\xi - \eta\|_q \cdot \|\xi\|_{q'} \geqslant \mathsf{E}\langle \xi - \eta, \xi\rangle = \mathsf{E}|\xi|^2 - \mathsf{E}\langle \xi, \eta\rangle = N - \mathsf{E}\langle \xi, \eta\rangle. \tag{1.58}$$

Оценим левую часть (1.58), пользуясь неравенством Гёльдера:

$$\mathsf{E}\,\|\xi - \eta\|_q \cdot \|\xi\|_{q'} \leqslant (\mathsf{E}\|\xi - \eta\|_q^q)^{1/q}(\mathsf{E}\|\xi\|_{q'}^{q'})^{1/q'}.$$

Следовательно,

$$(\mathsf{E}\|\xi - \eta\|_q^q)^{1/q} \geqslant \frac{N - \mathsf{E}\langle \xi, \eta\rangle}{(\mathsf{E}\|\xi\|_{q'}^{q'})^{1/q'}}. \tag{1.59}$$

Оценим $\mathsf{E}\langle \xi, \eta\rangle$. Если $\mathsf{P}(\eta \in Q_n) = 1$, $\dim Q_n \leqslant n$, то

$$\mathsf{E}\langle \xi, \eta\rangle = \mathsf{E}\langle \mathrm{proj}_{Q_n}\xi, \eta\rangle \leqslant (\mathsf{E}|\,\mathrm{proj}_{Q_n}\xi|^2)^{1/2}(\mathsf{E}|\eta|^2)^{1/2}.$$

В силу изотропности, $\mathsf{E}|\,\mathrm{proj}_{Q_n}\xi|^2 = n$. Оценим второй множитель:

$$(\mathsf{E}|\eta|^2)^{1/2} \leqslant (\mathsf{E}|\eta|^q)^{1/q} \leqslant N^{1/2-1/q}(\mathsf{E}\|\eta\|_q^q)^{1/q}.$$

Ясно, что

$$(\mathsf{E}\|\eta\|_q^q)^{1/q} \leqslant (\mathsf{E}\|\xi\|_q^q)^{1/q} + (\mathsf{E}\|\xi - \eta\|_q^q)^{1/q} \leqslant 2(\mathsf{E}\|\xi\|_q^q)^{1/q}.$$

В итоге, получаем

$$\mathsf{E}\langle \xi, \eta\rangle \leqslant n^{1/2} \cdot N^{1/2-1/q} \cdot 2(\mathsf{E}\|\xi\|_q^q)^{1/q}.$$

При указанных n эта величина не превосходит $N/2$, откуда, в силу (1.59), получаем оценку величины $(\mathsf{E}\|\xi - \eta\|_q^q)^{1/q}$. В силу произвольности η, та же оценка справедлива и для поперечника. □

Замечание. Для любого случайного вектора $N^{2/q}(\mathsf{E}|\xi|^2)(\mathsf{E}\|\xi\|_q^q)^{-2/q} \leqslant N$, поэтому условие на n в утверждении нетривиально.

Следствие 1.4. *Пусть $2 < q < \infty$, $k \leqslant N$. Справедливо неравенство*

$$d_n(V_k^N, \ell_q^N) \geqslant \frac{1}{2} k^{1/q} \quad \text{при } n \leqslant \tfrac{1}{16} N^{2/q} k^{1-2/q}.$$

Доказательство. Возьмём $\xi \sim \operatorname{Unif} \operatorname{extr} V_k^N$, т.е. вектор вида $\xi = \pm e_{i_1} \pm \ldots \pm e_{i_k}$ со случайно выбранными различными i_j. Ясно, что для любого $p \in [1, \infty)$ имеем $\mathsf{E}|\xi_i|^p = \mathsf{E}|\xi_i| = k/N$, $\mathsf{E}\|\xi\|_p^p = k$. Остаётся воспользоваться неравенством $d_n(V_k^N, \ell_q^N) \geqslant d_n^{\text{avg}}(\xi, \ell_q^N)_q$ и применить к ξ утверждение 1.10. □

У Глускина [24] была получена аналогичная оценка поперечника, но константа в условии на n зависела от q. Васильева (препринт `arXiv:2411.14799`) отметила, что зависимость от q можно убрать. Приведённое здесь доказательство, по-видимому, наиболее простое и даёт явные константы.

Случай L_0. В этом подпараграфе мы получим нижние оценки в самой слабой метрике из рассматриваемых: метрике Ки–Фаня (1.2), (1.4).

Утверждение 1.11. *Для любого $\varepsilon \in (0, 1)$ существует $\delta > 0$ такое, что если случайные величины $\xi_1, \ldots, \xi_N$ независимы и*

$$\inf_{c \in \mathbb{R}} \|\xi_i - c\|_{L_0} \geqslant \varepsilon, \quad i = 1, \ldots, N,$$

то для любого подпространства $Q_n \subset \mathbb{R}^N$ размерности $n \leqslant \delta N$ имеем

$$\mathsf{P}(\rho(\xi, Q_n)_{L_0^N} \leqslant \delta) \leqslant 2 \exp(-\delta N), \tag{1.60}$$

где $\xi = (\xi_1, \ldots, \xi_N)$.

Условие $\|\xi_i - c\|_{L_0} \geqslant \varepsilon$ существенно, поскольку оно запрещает аппроксимацию вектора ξ константным вектором. Это условие позволяет в некотором смысле “отделить” значения ξ_i.

Лемма 1.2. *Пусть $\varepsilon, \tau \in (0, 1]$, случайная величина ζ удовлетворяет условию $\inf_{c \in \mathbb{R}} \|\zeta - c\|_{L_0} \geqslant \varepsilon$. Тогда найдутся числа $a < b$ такие, что $\mathsf{P}(\zeta \leqslant a) \geqslant \varepsilon/3$, $\mathsf{P}(\zeta \geqslant b) \geqslant \varepsilon/3$, $\mathsf{P}(a < \zeta < b) \leqslant \tau$, $|b - a| \geqslant \tau\varepsilon/2$.*

Доказательство. Рассмотрим квантили

$$q_- := \inf\{x \colon \mathsf{P}(\zeta \leqslant x) \geqslant \varepsilon/3\}, \quad q_+ := \sup\{x \colon \mathsf{P}(\zeta \geqslant x) \geqslant \varepsilon/3\}.$$

Отметим, что $\mathsf{P}(\zeta \leqslant q_-) \geqslant \varepsilon/3$ и $\mathsf{P}(\zeta \geqslant q_+) \geqslant \varepsilon/3$. Видно, что $|q_+ - q_-| \geqslant \varepsilon$, поскольку в противном случае $\|\zeta - c\|_{L_0} \leqslant \frac{2}{3}\varepsilon$ для $c = \frac{1}{2}(q_- + q_+)$. Положим $k := \lceil 1/\tau \rceil$, разделим $[q_-, q_+]$ на k равных интервалов и выберем в качестве (a, b) интервал с наименьшей вероятностью $\mathsf{P}(a < \zeta < b)$. Тогда эта вероятность не превосходит $1/k \leqslant \tau$. Наконец, $|b - a| \geqslant |q_+ - q_-|/k \geqslant \frac{1}{2}\tau\varepsilon$. □

Доказательство утверждения 1.11. Пусть $\delta > 0$ достаточно мало (уточним выбор позже); оценим $\mathsf{P}(\rho(\xi, Q_n)_{L_0^N} < \delta)$, где $\dim Q_n = n \leqslant \delta N$. Рассмотрим (почти) наилучшую L_0^N аппроксимацию вектора $(\xi_1, \ldots, \xi_N)$ случайным вектором $(\eta_1, \ldots, \eta_N)$ из подпространства Q_n. Обозначим через

$$\Lambda := \{k\colon |\xi_k - \eta_k| > \delta\}$$

множество (случайное) координат с плохой аппроксимацией. Если $\|\xi - \eta\|_{L_0^N} < \delta$, то $|\Lambda| < \delta N$; поэтому нам достаточно оценить вероятность события $\{|\Lambda| \leqslant \delta N\}$.

Применим лемму 1.2 к ε, $\tau := 5\delta/\varepsilon$ и ξ_k, получим интервалы (a_k, b_k) такие, что: $\mathsf{P}(a_k < \xi_k < b_k) \leqslant \tau$, $|b_k - a_k| \geqslant \tau\varepsilon/2 = 5\delta/2$, $\max\{\mathsf{P}(\xi_k \leqslant a_k), \mathsf{P}(\xi_k \geqslant b_k)\} \leqslant 1 - \varepsilon/3$. Обозначим через

$$\Gamma := \{k\colon \xi_k \in (a_k, b_k)\}$$

множество координат с "промежуточными" значениями. Имеем $\mathsf{E}|\Gamma| \leqslant \tau N$; в силу следствия неравенства Бернштейна (1.26), $\mathsf{P}(|\Gamma| > 2\tau N) \leqslant 2\exp(-c\tau N)$. Следовательно,

$$\mathsf{P}(|\Lambda| \leqslant \delta N) \leqslant \mathsf{P}(|\Lambda| \leqslant \delta N,\ |\Gamma| \leqslant 2\tau N) + 2\exp(-c\tau N). \tag{1.61}$$

Зафиксируем некоторые множества координат $\Lambda^\circ, \Gamma^\circ \subset \{1, \ldots, N\}$ размеров $|\Lambda^\circ| \leqslant \delta N$, $|\Gamma^\circ| \leqslant 2\tau N$. Положим $N' := N - |\Lambda^\circ \cup \Gamma^\circ|$; тогда $N' \geqslant N/2$. Для $x \in \mathbb{R}^N$ через $x' \in \mathbb{R}^{N'}$ будем обозначать вектор x, из которого удалили координаты из множества $\Lambda^\circ \cup \Gamma^\circ$; аналогичные обозначения будем применять для случайных векторов и подмножеств $\mathbb{R}^N$. Рассмотрим событие $\{\Lambda = \Lambda^\circ,\ \Gamma = \Gamma^\circ\}$. При наступлении этого события имеем $\|\xi' - \eta'\|_{\ell_\infty^{N'}} \leqslant \delta$, причём $\eta' \in Q_n'$. Следовательно,

$$d_n(K, \ell_\infty^{N'}) \leqslant \delta, \quad \text{где } K := \{\xi'(\omega)\colon \Lambda(\omega) = \Lambda^\circ,\ \Gamma(\omega) = \Gamma^\circ\}. \tag{1.62}$$

Рассмотрим случайный вектор π с координатами $\pi_k := -1$ при $\xi_k \leqslant a_k$; $\pi_k := 0$ при $a_k < \xi_k < b_k$; $\pi_k := 1$ при $\xi_k \geqslant b_k$. По построению $\pi' \in \{-1, 1\}^{N'}$. Докажем, что множество $S := \{\pi'(\omega)\colon \Lambda(\omega) = \Lambda^\circ,\ \Gamma(\omega) = \Gamma^\circ\}$ возможных значений π' имеет мощность не больше $(eN'/n)^n$. Действительно, в противном случае из (1.34), (1.35) следует, что S содержит "куб" размерности $n+1$ и поэтому

$$\operatorname{vc}(K, t) \geqslant n+1, \quad t := \frac{1}{2}\min(b_k - a_k) \geqslant 5\delta/4,$$

откуда, в силу (1.33) имеем $d_n(K, \ell_\infty^{N'}) \geqslant 5\delta/4$, что противоречит (1.62). Для любого $s \in S$, в силу выбора a_k, b_k и независимости координат, имеем $\mathsf{P}(\pi' = s) \leqslant (1-\varepsilon/3)^{N'}$, откуда

$$\mathsf{P}(\Lambda = \Lambda^\circ,\ \Gamma = \Gamma^\circ) \leqslant (eN'/n)^n (1-\varepsilon/3)^{N'} \leqslant (eN/n)^n (1-\varepsilon/3)^{N/2}.$$

Теперь мы можем оценить вероятность в (1.61):

$$\mathsf{P}(|\Lambda| \leqslant \delta N,\ |\Gamma| \leqslant 2\tau N) = \sum_{\substack{|\Lambda^\circ| \leqslant \delta N \\ |\Gamma^\circ| \leqslant 2\tau N}} \mathsf{P}(\Lambda = \Lambda^\circ,\ \Gamma = \Gamma^\circ) \leqslant$$
$$\leqslant (eN/(\delta N))^{\delta N} \cdot (eN/(2\tau N))^{2\tau N} \cdot (eN/n)^n \cdot (1-\varepsilon/3)^{N/2}.$$

Остаётся выбрать параметр δ достаточно малым, чтобы гарантировать экспоненциальную оценку для этого выражения. □

Утверждение 1.12. *В условиях утверждения 1.11 мы имеем для $\delta_1 := \delta_1(\varepsilon)$:*

$$d_n^{\text{avg}}(\{\xi_1, \ldots, \xi_N\}, L_0) \geqslant \delta_1 \quad \textit{при } n \leqslant \delta_1 N.$$

Доказательство. Докажем, что для любого случайного вектора ξ в $\mathbb{R}^N$

$$d_n^{\text{avg}}(\{\xi_1, \ldots, \xi_N\}, L_0) \geqslant c\, d_n^{\text{avg}}(\xi, L_0^N)^4. \tag{1.63}$$

Пусть $d_n^{\text{avg}}(\{\xi_1, \ldots, \xi_N\}, L_0) < \varepsilon$. Тогда для некоторых случайных величин $\eta_1, \ldots, \eta_N$ из n-мерного подпространства L_0 мы имеем $\sum \|\xi_k - \eta_k\|_{L_0} \leqslant N\varepsilon$. Определим случайные множества $\Lambda_t := \{k\colon |\xi_k - \eta_k| > t\}$; ясно, что $\|\xi - \eta\|_{L_0^N} > t$ тогда и только тогда, когда $|\Lambda_t| > tN$.

Возьмём $\delta := \sqrt{\varepsilon}$ и оценим

$$\mathsf{E}|\Lambda_\delta| = \sum_{k=1}^N \mathsf{P}(|\xi_k - \eta_k| > \delta) \leqslant \delta N + |\{k\colon \mathsf{P}(|\xi_k - \eta_k| > \delta) > \delta\}| \leqslant$$
$$\leqslant \delta N + |\{k\colon \|\xi_k - \eta_k\|_{L_0} > \delta\}| \leqslant \delta N + \varepsilon N/\delta = 2\delta N.$$

Далее, для $\gamma := \sqrt{2\delta}$ мы имеем

$$\mathsf{P}(\|\xi - \eta\|_{L_0^N} > \gamma) \leqslant \mathsf{P}(|\Lambda_\gamma| > \gamma N) \leqslant \frac{\mathsf{E}|\Lambda_\gamma|}{\gamma N} \leqslant \frac{\mathsf{E}|\Lambda_\delta|}{\gamma N} \leqslant \frac{2\delta N}{\gamma N} = \gamma.$$

Следовательно, $\mathsf{E}\|\xi - \eta\|_{L_0^N} \leqslant 2\gamma = c\varepsilon^{1/4}$.

Неравенство (1.63) доказано. Воспользуемся теперь утверждением 1.11. Если N достаточно велико, то $2\exp(-\delta N) < 1/2$, и для любого n-мерного подпространства имеем

$$\mathsf{E}\rho(\xi, Q_n)_{L_0^N} \geqslant \delta\, \mathsf{P}(\rho(\xi, Q_n)_{L_0^N} > \delta) \geqslant \delta/2,$$

что даёт оценку $d_n(\{\xi_1, \ldots, \xi_N\}, L_0) \geqslant c(\delta/2)^4$. Если же $N \leqslant N(\delta)$, то за счёт выбора малого δ_1 неравенство $n \leqslant \delta_1 N$ влечёт $n = 0$, и оценка поперечника тривиальна. □

Лакунарные системы. Пусть $\lambda > 1$. Будем говорить, что последовательность натуральных чисел $k_1, \ldots, k_N$ является λ-лакунарной, если $k_{j+1}/k_j \geqslant \lambda$ при $j = 1, \ldots, N-1$.

Утверждение 1.13. *Пусть φ — борелевская интегрируемая по Риману функция на $\mathbb{T}$ и для любого $a \in \mathbb{R}$ имеем $\mu\{x \in \mathbb{T}\colon \varphi(x) \neq a\} > 0$.*

1. *Существует $\lambda = \lambda(\varphi) > 1$ такое, что для любой λ-лакунарной последовательности $k_1, \ldots, k_N$ имеем*

$$d_n^{\mathrm{avg}}(\{\varphi(k_1 x), \ldots, \varphi(k_N x)\}, L_0(\mathbb{T})) \geqslant c(\varphi), \quad \text{при } n \leqslant c(\varphi)N.$$

2. *Для любого $\varepsilon > 0$ существует $\lambda = \lambda(\varphi, \varepsilon) > 1$ такое, что для всякой λ-лакунарной последовательности $k_1, \ldots, k_N$, имеем*

$$d_n^{\mathrm{avg}}(\{\varphi(k_1 x), \ldots, \varphi(k_N x)\}, L_1(\mathbb{T})) \geqslant c(\varphi, \varepsilon), \quad \text{при } n \leqslant N(1-\varepsilon).$$

Мы используем следующий количественный результат.

Теорема 1.C ([90], Theorem 4.3)**.** *Пусть (k_j) — возрастающая последовательность натуральных чисел. Существует последовательность (g_j) измеримых функций на $(0,1)$, независимых и равномерно распределённых на $(0,1)$, для которых*

$$\mu\{x \in [0,1]\colon |\{k_j x\} - g_j(x)| > 2k_j/k_{j+1}\} \leqslant 2k_j/k_{j+1}, \quad j = 1, 2, \ldots.$$

Таким образом, g_k определены на $(0,1)$, принимают значения в $(0,1)$, $\mu\{x\colon g_k(x) < t\} = t$ при $t \in (0,1)$, и случайные величины $g_k(\omega)$, $\omega \in \Omega := [0,1]$, независимы.

Перейдём к доказательству.

Доказательство. Рассмотрим λ-лакунарную последовательность $k_1, \dots, k_N$ для некоторого $\lambda > 1$. Применяя теорему 1.C, получим функции $g_1, \dots, g_N$, для которых $\|\{k_j x\} - g_j(x)\|_{L_0[0,1]} \leqslant 2/\lambda$. Перенесём их на окружность: $\widetilde{g}_j(x) := 2\pi g(x/(2\pi))$, тогда $\|(k_j x \bmod 2\pi) - \widetilde{g}_j(x)\|_{L_0(\mathbb{T})} \leqslant 4\pi/\lambda$.

Положим $f_j(x) := \varphi(\widetilde{g}_j(x))$. Доказательство построено следующим образом: сначала мы покажем, что расстояние между функциями f_j и $\varphi(k_j x)$ можно сделать сколь угодно малым за счёт достаточно большого λ, потом мы докажем жёсткость системы $\{f_j\}$, что повлечёт жёсткость $\{\varphi(k_j x)\}$.

Имеем

$$\int_{\mathbb{T}} |\varphi(k_j x) - f_j(x)|\,dx = \int\limits_{A} |\varphi(k_j x) - \varphi(\widetilde{g}_j(x))|\,dx + \int\limits_{\mathbb{T}\setminus A} |\varphi(k_j x) - \varphi(\widetilde{g}_j(x))|\,dx,$$

где $A = \{x\colon |(k_j x \bmod 2\pi) - \widetilde{g}_j(x)| > 4\pi/\lambda\}$. Первый интеграл не превосходит $2\|\varphi\|_\infty \mu(A) \leqslant \|\varphi\|_\infty \cdot 8\pi/\lambda$. Второй интеграл не больше

$$\int_{\mathbb{T}} \sup_{|x-t|\leqslant 4\pi/\lambda} |\varphi(t) - \varphi(x)|\,dx =: \tau(\varphi, 8\pi/\lambda),$$

где τ — усреднённый модуль непрерывности. Известно, что для интегрируемых по Риману функций $\lim_{h\to 0} \tau(\varphi, h) = 0$. Следовательно, для достаточно больших λ число $\max_j \|\varphi(k_j x) - \varphi(\widetilde{g}_j(x))\|_1$ мало. То же верно и для метрики L_0 в силу (1.3).

В случае L_0 мы можем воспользоваться утверждением 1.12, которое обеспечивает жёсткость $\{f_j\}$ и, в силу близости f_j и $\varphi(k_j x)$, требуемую жёсткость $\{\varphi(k_j x)\}$.

Перейдём ко случаю L_1. Предположим сначала, что $\int_{\mathbb{T}} \varphi = 0$. Тогда f_j независимы, $\int_{\mathbb{T}} f_j = 0$, $\int_{\mathbb{T}} |f_j| = \int_{\mathbb{T}} |\varphi|$ и применима теорема 1.2, которая обеспечивает жёсткость $\{f_j\}$. Общий случай сводится к случаю нулевого среднего следующим образом. Пусть $I := (1/2\pi)\int_{\mathbb{T}} \varphi$, тогда $f_j = \widetilde{f}_j + I$, где $\widetilde{f}_j$ имеют нулевое среднее. Имеем

$$d_n(\{f_1, \dots, f_N\}, L_1) \geqslant d_{n+1}(\{\widetilde{f}_1, \dots, \widetilde{f}_N\}, L_1).$$

Если N достаточно велико, скажем, $N > N_0 := \lfloor 2/\varepsilon \rfloor$, то из неравенства $n \leqslant N(1-\varepsilon)$ следует $n+1 \leqslant N(1-\varepsilon/2)$ и применимы прежние рассуждения. Если $N \leqslant N_0$, то достаточно доказать неравенство

$$d_{N_0-1}(\{f_1, \dots, f_{N_0}\}, L_1) \geqslant c(N_0, \varphi).$$

Поскольку f_j независимы, между ними нет и линейной зависимости и поперечник отделён от нуля. □

1.4 Приближение специальных систем функций

Этот параграф посвящён приближению двух систем характеров — системы Уолша и тригонометрической системы. В нормированном случае L_q, $q \geqslant 1$ удалось получить результат только для системы Уолша.

Напомним определение системы Уолша в нумерации Пэли. Это система функций $w_0, w_1, \dots, w_N, \dots$, заданных на отрезке $[0,1]$ и принимающих значения ± 1. Пусть $m \in \mathbb{Z}_+$ и $m = \sum_{j=1}^{k} m_j 2^{j-1}$, где $m_j \in \{0,1\}$, тогда функция w_m определяется следующим образом:

$$w_m(t) = (-1)^{\sum_{j=1}^{k} m_j t_j}, \quad \text{где } t = \sum_{j \geqslant 1} 2^{-j} t_j,\ t_j \in \{0,1\}. \tag{1.64}$$

Теорема 1.4. *Пусть $w_0, w_1, w_2, \dots$ — система Уолша в нумерации Пэли. Для любого $q \in [1,2)$ существуют $\delta_1 = \delta(q) > 0$, $\delta_2 = \delta_2(q) > 0$ такие, что справедливо неравенство*

$$d_n(\{w_0, w_1, \dots, w_{N-1}\}, L_q[0,1]) \leqslant C_q N^{-\delta_1}, \quad \textit{при } n \geqslant N^{1-\delta_2}.$$

Кроме того,

$$d_n(\{w_0, w_1, \dots, w_{N-1}\}, L_q[0,1]) \geqslant c \min\{1, n^{-1/2} N^{1/q'}\}, \quad \textit{при } n \leqslant N/4.$$

Из доказательства видно, что в качестве δ_1 и δ_2 можно взять любые числа из интервала $(0,1)$, удовлетворяющие неравенству

$$\delta_1 + C\delta_2^{1/2} \log(2/\delta_2) \leqslant \frac{1}{q} - \frac{1}{2}.$$

Этот результат показывает содержательность нижних оценок, полученных в §1.3. Ортонормированности и равномерной ограниченности не достаточно для жёсткости в L_q при $q < 2$ — нужны более сильные условия.

Из теоремы мы видим, что при $q \in (1,2)$ порядок размерности n, для которой n-поперечник N функций системы Уолша стремится к нулю, находится между $N^{2/q'}$ и $N^{1-\delta_2}$.

Характеры. Напомним, что характером конечной абелевой группы G называется гомоморфизм $G \to \mathbb{C}^*$. Характеры образуют абелеву группу по умножению; она обозначается $\widehat{G}$. При этом $|G| = |\widehat{G}|$. Мы рассматриваем два случая: $G = \mathbb{Z}_2^k$, соответствующий системе Уолша, и $G = \mathbb{Z}_N$, соответствующей дискретной тригонометрической системе.

Группу $G = \mathbb{Z}_2^k$ мы отождествляем с $\{0,1\}^k$. Характеры в этом случае вещественны и имеют вид:

$$W_x \colon y \mapsto (-1)^{\langle x,y\rangle}, \quad x, y \in \{0,1\}^k, \tag{1.65}$$

где $\langle x, y\rangle = \sum_{i=1}^k x_i y_i$.

В этом параграфе мы используем обозначение

$$e(\theta) := \exp(2\pi i\theta).$$

Характеры группы $\mathbb{Z}_N$ мы будем использовать чуть позже; они имеют вид

$$y \mapsto e(xy/N), \quad x, y \in \mathbb{Z}_N.$$

Приближение матрицы Уолша–Адамара. Матрицу значений характеров на группе $\mathbb{Z}_2^k$ обозначим через W^k:

$$W^k(x,y) := (-1)^{\langle x,y\rangle}, \quad x, y \in \{0,1\}^k$$

(здесь мы индексируем строки и столбцы матрицы булевыми векторами). Эта матрица называется матрицей Уолша–Адамара.

Утверждение 1.14. *Пусть $\varepsilon \in (0, \frac{1}{2})$. Для любого достаточно большого k существует матрица $B \in \mathbb{R}^{N\times N}$, $N = 2^k$, для которой выполнены свойства:*

(i) в каждой строке B отличается от W^k не более чем в $N^{c_1\varepsilon\log(1/\varepsilon)}$ позициях,

(ii) $\operatorname{rank} B \leqslant N^{1-c_2\varepsilon^2}$,

(iii) $\|B\|_C \leqslant N^{\frac{1}{2}+c_3\varepsilon\log(1/\varepsilon)}$.

Напомним, что $\|B\|_C = \max|B(x,y)|$. Как и W^k, матрица B индексируется булевыми векторами.

Выведем теорему из этого утверждения.

Доказательство теоремы 1.4. Оценка снизу вытекает из утверждения 1.5 (при $q = 1$ её можно получить предельным переходом). Докажем оценку сверху.

Можем считать, что $N = 2^k$. По определению (1.64), для каждого $m < 2^k$ значения функции w_m на двоичных отрезках длины 2^{-k} получаются из элементов соответствующей строки матрицы W^k. Возьмём некоторое $\varepsilon > 0$, применим утверждение 1.14 и оценим погрешность приближения строк в L_q^N:

$$(N^{-1}\sum_{y\in\{0,1\}^k}|B(x,y)-W^k(x,y)|^q)^{\frac{1}{q}} \leqslant$$
$$\leqslant N^{-\frac{1}{q}}(1+\|B\|_C)\cdot|\{y\colon B(x,y)\neq W^k(x,y)\}|^{\frac{1}{q}} \leqslant$$
$$\leqslant N^{-\frac{1}{q}+\frac{1}{2}+c\varepsilon\log(1/\varepsilon)} \leqslant N^{-\delta_1},$$

если $\varepsilon = \varepsilon(q)$ достаточно мало. При этом $n = \operatorname{rank} B \leqslant N^{1-c_2\varepsilon^2} = N^{1-\delta_2}$. □

Лемма 1.A ([83], Lemma 3.3)**.** *Для любых параметров $a, b \in (0, 1/5)$, и для любого вектора $x \in \{0,1\}^k$ с $\|x\|_1 \in [(1/2-a)k, (1/2+a)k]$, количество векторов $y \in \{0,1\}^k$, удовлетворяющих обоим условиям*

- $\|y\|_1 \in [(1/2-a)k, (1/2+a)k]$,
- $\sum_{j=1}^k x_j y_j \leqslant bk$,

не превосходит $(2ak+1)(bk+1)2^{f(a,b)k}$, где функция $f(a,b) \to 0$ при $a, b \to 0$.

В доказательстве авторы выписывают оценку сверху на функцию f, из которой следует, что

$$f(a,b) \leqslant C(a+b)\log\frac{1}{a+b}. \tag{1.66}$$

Докажем теперь утверждение 1.14. Рассуждение в основном повторяет конструкцию из работы [83], в которой были установлены свойства (i) и (ii).

Доказательство. Зафиксируем достаточно большое k и достаточно малое ε. Положим

$$l := \lceil 2k\varepsilon \rceil - 1, \quad d := \lceil (1/2 - \varepsilon)k \rceil.$$

Определим множество "типичных" точек куба:

$$\mathcal{C}_{k,\varepsilon} := \{x \in \{0,1\}^k \colon \|x\|_1 \in [(1/2-\varepsilon)k, (1/2+\varepsilon)k]\}.$$

Построим многочлен $q(z_1, \ldots, z_k)$, $\deg q \leqslant d$, такой что

$$q(z) = (-1)^{\|z\|_1}, \quad \text{если } z \in \{0,1\}^k \text{ и } l+1 \leqslant \|z\|_1 \leqslant l+d+1.$$

В [83] строился многочлен с целыми коэффициентами, чтобы рассуждение проходило над любым полем. Поскольку нас интересует поле $\mathbb{R}$, мы возьмём просто $q(z) = (-1)^{l+1} Q_d(z_1 + \ldots + z_k - l - 1)$, где многочлен $Q_d \in \mathcal{P}_d$ определяется условиями

$$Q_d(j) = (-1)^j, \ j = 0, 1, \ldots, d.$$

Положим $p(x,y) := q(x_1y_1, \ldots, x_ky_k)$. Заменим в p все выражения вида $x_i^m y_i^m$, $m > 1$, на x_iy_i и получим другой многочлен $\widetilde{p}$; при этом, однако, $\widetilde{p}$ совпадает с p на булевом кубе $\{0,1\}^k \times \{0,1\}^k$. Каждый моном в $\widetilde{p}$ имеет вид $\prod_{i \in I} x_iy_i$, $|I| \leqslant d$, поэтому количество мономов M не превосходит (см. (1.36))

$$M(\widetilde{p}) \leqslant \sum_{s=0}^{d} \binom{k}{s} \leqslant 2^{kh_2(d/k)} \leqslant 2^{k(1-c\varepsilon^2)}. \tag{1.67}$$

Положим $\widetilde{B}(x,y) := \widetilde{p}(x,y)$. Ключевое соображение состоит в том, что

$$\operatorname{rank} \widetilde{B} \leqslant M(\widetilde{p}). \tag{1.68}$$

Действительно, каждый моном задаёт одноранговую матрицу.

По построению, $\widetilde{B}(x,y) = W^k(x,y)$, если $\langle x, y \rangle \in [2\varepsilon k, (1/2+\varepsilon)k]$.

Исправим $\widetilde{B}$: оставляем $B(x,y) := \widetilde{B}(x,y)$, если $x, y \in \mathcal{C}_{k,\varepsilon}$, а для остальных пар (x,y) полагаем $B(x,y) := W^k(x,y)$. Изменения коснутся малого количества строк или столбцов, поэтому ранг B ограничивается сверху:

$$\operatorname{rank} B \leqslant \operatorname{rank} \widetilde{B} + 2|\{0,1\}^k \setminus \mathcal{C}_{k,\varepsilon}|, \tag{1.69}$$

причём по неравенству Хёфдинга (1.23), имеем $2^{-k}|\{0,1\}^k \setminus \mathcal{C}_{k\varepsilon}| \leqslant 2\exp(-2k\varepsilon^2)$.

Построенная матрица B и есть требуемая матрица.

Свойство (ii) следует из (1.67), (1.68) и (1.69).

Заметим, что B может отличаться от W^k только при $x, y \in \mathcal{C}_{k,\varepsilon}$, $\langle x, y\rangle \notin [2k\varepsilon, (1/2+\varepsilon)k]$; такие пары (x, y) удовлетворяют условиям

$$\begin{cases} \langle x, y\rangle < 2\varepsilon k, \\ \quad x, y \in \mathcal{C}_{k,\varepsilon}. \end{cases} \tag{1.70}$$

В силу леммы 1.A ($a := \varepsilon$, $b := 2\varepsilon$) и оценки (1.66), для каждого x существует не более $2^{ck\varepsilon \log(1/\varepsilon)}$ векторов y, для которых выполнено (1.70). Отсюда следует свойство (i).

Оценим теперь $\max|B(x, y)|$. Элементы матрицы B либо совпадают с элементами W^k, либо получаются из многочлена p. Следовательно,

$$\|B\|_C \leqslant \max\{1, \max_{(x,y)\colon (1.70)} |p(x, y)|\}.$$

Имеем:

$$\max_{(x,y)\colon (1.70)} |p(x, y)| \leqslant \max_{\substack{z\in\{0,1\}^k \\ \sum z_i < 2\varepsilon k}} |q(z)| = \max_{\substack{z\in\{0,1\}^k \\ \sum z_i \leqslant l}} |q(z)| = \max_{s=1,\dots,l+1} |Q_d(-s)|.$$

Выпишем многочлен Q_d явно, пользуясь интерполяционной формулой Ньютона:

$$Q_d(t) = \sum_{j=0}^{d} (-2)^j \binom{t}{j}, \quad \text{где } \binom{t}{j} = \frac{t(t-1)\cdots(t-j+1)}{j!}.$$

Имеем (при $1 \leqslant s \leqslant l+1$):

$$\left|\binom{-s}{j}\right| = \frac{|-s|\cdots|-s-j+1|}{j!} = \binom{s+j-1}{j} \leqslant \binom{l+j}{j} = \binom{l+j}{l} \leqslant \binom{l+d}{l}.$$

Следовательно, $|Q_d(-s)| \leqslant 2^{d+1}\binom{l+d}{l}$. Наконец, используя оценку биномиальных коэффициентов (1.36), мы получим:

$$\max_{1\leqslant s\leqslant l+1} |Q(-s)| \leqslant 2^{d+1}\binom{l+d}{l} \leqslant 2^{d+1+(l+d)h_2(\frac{l}{l+d})}.$$

Оценим показатель (при достаточно большом k):

$$\begin{aligned} d+1+(l+d)h_2(\frac{l}{l+d}) \leqslant \\ \leqslant k(1/2-\varepsilon)+2+((1/2+\varepsilon)k+1)h_2(\frac{2k\varepsilon}{(1/2+\varepsilon)k-1}) \leqslant \\ \leqslant k(1/2+h_2(4\varepsilon)+o(1)) \leqslant k(1/2+C\varepsilon\log(1/\varepsilon)). \end{aligned}$$

Собирая вместе полученные неравенства, выводим требуемую оценку на $\|B\|_C$. $\square$

Приближение пространствами малой размерности в L_0 и L_{H}. Мы будем приближать характеры (как функции на G) в метриках $L_0 := L_0(G)$ и $L_{\mathrm{H}} := L_{\mathrm{H}}(G)$, см. §1.1 (считаем, что на G задана равномерная вероятностная мера).

Напомним следующий известный факт (см., например, [138, §2.1]).

Лемма 1.B. *Пусть G — конечная абелева группа, $A, B \subset G$ и $|A|+|B| > |G|$. Тогда $A + B = \{a + b\colon a \in A,\, b \in B\} = G$.*

Из него немедленно вытекает следствие.

Лемма 1.3. *Пусть $A, B \subset \widehat{G}$ и $|A| + |B| > |G|$. Тогда для любых $m, n \in \mathbb{N}$ имеем*

$$d_{mn}(\widehat{G}, L_{\mathrm{H}}) \leqslant d_m(A, L_{\mathrm{H}}) + d_n(B, L_{\mathrm{H}}).$$

Доказательство. Если характеры $\chi \in A$ приближены функциями $g \in Q$, $\dim Q \leqslant m$, с погрешностью ε, а характеры $\zeta \in B$ приближены функциями $h \in V$, $\dim V \leqslant n$, с погрешностью δ, то попарные произведения $\chi\zeta \in AB$ приближены произведениями $gh \in QV$, $\dim QV \leqslant mn$, с погрешностью не более $\varepsilon + \delta$. Применяя лемму 1.B к множествам A, B и группе $\widehat{G}$, получаем, что $AB = \widehat{G}$ и все характеры приближены. $\square$

Следующее утверждение, как и теорема 1.4, доказывается с использованием методов [83]. Это модификация теоремы 1.2 из той работы.

Утверждение 1.15. *Пусть $k \in \mathbb{N}$. Для любого $1 \leqslant \lambda \leqslant \frac{1}{4}\sqrt{k}$ справедливо неравенство*

$$d_n(\widehat{\mathbb{Z}_2^k}, L_{\mathrm{H}}) \leqslant 4\exp(-2\lambda^2), \quad \textit{при } n \geqslant (\sqrt{k}/\lambda)^{8\lambda\sqrt{k}}.$$

Доказательство. Требуется приблизить характеры $\{W_x\}$, см. (1.65), в метрике Хэмминга. Будем применять неравенство Хёфдинга (1.23), которое даёт, что

$$\mathsf{P}(|\sum_{i=1}^{k} X_i - k/2| > tk^{1/2}) \leqslant 2\exp(-2t^2), \quad X \sim \mathrm{Unif}\{0,1\}^k. \tag{1.71}$$

Пусть $\mathcal{X}$ — множество векторов x таких, что $|\sum_{i=1}^k x_i - k/2| \leqslant \sqrt{k}$. Из (1.71) при $t = 1$ получим, что

$$|\mathcal{X}| \geqslant 2^k(1 - 2/e^2) > 2^{k-1}. \tag{1.72}$$

Мы приблизим сначала W_x для $x \in \mathcal{X}$.

Зафиксируем $x \in \mathcal{X}$. Обозначим $I := \{i\colon x_i = 1\}$,

$$Y_I := \left\{ y \in \{0,1\}^k\colon \left| \sum_{i\in I} y_i - \frac{|I|}{2} \right| \leqslant \lambda |I|^{1/2} \right\}.$$

Тогда (см. (1.71)) имеем $2^{-k}|Y_I| \geqslant 1 - 2\exp(-2\lambda^2)$. Кроме того, для всех $y \in Y_I$ имеем

$$|\langle x, y\rangle - k/4| = |\sum_{i\in I} y_i - k/4| \leqslant \left|\sum_{i\in I} y_i - \frac{|I|}{2}\right| + \left|\frac{|I|}{2} - k/4\right| \leqslant$$
$$\leqslant \lambda\sqrt{|I|} + \sqrt{k}/2 \leqslant \lambda\sqrt{k/2 + \sqrt{k}} + \sqrt{k}/2 \leqslant 2\lambda\sqrt{k}.$$

Рассмотрим полином $q(t)$, интерполирующий функцию $(-1)^t$ для $t \in \mathbb{Z} \cap [k/4 - 2\lambda\sqrt{k}, k/4 + 2\lambda\sqrt{k}]$. Его степень d не превосходит $4\lambda\sqrt{k}$; отметим, что $4\lambda\sqrt{k} \leqslant k$. Тогда $W_x(y) = q(\langle x, y\rangle) = (-1)^{\langle x,y\rangle}$ для всех $y \in Y_I$, поэтому

$$\|W_x - q(\langle x, \cdot\rangle)\|_{L_\text{н}} \leqslant 2\exp(-2\lambda^2).$$

Оценим линейную размерность семейства функций $\{q(\langle x, \cdot\rangle)\}_x$ (тут достаточно брать $x \in \mathcal{X}$, но нам проще рассмотреть все $x \in \{0,1\}^k$). Размерность равна рангу матрицы $(q(\langle x, y\rangle))_{x,y\in\{0,1\}^k}$, где x индексирует строки, а y столбцы. Этот ранг не превосходит количества мономов (после приведения $x_i^m y_i^m \mapsto x_i y_i$) в полиноме $q(x_1y_1 + \ldots + x_ky_k)$, аналогично (1.68). Моном это произведение некоторых x_iy_i, и их количество оценивается:

$$\operatorname{rank}(q(\langle x, y\rangle)_{x,y\in\{0,1\}^k}) \leqslant \binom{k}{0} + \ldots + \binom{k}{d} \leqslant (ek/d)^d \leqslant$$
$$\leqslant (e\sqrt{k}/4\lambda)^{4\lambda\sqrt{k}} \leqslant (\sqrt{k}/\lambda)^{4\lambda\sqrt{k}}.$$

Для завершения доказательства применим лемму 1.3 ко множеству $A = B = \{W_x\colon x \in \mathcal{X}\}$; условие леммы выполнено в силу (1.72). □

Следствие 1.5. *Пусть $a, b > 0$, $a + 2b < 1$. Тогда при достаточно больших N имеем*

$$d_n(\{w_0, \dots, w_{N-1}\}, L_{\mathrm{H}}[0,1]) \leqslant \exp(-\log^a N), \quad \textit{при } n \geqslant \exp(\log^{1-b} N).$$

Доказательство. Можно считать, что $N = 2^k$ и приближать характеры $\widehat{\mathbb{Z}_2^k}$. Возьмём $\lambda \asymp k^{a/2}$ в утверждении 1.15, это даст требуемую оценку вероятности. При этом

$$(\sqrt{k}/\lambda)^{8\lambda\sqrt{k}} = \exp(k^{1/2+a/2+o(1)}),$$

и по условию $1/2 + a/2 < 1 - b$. □

Перейдём к приближению тригонометрической системы. До сих пор мы рассматривали приближение вещественных функций в вещественных пространствах; теперь характеры будут комплекснозначные, и требуется комментарий. Нам будет удобно работать в пространствах над полем $\mathbb{C}$; будем указывать поле в обозначениях: $L_q^{\mathbb{C}}$, $L_0^{\mathbb{C}}$, $L_{\mathrm{H}}^{\mathbb{C}}$. Если n-поперечник рассматривается в комплексном пространстве, естественно рассматривать комплексные подпространства:

$$d_n(K, X^{\mathbb{C}}) = \inf_{y_1,\dots,y_n \in X^{\mathbb{C}}} \sup_{x \in K} \inf_{c_1,\dots,c_n \in \mathbb{C}} \|x - \sum_{k=1}^{n} c_k y_k\|.$$

Мы применяли лемму 1.3 в случае, когда все характеры вещественны; лемма справедлива и над полем $\mathbb{C}$. Впрочем, вещественная и комплексная размерность отличаются в 2 раза и для нас это отличие не принципиально.

В следующем утверждении рассматриваются характеры на группе $\mathbb{Z}_N$ для $N = 2^k$.

Утверждение 1.16. *Пусть $k \in \mathbb{N}$. Для любого $1 \leqslant \lambda \leqslant c_0 k^{1/4}$ имеем*

$$d_n(\widehat{\mathbb{Z}_{2^k}}, L_0^{\mathbb{C}}) \leqslant c_1 k \exp(-c_2\lambda^2), \quad \textit{при } n \geqslant (\sqrt{k}/\lambda)^{c_3\lambda^3\sqrt{k}}.$$

Доказательство. Наша задача — приблизить характеры $y \mapsto e(xy/2^k)$. Запишем $x = \sum_{i=0}^{k-1} x_i 2^i$, $y = \sum_{j=0}^{k-1} y_j 2^j$ и идентифицируем $x, y \in \mathbb{Z}_{2^k}$ с двоичными векторами $(x_0, \dots, x_{k-1})$, $(y_0, \dots, y_{k-1})$. Пользуясь 1-периодичностью функции $e(\cdot)$, получим

$$e(xy/2^k) = e(\sum_{i,j=0}^{k-1} x_i y_j 2^{i+j-k}) = e(\sum_{s=1}^{k} 2^{-s} \sum_{i+j=k-s} x_i y_j).$$

Зафиксируем число $s_0 \leqslant \sqrt{k}$ и оценим "хвост" суммы:

$$\sum_{s>s_0} 2^{-s} \sum_{i+j=k-s} x_i y_j \leqslant k \sum_{s>s_0} 2^{-s} \leqslant k 2^{-s_0}.$$

Поскольку функция $e(\cdot)$ 2π-липшицева, имеет место равномерное приближение

$$|e(xy/2^k) - \psi_x(y)| \leqslant 2\pi k 2^{-s_0}, \quad \text{где } \psi_x(y) := e(\sum_{s=1}^{s_0} 2^{-s} \sum_{i+j=k-s} x_i y_j). \qquad (1.73)$$

Следовательно, достаточно приблизить функции ψ_x.

Так же, как и в предыдущем доказательстве, сперва мы рассмотрим векторы, лежащие во множестве $\mathcal{X} := \{x\colon |\sum_{i=0}^{k-1} x_i - k/2| \leqslant \sqrt{k}\}$. Зафиксируем $x \in \mathcal{X}$.

Пусть $s \in \{1, \dots, s_0\}$. Имеем $|\sum_{i=0}^{k-s} x_i - k/2| \leqslant \sqrt{k} + s \leqslant 2\sqrt{k}$. Обозначим $J_s := \{j \in \{0, \dots, k-s\}\colon x_{k-s-j} = 1\}$, тогда $||J_s| - k/2| \leqslant 2\sqrt{k}$. Для случайного вектора $y \in \{0,1\}^k$ с вероятностью не менее $1 - 2\exp(-2\lambda^2)$ выполнено неравенство

$$|\sum_{j\in J_s} y_j - \frac{1}{2}|J_s|| \leqslant \lambda |J_s|^{1/2}.$$

Для таких y имеем:

$$\left|\sum_{i+j=k-s} x_i y_j - \frac{k}{4}\right| = \left|\sum_{j\in J_s} y_j - \frac{k}{4}\right| \leqslant \left|\sum_{j\in J_s} y_j - \frac{|J_s|}{2}\right| + \left|\frac{|J_s|}{2} - \frac{k}{4}\right| \leqslant$$

$$\leqslant \lambda\sqrt{\frac{k}{2} + 2\sqrt{k}} + \sqrt{k} \leqslant 2\lambda\sqrt{k}.$$

(Здесь мы использовали то, что можно считать k достаточно большим, т.к. $c_0 k^{1/4} \geqslant 1$.)

Как и раньше, воспользуемся техникой интерполяции. Существует полином q_s степени $d \leqslant 4\lambda\sqrt{k}$, который интерполирует функцию $t \mapsto e(2^{-s}t)$ в точках $\mathbb{Z} \cap [k/4 - 2\lambda\sqrt{k}, k/4 + 2\lambda\sqrt{k}]$. Тогда функция $e(2^{-s}\sum_{i+j=k-s} x_i y_j)$ приближается функцией $\phi_x^s(y) := q_s(\sum_{i+j=k-s} x_i y_j)$ в метрике Хэмминга с погрешностью не более $2\exp(-2\lambda^2)$.

Размерность множества построенных функций, $\dim\{\phi_x^s(\cdot)\}_x$, не превосходит, как и в предыдущем доказательстве, $(ek/d)^d \leqslant (\sqrt{k}/\lambda)^{4\lambda\sqrt{k}}$.

Заметим, что приближаемая функция представляет собой произведение:

$$\psi_x(y) = \prod_{s=1}^{s_0} e(2^{-s} \sum_{i+j=k-s} x_i y_j). \qquad (1.74)$$

Мы приблизим это произведение произведением соответствующих полиномов: $\psi_x \approx \prod_{s=1}^{s_0} \phi_x^s$. При этом погрешность в метрике Хэмминга не превосходит $2s_0 \exp(-2\lambda^2)$. Ранг произведения не превосходит произведения рангов, т.е. $(\sqrt{k}/\lambda)^{4s_0\lambda\sqrt{k}}$.

Далее, используем лемму 1.3, чтобы приблизить ψ_x в L_{H} для всех x.

Наконец, неравенство (1.73) даёт приближение исходных характеров в L_0 с погрешностью $\max\{2\pi k 2^{-s_0}, 4s_0 \exp(-2\lambda^2)\}$, причём используемая размерность не больше $(\sqrt{k}/\lambda)^{8s_0\lambda\sqrt{k}}$. Остаётся положить $s_0 := \lfloor \lambda^2 \rfloor$. □

Следствие 1.6. *Пусть $a, b > 0$, $3a + 2b < 1$. При достаточно больших N справедливо неравенство*

$$d_n(\{e^{i\nu x}\}_{\nu=-N}^{N}, L_0^{\mathbb{C}}(\mathbb{T})) \leqslant \exp(-\log^a N), \quad \textit{при } n \geqslant \exp(\log^{1-b} N).$$

Доказательство. Нам будет удобнее приближать систему $(e^{i\nu x})_{\nu=0}^{2N}$. Аппроксимируем эти функции, приблизив их кусочно-постоянными с дроблением $\mathbb{T}$ на $M := 2^k \asymp N^2$ отрезков; это даст небольшую равеномерную погрешность $O(N/M) = O(N^{-1})$. Мы получим матрицу

$$F(\nu, m) := e(\nu m/M), \quad 0 \leqslant \nu \leqslant 2N,\ 0 \leqslant m < M.$$

Приблизим строки этой матрицы в L_0, используя утверждение 1.16. Возьмём $\lambda \asymp k^{a/2}$, тогда $(\sqrt{k}/\lambda)^{c_3\lambda^3\sqrt{k}} = k^{1/2+3a/2+o(1)}$, и по условию $1/2+3a/2 < 1-b$. □

Таким образом, в L_0 возможна аппроксимация пространствами очень малой размерности.

Тригонометрическая система в L_q, $q < 1$. Нам потребуются пространства комплексных тригонометрических полиномов:

$$\mathcal{T}^{\mathbb{C}}(\Lambda) := \left\{ \sum_{k\in\Lambda} c_k e^{ikx} \colon c_k \in \mathbb{C} \right\}, \quad \Lambda \subset \mathbb{Z}.$$

Тригонометрический поперечник d_n^T соответствует аппроксимации подпространствами, натянутыми на n функций тригонометрической системы:

$$d_n^T(K, L_q^{\mathbb{C}}(\mathbb{T})) := \inf_{\substack{\Lambda\subset\mathbb{Z}\\|\Lambda|=n}} \sup_{f\in K} \inf_{T\in\mathcal{T}^{\mathbb{C}}(\Lambda)} \|f - T\|_q.$$

Утверждение 1.17. *Для любого $q \in (0,1)$ при достаточно больших N справедливо неравенство*

$$d_n^T(\{e^{ikx}\}_{k=-N}^N, L_q^{\mathbb{C}}(\mathbb{T})) \leqslant \log^{-c(q)} N, \quad \text{при } n \geqslant N\exp(-\log^c N). \tag{1.75}$$

Отметим, что определение поперечников давалось для метрических пространств; при $q < 1$ возникает квазиметрика, но это ничему не мешает.

Лемма 1.4. *Для любого $q \in (0,1)$ существует $\delta = \delta(q) > 0$ такое, что для всех $m \in \mathbb{N}$ имеем*

$$\min\{\|T\|_q \colon T \in \mathcal{T}_m,\ \widehat{T}(0) = 1\} \leqslant C(q)m^{-\delta},$$

Доказательство. В силу монотонности L_q-нормы по q, достаточно рассмотреть случай $q > 1/2$. Возьмём в качестве T ядро Фейера K_m и воспользуемся оценкой (1.8). Имеем

$$\int_{-\pi}^{\pi} |K_m(x)|^q\,dx = \int_{|x|\leqslant 1/m} |K_m(x)|^q\,dx + \int_{|x|>1/m} |K_m(x)|^q\,dx \lesssim$$
$$\lesssim m^{q-1} + \int_{1/m}^{\pi} m^{-q}x^{-2q}\,dx \leqslant m^{q-1} + m^{-q}m^{2q-1} = 2m^{q-1},$$

что даёт нужное неравенство с $\delta = 1/q - 1$. □

Лемма 1.5. *Пусть $m, N \in \mathbb{N}$, $N \geqslant Cm^2$. Существует множество $\Lambda \subset \mathbb{Z}$ мощности*

$$|\Lambda| \leqslant cN^{1-\frac{1}{2m}}\log^{\frac{1}{m}} N,$$

для которого выполнено следующее. Для любого $k \in \{-N, \dots, N\}$ найдётся $h \in \mathbb{Z}$, $h \neq 0$ такое, что $k \pm lh \in \Lambda$, $l = 1, \dots, m$.

Докажем лемму позже; сначала выведем из лемм утверждение 1.17.

Доказательство. Возьмём некоторое число m, применим лемму 1.5 и получим множество Λ, а с помощью леммы 1.4 получим полином $T \in \mathcal{T}_m$. Для аппроксимации e^{ikx}, $k \in \{-N, \dots, N\} \setminus \Lambda$ подпространством $\mathcal{T}(\Lambda)$, мы возьмём $h \in \mathbb{Z}$, $h \neq 0$, для которого $k \pm h, \dots, k \pm mh \in \Lambda$. Рассмотрим полином

$$t(x) := \sum_{l=-m}^{m} \hat{T}(l)e^{i(k+lh)x}.$$

Он равен $e^{ikx} + s(x)$, где $\operatorname{supp} \widehat{s} \subset \Lambda$, поэтому нам достаточно оценить $\|t\|_q$. Заметим, что $t(x) = e^{ikx} T(hx)$,

$$\|t\|_q = \|T(hx)\|_q = \|T(x)\|_q \lesssim m^{-\delta}.$$

Наконец, положим $m \asymp \log^{1/2} N$, это даст требуемую погрешность и $|\Lambda| \leqslant N \exp(-\log^c N)$. □

Теперь мы докажем лемму 1.5.

Доказательство. Введём обозначения $\mathcal{K} := \{-N, \dots, N\}$ и $\mathcal{L} := \{-2N, \dots, 2N\}$.

Пусть $\Lambda \subset \mathcal{L}$ — это случайное подмножество, каждая точка попадает в Λ независимо от остальных с некоторой вероятностью $\tau \in (0, 1)$.

Скажем, что шаг $h \in \mathbb{Z} \setminus \{0\}$ допустим для $k \in \mathcal{K}$, если $k \pm lh \in \Lambda$ при $l = 1, \dots, m$. Имеем (при условии, что $k \pm lh \in \mathcal{L}$):

$$\mathsf{P}(h \text{ допустим для } k) = \tau^{2m}.$$

Назовём элемент $k \in \mathcal{K}$ "плохим", если для него нет допустимых шагов. Предположим, что для любого $k \in \mathcal{K}$ есть как минимум S шагов $h_1, \dots, h_S \in \mathbb{Z} \setminus \{0\}$ таких, что прогрессии $k \pm lh_i$, $l = 1, \dots, m$ лежат во множестве $\mathcal{L}$ и не пересекаются (скоро мы оценим S). Тогда события "шаг h_i допустим для k" независимы для различных i. Следовательно, вероятность того, что k плохой, не превосходит вероятности того, что все h_i недопустимы, т.е. $(1 - \tau^{2m})^S$.

Ожидаемое количество плохих k не превосходит

$$|\mathcal{K}|(1 - \tau^{2m})^S \leqslant (2N + 1) \exp(-S\tau^{2m}),$$

поэтому с вероятностью $\geqslant 2/3$ это количество не превосходит $(6N + 3) \exp(-S\tau^{2m})$. Далее, $\mathsf{E}|\Lambda| = N\tau$, поэтому с вероятностью $\geqslant 2/3$ имеем $|\Lambda| \leqslant 3N\tau$. Следовательно, существует множество Λ, для которого оба неравенства справедливы; зафиксируем это Λ.

Если τ не слишком мало, так что выполнено неравенство

$$(6N + 3) \exp(-S\tau^{2m}) < 1, \tag{1.76}$$

то плохих k вообще не будет и Λ и есть наше искомое множество.

Остаётся оценить S снизу. В качестве шагов h мы можем взять простые числа из отрезка $[1, N/m]$; следовательно,

$$S \geqslant c(N/m)/\log(N/m) \gtrsim m^{-1}N/\log N.$$

Наконец, для выполнения неравенства (1.76) мы положим:

$$\tau = (\ln(7N)/S)^{\frac{1}{2m}} \lesssim (\log^2 N/N)^{\frac{1}{2m}}.$$

В силу построения, $|\Lambda| \leqslant 3N\tau$ и искомая оценка получена. $\square$

Глава 2

Поперечники классов функций критической гладкости

2.1 Класс Соболева W_1^1

Теорема 2.1. *Для любого $q \in (2, \infty)$ и всех натуральных чисел $n \geqslant 2$ справедливо неравенство*

$$c_1(q) n^{-1/2} \log n \leqslant d_n(W_1^1[0,1], L_q[0,1]) \leqslant c_2(q) n^{-1/2} \log n.$$

Рассмотрим "ступеньки" — функции χ_t, заданные как $\chi_t(x) := 1$, $0 \leqslant x < t$ и $\chi_t(x) := -1$, $t \leqslant x \leqslant 1$.

Утверждение 2.1. *Пусть $q \in (2, \infty)$, $\gamma > 1/q$. Существует $\delta = \delta(q, \gamma) > 0$ такое, что для достаточно больших n, произвольного n-мерного подпространства $Q_n \subset L_q[0,1]$ и произвольного семейства функций $\{\eta_t(\cdot)\}_{0 \leqslant t \leqslant 1}$, лежащих в этом подпространстве, выполнено хотя бы одно из двух неравенств:*

$$\left(\int_0^1 \|\chi_t - \eta_t\|_{L_2}^2 \, dt \right)^{1/2} \geqslant \delta n^{-1/2} \log n, \tag{2.1}$$

$$\left(\int_0^1 \|\chi_t - \eta_t\|_{L_q}^q \, dt \right)^{1/q} \geqslant n^{-\gamma}. \tag{2.2}$$

Теперь доказательство теоремы получается почти мгновенно.

Доказательство. Верхняя оценка хорошо известна, см. [50]. (Она также вытекает из вложения $\mathcal{W}_1^1[0,1] \subset \mathcal{B}_{1,\infty}^1[0,1]$ и теоремы 2.2, см. §2.2.) Докажем нижнюю оценку. Фиксируем произвольно $\gamma \in (1/q, 1/2)$, из утверждения 2.1 мы получим некоторое δ. Предположим, что $d_n(W_1^1, L_q) < (\delta/2)n^{-1/2}\log n$. Функции χ_t лежат в замыкании (в L_q) множества $2W_1^1$; следовательно, существует подпространство Q_n и функции $\eta_t \in Q_n$ такие, что $\sup_{0\leqslant t\leqslant 1}\|\chi_t - \eta_t\|_{L_q} < \delta n^{-1/2}\log n$. Тогда имеем

$$\left(\int_0^1 \|\chi_t - \eta_t\|_{L_2}^2\,dt\right)^{1/2} \leqslant \sup_t \|\chi_t - \eta_t\|_{L_q} < \delta n^{-1/2}\log n,$$

поэтому неравенство (2.1) не имеет места. Следовательно,

$$\sup_t \|\chi_t - \eta_t\|_{L_q} \geqslant \left(\int_0^1 \|\chi_t - \eta_t\|_{L_q}^q\,dt\right)^{1/q} \geqslant n^{-\gamma},$$

что противоречит нашему предположению при больших n (поскольку $\gamma < 1/2$). Нижняя оценка доказана. □

Несколько комментариев относительно утверждения 2.1. Во-первых, чтобы были определены интегралы в (2.1) и (2.2), нужна измеримость. Мы, конечно, можем предполагать, что η_t зависит от t непрерывным образом (в качестве η_t можно взять L_q-ближайшую к χ_t функцию в Q_n), или что η_t как функция от t кусочно–постоянна. Во-вторых, из утверждения следует оценка на поперечник $d_n \gtrsim n^{-1/2}\log n$ в метрике $\|f\|_2 + n^{-1/2+1/q+\varepsilon}\|f\|_q$, для любого $\varepsilon > 0$. При этом существует подпространство Q_n, для которого $\rho(W_1^1, Q_n)_2 \asymp d_n(W_1^1, L_2) \asymp n^{-1/2}$ и одновременно $\rho(W_1^1, Q_n)_q \asymp d_n(W_1^1, L_q) \asymp n^{-1/2}\log n$ (возьмём в качестве Q_n сумму экстремальных подпространств размерности $n/2$ для L_2 и L_q).

Доказательство утверждения 2.1. Предположим, неравенство (2.1) не выполнено (число δ выберем позже), докажем (2.2). Обозначим через $c_{k,j}(f)$ коэффициенты в разложении функции $f \in L_1$ по нормированной в L_∞ системе Хаара $\{h_{k,j}\}$:

$$h_{k,j}(x) := \begin{cases} 1, & x \in [\frac{j-1}{2^k}, \frac{j-1/2}{2^k}), \\ -1, & x \in (\frac{j-1/2}{2^k}, \frac{j}{2^k}], \\ 0, & \text{иначе}, \end{cases}$$

$$c_{k,j}(f) := 2^k \int_0^1 f(x)h_{k,j}(x)\,dx = 2^k\Big(\int_{\frac{j-1}{2^k}}^{\frac{j-1/2}{2^k}} f(x)\,dx - \int_{\frac{j-1/2}{2^k}}^{\frac{j}{2^k}} f(x)\,dx\Big), \qquad (2.3)$$

$k = 0, 1, \ldots$, $j = 1, \ldots, 2^k$. Через $P_k f := \sum_{j=1}^{2^k} c_{k,j}(f) h_{k,j}$ обозначим k-ю двоичную пачку ряда Фурье–Хаара f.

Зафиксируем некоторые номера $k_0 < k_1$ (см. (2.8)) и рассмотрим пространство векторов

$$\mathcal{F} := \{x = (x_{k,j})\colon k_0 \leqslant k \leqslant k_1,\ j = 1, \ldots, 2^k\}.$$

Зададим веса координат: $w(k, j) = 2^{-k}$. В полученном весовом пространстве определено скалярное произведение и ℓ_p-нормы:

$$\langle x, y\rangle_w := \sum_{k=k_0}^{k_1} 2^{-k} \sum_{j=1}^{2^k} x_{k,j} y_{k,j}, \quad \|x\|_{p,w} := \left(\sum_{k=k_0}^{k_1} 2^{-k} \sum_{j=1}^{2^k} |x_{k,j}|^p \right)^{1/p}.$$

При этом $\langle x, y\rangle_w \leqslant \|x\|_{p,w} \|y\|_{p',w}$.

В силу хорошо известной теоремы Марцинкевича–Пэли (см., например, [43, Гл. 3]),

$$\|f\|_{L_q}^q \asymp \|((\int_0^1 f(x)\,dx)^2 + \sum_{k \geqslant 0} (P_k f)^2)^{1/2}\|_{L_q}^q.$$

Используя неравенство (1.1) для $p = q/2$, мы получим

$$\|f\|_{L_q}^q \gtrsim \int_0^1 \left| \sum_{k=k_0}^{k_1} (P_k f(x))^2 \right|^{q/2} dx \geqslant \int_0^1 \sum_{k=k_0}^{k_1} |P_k f(x)|^q\,dx,$$

или, в терминах коэффициентов,

$$\|f\|_{L_q}^q \gtrsim \sum_{k=k_0}^{k_1} 2^{-k} \sum_{j=1}^{2^k} |c_{k,j}(f)|^q = \|c_{k,j}(f)\|_{q,w}^q. \tag{2.4}$$

Мы также используем простое неравенство

$$\|f\|_{L_2}^2 \geqslant \sum_{k=k_0}^{k_1} \|P_k f\|_{L_2}^2 = \|c_{k,j}(f)\|_{2,w}^2. \tag{2.5}$$

Положим

$$X_{k,j}(t) := c_{k,j}(\chi_t), \quad H_{k,j}(t) := c_{k,j}(\eta_t).$$

Будем смотреть на X, H и как на набор функций переменной $t \in [0, 1]$, и как на однопараметрическое семейство векторов $X(t), H(t) \in \mathcal{F}$. Нетрудно видеть, что

$X_{0,1}(t) = 1-|2t-1|$; остальные функции также кусочно-линейны и выражаются через неё: $X_{k,j}(t) = X_{0,1}(2^k t - (j-1))$ при $t \in [(j-1)2^{-k}, j2^{-k}]$ и $X_{k,j}(t) = 0$ вне этого отрезка.

Вектор $H(t)$ лежит в некотором n-мерном подпространстве $V_n \subset \mathcal{F}$ при всех t, поскольку η_t лежат в n-мерном подпространстве и отображение $f \mapsto (c_{k,j}(f))$ линейно.

Пусть $Z_{k,j}(t)$ — некоторые функции. Воспользуемся неравенством Гёльдера (дважды):

$$\int_0^1 \langle X(t) - H(t), Z(t)\rangle_w \, dt \leqslant \int_0^1 \|X(t) - H(t)\|_{q,w} \cdot \|Z(t)\|_{q',w} \, dt \leqslant$$
$$\leqslant \Big(\int_0^1 \|X(t) - H(t)\|_{q,w}^q \, dt\Big)^{1/q} \Big(\int_0^1 \|Z(t)\|_{q',w}^{q'} \, dt\Big)^{1/q'}. \quad (2.6)$$

Таким образом, используя (2.4), получаем

$$\Big(\int_0^1 \|\chi_t - \eta_t\|_q^q \, dt\Big)^{1/q} \gtrsim \Big(\int_0^1 \|X(t) - H(t)\|_{q,w}^q \, dt\Big)^{1/q} \geqslant \frac{I_1 - I_2}{I_3^{1/q'}}, \quad (2.7)$$

где $I_1 := \int_0^1 \langle X(t), Z(t)\rangle_w \, dt$, $I_2 := \int_0^1 \langle H(t), Z(t)\rangle_w \, dt$, $I_3 := \int_0^1 \|Z(t)\|_{q',w}^{q'} \, dt$.

Конкретизируем теперь наше построение, чтобы оценить величины I_1, I_2, I_3:

$$k_0 := \lfloor \log_2 n \rfloor, \quad k_1 := \lfloor \gamma q \log_2 n \rfloor, \quad (2.8)$$

$$Z_{k,j}(t) := \begin{cases} 2^k a(X_{k,j}(t) - 1/2), & x \in [(j-1)2^{-k}, j2^{-k}], \\ 0, & \text{иначе}, \end{cases}$$

где $a := \sqrt{12}$.

Отметим равенство $\int_0^1 Z_{k,j}(t) \, dt = 0$. Кроме того, $\int_0^1 (t - t_0) Z_{k,j}(t) \, dt = 0$ для любого t_0 — в силу предыдущего равенства, достаточно проверить это для t_0 в середине носителя, то есть $t_0 = (j - 1/2)/2^k$. В этом случае под интегралом произведение функций чётной относительно t_0 и нечётной; поэтому интеграл равен нулю. Докажем, что

$$\int_0^1 Z_{k,j}(t) Z_{k',j'}(t) \, dt = 0, \quad \int_0^1 Z_{k,j}(t)^2 \, dt = 2^k \quad (2.9)$$

при всех $(k,j) \neq (k',j')$. Проверим ортогональность. При $k = k'$ носители функций не пересекаются. Если $k < k'$ и носители пересекаются, то на носителе

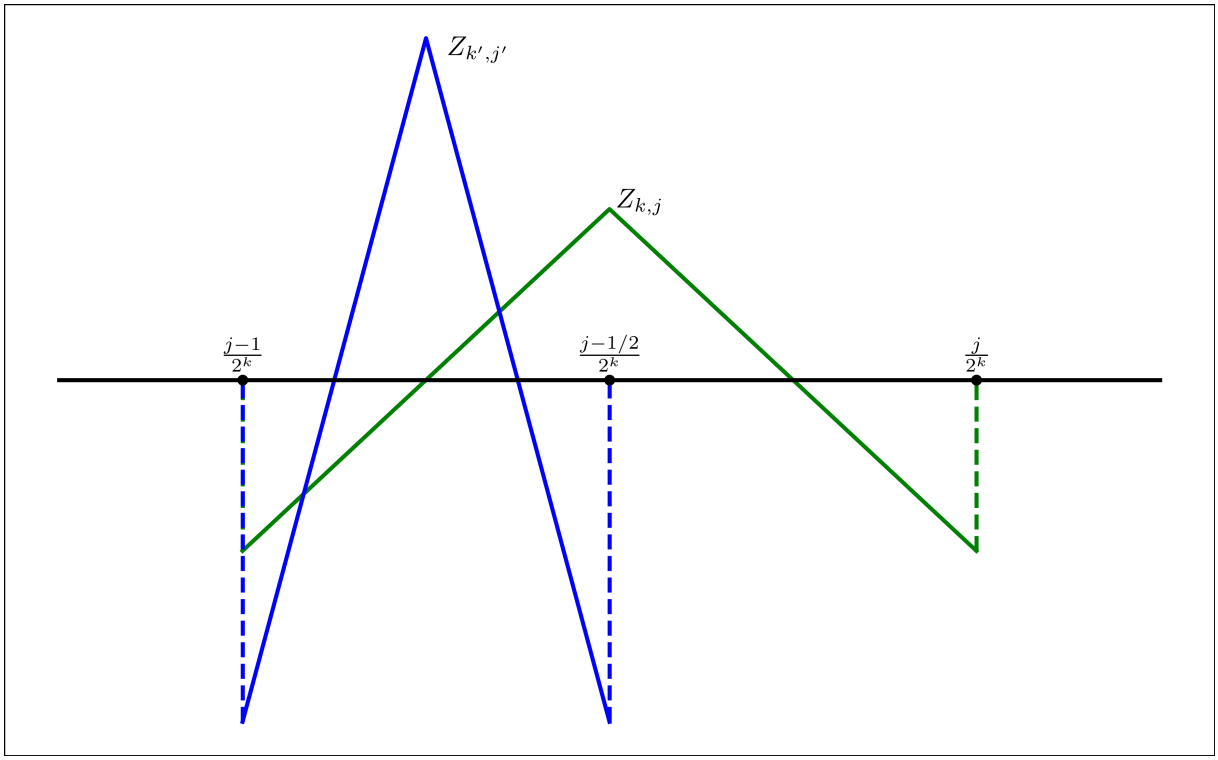


Рис. 2.1: Ортогональность $Z_{k,j}$.

$Z_{k',j'}$ функция $Z_{k,j}$ линейна, и поэтому интеграл от их произведения равен нулю, см. рис. 2.1. Далее,

$$\int_0^1 Z_{k,j}^2(t)\,dt = 2^{2k}a^2 \int_{\frac{j-1}{2^k}}^{\frac{j}{2^k}} (X_{k,j}(t) - 1/2)^2\,dt = 2^k a^2 \int_0^1 (X_{0,1}(t) - 1/2)^2\,dt = 2^k.$$

Перейдём к I_1, I_2, I_3. Сначала вычислим I_1:

$$\int_0^1 X_{k,j}(t)Z_{k,j}(t)\,dt = 2^k a \int_{\frac{j-1}{2^k}}^{\frac{j}{2^k}} (X_{k,j}(t) - 1/2)^2\,dt = a\int_0^1 (X_{0,1}(t) - 1/2)^2\,dt = 1/a,$$

$$I_1 = \int_0^1 \langle X(t), Z(t)\rangle_w\,dt = \sum_{k=k_0}^{k_1} 2^{-k} \sum_{j=1}^{2^k} \int_0^1 X_{k,j}(t)Z_{k,j}(t)\,dt = (k_1 - k_0 + 1)/a.$$

В силу (2.9), Z изотропен в $\mathcal{F}$, то есть $\int_0^1 \langle v, Z(t)\rangle_w^2\,dt = \langle v, v\rangle_w^2$ для любого $v \in \mathcal{F}$. Пусть proj_{V_n} — ортопроектор в $\mathcal{F}$ на подпространство V_n (содержащее векторы $H(t)$). В силу изотропности, имеем (где $v_1, \ldots, v_n$ — это ортонормированный в $\mathcal{F}$ базис V_n):

$$\int_0^1 \|\,\mathrm{proj}_{V_n} Z(t)\|_{2,w}^2\,dt = \int_0^1 \sum_{m=1}^{n} \langle v_m, Z(t)\rangle_w^2\,dt = n. \tag{2.10}$$

Это позволяет оценить I_2:

$$I_2 = \int_0^1 \langle H(t), Z(t)\rangle_w \, dt = \int_0^1 \langle H(t), \operatorname{proj}_{V_n} Z(t)\rangle_w \, dt \leqslant$$
$$\leqslant \Big(\int_0^1 \|H(t)\|_{2,w}^2 \, dt\Big)^{1/2} \Big(\int_0^1 \|\operatorname{proj}_{V_n} Z(t)\|_{2,w}^2 \, dt\Big)^{1/2} = n^{1/2} \Big(\int_0^1 \|H(t)\|_{2,w}^2 \, dt\Big)^{1/2}.$$

Используем неравенство (2.5):

$$\|H(t)\|_{2,w}^2 \leqslant 2\|X(t)\|_{2,w}^2 + 2\|X(t) - H(t)\|_{2,w}^2 \leqslant 2\|X(t)\|_{2,w}^2 + 2\|\chi_t - \eta_t\|_{L_2}^2.$$

Первое слагаемое оценивается при каждом t: на каждом двоичном уровне не более одного числа $X_{k,j}(t)$ отлично от нуля, поэтому $\|X(t)\|_{2,w}^2 \leqslant \sum_{k \geqslant k_0} 2^{-k} \leqslant 2^{1-k_0} < 4/n$. Второе слагаемое можно оценить в среднем, используя отрицание (2.1); итого,

$$\int_0^1 \|H(t)\|_{2,w}^2 \, dt \lesssim n^{-1} + \int_0^1 \|\chi_t - \eta_t\|_{L_2}^2 \, dt \lesssim \delta^2 n^{-1} \log^2 n.$$

Следовательно, $I_2 \lesssim \delta \log n$ при достаточно больших n.

Наконец,

$$I_3 = \int_0^1 \|Z(t)\|_{q',w}^{q'} \, dt = \sum_{k=k_0}^{k_1} 2^{-k} \sum_{j=1}^{2^k} 2^{kq'} a^{q'} \int_{\frac{j-1}{2^k}}^{\frac{j}{2^k}} |X_{k,j}(t) - 1/2|^{q'} \, dt \lesssim$$
$$\lesssim \sum_{k=k_0}^{k_1} 2^{k(q'-1)} \asymp 2^{k_1(q'-1)}.$$

Значит, $I_3^{1/q'} \lesssim 2^{k_1(q'-1)/q'} = 2^{k_1/q} \leqslant n^\gamma$.

Итого, числитель (2.7) не меньше

$$(k_1 - k_0 + 1)/a - C_q \delta \log_2 n \geqslant ((\gamma q - 1)/a - C_q \delta) \log_2 n \gtrsim \log n,$$

если δ достаточно мало. Знаменатель дроби $\lesssim n^\gamma$. Утверждение доказано. $\square$

Отметим, что в доказательстве используется стандартное двоичное разложение (2.3), (2.4) так же, как и у Куланина [50]. В его рассуждениях двоичные уровни рассматривались независимо, и на каждом из них возникало подпространство размерности $\leqslant n$, отсюда происходила "потеря" степени логарифма в оценке снизу. Мы же контролируем все двоичные уровни одновременно, используя специально построенное усреднение и двойственность (2.6), (2.7) в духе работы Глускина [24]; см. также утверждение 1.10.

Следствия для конечномерных множеств. В.Н. Коновалов отмечал, что поперечник $d_n(W_1^1, L_q)$ допускает естественную дискретизацию (делим $[0,1]$ на N равных отрезков и усредняем все функции), а именно,

$$d_n(W_1^1, L_q) \asymp N^{-1/q} d_n(\widetilde{B}_1^N, \ell_q^N), \quad \text{если } N \geqslant n^{q/2}.$$

Здесь $\widetilde{B}_1^N$ — косой октаэдр, являющийся выпуклой оболочкой векторов вида $\pm(1,1,\dots,1,0,0,\dots,0)$, т.е. дискретных аналогов ступенек χ_t. Из теоремы 2.1 следует порядковое равенство

$$d_n(\widetilde{B}_1^N, \ell_q^N) \asymp N^{1/q} n^{-1/2} \log n, \quad \text{при } n \leqslant N^{2/q}.$$

Из утверждения 2.1 следует порядковое равенство для усрёднённого поперечника случайной ступеньки:

$$d_n^{\text{avg}}(\chi_{\mathbf{t}}, L_q)_q \asymp n^{-1/2} \log n, \quad \mathbf{t} \sim \text{Unif}[0,1].$$

Используя дискретизацию, можно вывести следующий результат о матрицах. Пусть $\Delta^{(N)}$ — это верхнетреугольная матрица с $\Delta^{(N)}_{i,j} = 1$ при $i \leqslant j$ и $\Delta^{(N)}_{i,j} = 0$ иначе. Тогда

$$\min_{\operatorname{rank} B \leqslant n} \|\Delta^{(N)} - B\|_{L_q^{N\times N}} \asymp n^{-1/2} \log n, \quad \text{при } n \leqslant N^{2/q}. \tag{2.11}$$

где $\|X\|_{L_q^{N\times N}} := (N^{-2} \sum_{i,j=1}^N |X_{i,j}|^q)^{1/q}$.

2.2 Классы Бесова $B^1_{1,\theta}$

Поперечники произведения октаэдров. В этом пункте мы получим оценку снизу поперечника произведения октаэдров для некоторого класса норм. Результат получится более общий, чем требуется для нахождения порядков поперечников классов Бесова; однако он может представлять самостоятельный интерес.

Рассмотрим пространство $\mathbb{R}^N$ с весовой нормой

$$\|x\|_{\ell^N_{q,w}} := \left(\sum_{i=1}^N w_i |x_i|^q \right)^{1/q}.$$

Пусть координаты разбиты на m блоков: $\{1,\dots,N\} = \sqcup_{j=1}^m \Delta_j$, $|\Delta_j| = N_j$. Рассмотрим в каждом блоке октаэдр $B_1^{N_j}$ и их декартово произведение

$$\prod_{j=1}^m B_1^{N_j} = \{x \in \mathbb{R}^N : \sum_{i\in\Delta_j} |x_i| \leqslant 1,\ j = 1,\dots,m\}.$$

Утверждение 2.2. *Пусть заданы веса $w_1,\dots,w_N \geqslant 0$, разбиение на m блоков $\{1,\dots,N\} = \sqcup_{j=1}^m \Delta_j$, $N_j = |\Delta_j|$, а также число $n \in \mathbb{N}$, для которых выполняются следующие условия:*

1) веса достаточно малы: $\max\limits_{1\leqslant i\leqslant N} w_i \leqslant (4n)^{-1}\sum_{i=1}^N w_i$;

2) весовая доля каждого блока не слишком мала:

$$\gamma_j := \frac{\sum_{i\in\Delta_j} w_i}{\sum_{i=1}^N w_i} \geqslant \frac{C\log 4m}{n}, \quad j = 1,\dots,m.$$

Предположим, для некоторых $q \in [2,\infty)$ и $h > 0$ норма в пространстве $X = (\mathbb{R}^N, \|\cdot\|)$ удовлетворяет неравенству

$$\|x\|_X \geqslant \max\{\|x\|_{\ell_{q,w}^N}, h\|x\|_\infty\}, \quad x \in \mathbb{R}^N.$$

Тогда для поперечника произведения октаэдров в X справедлива оценка

$$d_n(\prod_{j=1}^m B_1^{N_j}, X) \geqslant \min\{c_q\,(n \max_{1\leqslant j\leqslant m}\gamma_j)^{-1/2}(\sum_{i=1}^N w_i)^{1/q}, h/2\}.$$

Заметим, что для выполнения условия 2) необходимо, чтобы блоков было не слишком много: $Cm\log(4m) \leqslant n$.

Под пересечением $X \cap Y$ двух нормированных пространств мы понимаем пространство с нормой $\|x\|_{X\cap Y} := \max(\|x\|_X, \|x\|_Y)$.

Следствие 2.1. *Пусть $2 \leqslant q_1 < \infty$, $1 \leqslant q_2 \leqslant \infty$, веса $w_1,\dots,w_N$ удовлетворяют условиям 1), 2) утверждения 2.2. Тогда справедливо неравенство*

$$d_n(\prod_{j=1}^m B_1^{N_j}, \ell_{q_1,w}^N \cap \ell_{q_2,w}^N) \geqslant \min\left\{c_{q_1}(n\max_{1\leqslant j\leqslant m}\gamma_j)^{-\frac12}(\sum_{i=1}^N w_i)^{\frac{1}{q_1}}, \frac12 \min w_i^{\frac{1}{q_2}}\right\}. \quad (2.12)$$

Действительно, достаточно заметить, что $\|x\|_{\ell_{q_2,w}^N} \geqslant \min_i w_i^{1/q_2}\|x\|_\infty$. Это следствие и будет использоваться в дальнейшем (причём для $q_1 = 2$ и q_2 равному q из теоремы 2.2).

Доказательство утверждения 2.2 представляет собой развитие метода [40]. По поводу произведения октаэдров см. также §5.1.

Доказательство. Без ограничения общности

$$\|x\|_X = \max\{\|x\|_{\ell^N_{q,w}}, h\|x\|_\infty\}.$$

Обозначим для краткости $B := \prod_{j=1}^m B_1^{N_j}$ $d_n := d_n(B, X)$, и пусть Q_n — экстремальное подпространство для этого поперечника. Если $d_n \geqslant h/2$, то утверждение доказано; поэтому мы будем считать, что $d_n < h/2$.

Нас интересует приближение крайних точек B; легко видеть, что $\operatorname{extr} B = \{\sum_{j=1}^m \pm e_{i_j} \colon i_j \in \Delta_j\}$. Покажем, что для каждого $u \in \operatorname{extr} B$ найдётся $v = v(u) \in Q_n$ такой, что

$$\|u - v\|_X \leqslant 2d_n, \tag{2.13}$$

$$u_i = v_i \quad \forall\, i \in \operatorname{supp} u. \tag{2.14}$$

Для фиксированного $u \in \operatorname{extr} B$ возьмём $v_0 \in Q_n$ такой, что $\|u - v_0\|_X \leqslant d_n$. Запишем погрешность $u - v_0$ в виде $u_1 + \beta_1$, где $\operatorname{supp} u_1 \subset \operatorname{supp} u$, $\operatorname{supp} \beta_1 \cap \operatorname{supp} u = \varnothing$. Имеем $\|\beta_1\|_X \leqslant \|u - v_0\|_X \leqslant d_n$,

$$\|u_1\|_\infty \leqslant h^{-1}\|u_1\|_X \leqslant h^{-1}\|u - v_0\|_X \leqslant d_n/h \leqslant 1/2.$$

Тем самым в каждом блоке у u_1 не более одной ненулевой координаты, причём по модулю не больше $1/2$; отсюда $u_1 \in (1/2)B$. Значит, мы можем приблизить u_1 вектором $v_1 \in Q_n$ с погрешностью $\leqslant d_n/2$. Далее пишем $u_1 - v_1 = u_2 + \beta_2$, $\operatorname{supp} u_2 \subset \operatorname{supp} u$, и переходим к приближению $u_2 \in (1/4)B$. Повторяя этот процесс, мы получим приближение вектора u суммой $v_0 + v_1 + \ldots \in Q_n$ со свойством (2.14) и с погрешностью $\|\beta_1 + \beta_2 + \ldots\|_X \leqslant 2d_n$.

Итак, зафиксируем (какое-либо) отображение $v(\cdot)\colon \operatorname{extr} B \to Q_n$, удовлетворяющее условиям (2.13) и (2.14).

Нормируем веса условием $\sum_{i=1}^N w_i = 1$ и положим $p_i := 2nw_i$. Рассмотрим случайное множество $\mathbf{\Omega} \subset \{1, \ldots, N\}$ такое, что каждая точка i попадает в $\mathbf{\Omega}$ независимо от других с вероятностью p_i. Условие 1) утверждения даёт $p_i \in [0, 1]$ (даже $p_i \leqslant 1/2$, но это потребуется позже). Отметим, что выбор вероятностей пропорционально весам позволяет записать весовую норму в виде математического ожидания:

$$\mathsf{E} \sum_{i \in \mathbf{\Omega}} |x_i|^q = 2n\|x\|^q_{\ell^N_{q,w}}.$$

Коэффициент пропорциональности обеспечивает равенство $\mathsf{E}|\mathbf{\Omega}| = \sum_{i=1}^N p_i = 2n$; кроме того, $\mathsf{E}|\mathbf{\Omega} \cap \Delta_j| = 2n\gamma_j$.

Пусть $\mathbf{u}$ — случайная точка, выбранная равномерно из $\operatorname{extr} B$ и независимо от $\mathbf{\Omega}$. Другими словами, в каждом блоке Δ_j одна случайная координата $\mathbf{u}$ равна ± 1, остальные равны нулю. Рассмотрим случайную величину

$$\xi := \sum_{i \in \mathbf{\Omega}} |\mathbf{u}_i - v(\mathbf{u})_i|^q \cdot \mathbf{1}\{\operatorname{supp} \mathbf{u} \subset \mathbf{\Omega}\}.$$

Забегая вперёд, поясним смысл условия $\{\operatorname{supp} \mathbf{u} \subset \mathbf{\Omega}\}$. При условии наступления этого события мы переходим к ситуации малой размерности: $|\mathbf{\Omega}| \asymp n$, в которой работает L_2-оценка.

Оценим $\mathsf{E}\xi$ двумя способами. Фиксируем $u \in \operatorname{extr} B$ и усредняем по Ω:

$$\mathsf{E}(\xi | \mathbf{u} = u) = \sum_{i=1}^N |u_i - v(u)_i|^q \cdot \mathsf{P}(\operatorname{supp} u \subset \mathbf{\Omega},\ i \in \mathbf{\Omega}).$$

Если $i \in \operatorname{supp} u$, то, в силу (2.14), соответствующее слагаемое равно нулю. Иначе, в силу независимости,

$$\frac{\mathsf{P}(\operatorname{supp} u \subset \mathbf{\Omega},\ i \in \mathbf{\Omega})}{\mathsf{P}(\operatorname{supp} u \subset \mathbf{\Omega})} = \mathsf{P}(i \in \mathbf{\Omega}) = 2nw_i.$$

Суммируя по i и используя (2.13), получаем

$$\frac{\mathsf{E}(\xi | \mathbf{u} = u)}{\mathsf{P}(\operatorname{supp} u \subset \mathbf{\Omega})} = 2n \sum_{i=1}^N w_i |u_i - v(u)_i|^q = 2n \|u - v(u)\|_{\ell_{q,w}^N}^q \lesssim n d_n^q.$$

Домножим на вероятность и усредним по u, получим

$$\mathsf{E}\xi \lesssim n d_n^q \cdot \mathsf{P}(\operatorname{supp} \mathbf{u} \subset \mathbf{\Omega}).$$

Теперь фиксируем Ω и усредняем по u:

$$\mathsf{E}(\xi | \mathbf{\Omega} = \Omega) = \mathsf{E} \sum_{i \in \Omega} |\mathbf{u}_i - v(\mathbf{u})_i|^q \, \mathbf{1}\{\operatorname{supp} \mathbf{u} \subset \Omega\} =$$

$$= \mathsf{E}\left(\sum_{i \in \Omega} |\mathbf{u}_i - v(\mathbf{u})_i|^q \,\middle|\, \operatorname{supp} \mathbf{u} \subset \Omega \right) \cdot \mathsf{P}(\operatorname{supp} \mathbf{u} \subset \Omega).$$

Через $\widetilde{\mathbf{u}}$ обозначим случайный вектор с условным распределением

$$\mathsf{Law}(\widetilde{\mathbf{u}}) = \mathsf{Law}(\mathbf{u} | \operatorname{supp} \mathbf{u} \subset \Omega),$$

то есть вектор $\widetilde{\mathbf{u}}$ устроен следующим образом: в каждом множестве $\Omega_j := \Omega \cap \Delta_j$ одна случайная координата равна ± 1, а остальные равны нулю. Тогда

$$\mathsf{E}\left(\sum_{i\in\Omega} |\mathbf{u}_i - v(\mathbf{u})_i|^q \,\middle|\, \operatorname{supp}\mathbf{u} \subset \Omega\right) = \mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}})\|^q_{\ell_q(\Omega)}.$$

Скажем, что множество Ω *регулярно*, если мощность множеств $\Omega_j := \Omega \cap \Delta_j$ удовлетворяет неравенствам $\frac{3}{2}n\gamma_j \leqslant |\Omega_j| \leqslant An\gamma_j$ при всех j; константу A выберем достаточно большой для дальнейших рассуждений (см. лемму 2.1).

Предположим, что Ω регулярно. Ясно, что $\frac{3}{2}n \leqslant |\Omega| \leqslant An$.

Перейдём от q-среднего к среднеквадратическому и от метрики $\ell_q(\Omega)$ к метрике $\ell_2(\Omega)$ (здесь мы пользуемся тем, что $q \geqslant 2$):

$$(\mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}})\|^q_{\ell_q(\Omega)})^{1/q} \geqslant (\mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}})\|^2_{\ell_q(\Omega)})^{1/2} \geqslant |\Omega|^{1/q-1/2}(\mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}})\|^2_{\ell_2(\Omega)})^{1/2} \gtrsim$$
$$\gtrsim n^{1/q-1/2}(\mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}}))\|^2_{\ell_2(\Omega)})^{1/2} \geqslant n^{1/q-1/2} d_n^{\mathrm{avg}}(\widetilde{\mathbf{u}}, \ell_2(\Omega))_2.$$

Оценим поперечник, пользуясь утверждением 1.3. Имеем в стандартном базисе: $\mathsf{E}\widetilde{\mathbf{u}}_i\widetilde{\mathbf{u}}_j = 0$ и $\mathsf{E}\widetilde{\mathbf{u}}_i^2 = |\Omega_j|^{-1} \asymp (n\gamma_j)^{-1}$, $i \in \Omega_j$. Обозначим $\gamma := \max \gamma_j$, отсюда

$$d_n^{\mathrm{avg}}(\widetilde{\mathbf{u}}, \ell_2(\Omega))_2^2 \gtrsim \sum_{k=n+1}^{3n/2} (n\gamma)^{-1} \gtrsim \gamma^{-1}.$$

В итоге, для регулярного Ω имеем

$$\mathsf{E}(\xi|\boldsymbol{\Omega} = \Omega) \geqslant \mathsf{P}(\operatorname{supp}\mathbf{u} \subset \Omega) \cdot \mathsf{E}\|\widetilde{\mathbf{u}} - v(\widetilde{\mathbf{u}})\|^q_{\ell_q(\Omega)} \gtrsim \mathsf{P}(\operatorname{supp}\mathbf{u} \subset \Omega) \cdot n^{1-q/2} \cdot \gamma^{-q/2}.$$

Усредняя теперь по Ω, получим

$$\mathsf{E}\xi \gtrsim n^{1-q/2}\gamma^{-q/2} \cdot \mathsf{P}(\operatorname{supp}\mathbf{u} \subset \boldsymbol{\Omega},\ \boldsymbol{\Omega} \text{ регулярно}).$$

Чуть позже мы покажем, что можно разделить условия, входящие в вероятность:

$$\mathsf{P}(\operatorname{supp}\mathbf{u} \subset \boldsymbol{\Omega},\ \boldsymbol{\Omega} \text{ регулярно}) \geqslant \mathsf{P}(\operatorname{supp}\mathbf{u} \subset \boldsymbol{\Omega})\mathsf{P}(\boldsymbol{\Omega} \text{ регулярно}). \qquad (2.15)$$

Покажем, что $\boldsymbol{\Omega}$ регулярно с вероятностью $\geqslant 1/2$. Обозначим $\boldsymbol{\Omega}_j := \boldsymbol{\Omega}_j \cap \Delta_j$. Применим следствие из неравенства Бернштейна, (1.26):

$$\mathsf{P}(||\boldsymbol{\Omega}_j| - 2n\gamma_j| > n\gamma_j/2) \leqslant 2\exp(-cn\gamma_j).$$

Из условия 2) утверждения следует, что эта вероятность не больше $\frac{1}{2m}$. Следовательно, вероятность нерегулярности $\mathbf{\Omega}$ не превосходит $1/2$. В итоге:

$$\mathsf{E}\xi \gtrsim n^{1-q/2}\gamma^{-q/2}\mathsf{P}(\operatorname{supp}\mathbf{u} \subset \mathbf{\Omega}).$$

Сравнивая нижнюю и верхнюю оценки $\mathsf{E}\xi$, получаем требуемое неравенство на d_n.

Остаётся обосновать (2.15). Для этого понадобится одна техническая лемма.

Лемма 2.1. *Пусть $\delta \in (0,1)$, $\mathbf{\Omega}$ — случайное подмножество $\{1,\dots,M\}$, каждая точка i попадает в $\mathbf{\Omega}$ независимо от остальных с вероятностью $p_i \leqslant 1/2$. Тогда, если число A и величина $\omega := \mathsf{E}\#\mathbf{\Omega} = \sum_{i=1}^M p_i$ достаточно велики: $A \geqslant C_1$, $\omega \geqslant C_2(\delta)$, то*

$$\mathsf{E}(\#\mathbf{\Omega}|\mathcal{R}) \geqslant \mathsf{E}\#\mathbf{\Omega},$$

где событие $\mathcal{R}$ ("регулярность") определяется следующим образом:

$$\mathcal{R} := \{(1-\delta)\omega \leqslant \#\mathbf{\Omega} \leqslant A\omega\}.$$

Выведем из леммы неравенство (2.15). В силу независимости, достаточно доказывать неравенство для каждого блока по отдельности:

$$\mathsf{P}(\operatorname{supp}\mathbf{u} \cap \Delta_j \subset \mathbf{\Omega}_j \big| 3n\gamma_j/2 \leqslant \#\mathbf{\Omega}_j \leqslant An\gamma_j) \geqslant \mathsf{P}(\operatorname{supp}\mathbf{u} \cap \Delta_j \subset \mathbf{\Omega}_j).$$

Величина справа равна $N_j^{-1}\mathsf{E}|\mathbf{\Omega}_j|$, а величина слева — соответствующему условному математическому ожиданию. Остаётся применить лемму для $\delta = 1/4$ и $\mathbf{\Omega}\cap\Delta_j$. (Отметим, что $\omega = 2n\gamma_j \geqslant 2C\log 4m$ достаточно велико за счёт выбора C.)

Докажем лемму. Вероятностное пространство распадается на три непересекающихся события: $\mathcal{R}$, $\mathcal{C}_- := \{\#\mathbf{\Omega} < (1-\delta)\omega\}$ и $\mathcal{C}_+ := \{\#\mathbf{\Omega} > A\omega\}$. Положим также $\mathcal{C} := \mathcal{C}_- \sqcup \mathcal{C}_+$. Доказываемое неравенство эквивалентно тому, что $\mathsf{E}(\#\mathbf{\Omega}|\mathcal{R}) \geqslant \mathsf{E}(\#\mathbf{\Omega}|\mathcal{C})$. Используя неравенство (1.26), получим

$$\mathsf{P}(\#\mathbf{\Omega} \in [(1-\delta/3)\omega, (1+\delta/3)\omega]) \geqslant 1 - 2\exp(-c(\delta)\omega) > 9/10$$

при больших ω. Тем более это верно при условии события $\mathcal{R}$. Отсюда

$$\mathsf{E}(\#\mathbf{\Omega}|\mathcal{R}) \geqslant \frac{9}{10}(1-\delta/3)\omega + \frac{1}{10}(1-\delta)\omega > (1-\delta/2)\omega.$$

Оценим $\mathsf{E}(\#\mathbf{\Omega}|\mathcal{C}) = \mathsf{E}(\#\mathbf{\Omega}\mathbf{1}_{\mathcal{C}_-})/\mathsf{P}(\mathcal{C}) + \mathsf{E}(\#\mathbf{\Omega}\mathbf{1}_{\mathcal{C}_+})/\mathsf{P}(\mathcal{C})$.

Первое слагаемое не превосходит $(1-\delta)\omega\mathsf{P}(\mathcal{C}_-)/\mathsf{P}(\mathcal{C}) \leqslant (1-\delta)\omega$.

Второе слагаемое равно

$$\frac{\mathsf{E}\#\mathbf{\Omega}\mathbf{1}_{\mathcal{C}_+}}{\mathsf{P}(\mathcal{C})} = \frac{1}{\mathsf{P}(\mathcal{C})}\sum_{k>A\omega} k\mathsf{P}(\#\mathbf{\Omega} = k). \tag{2.16}$$

Заметим, что $\mathsf{P}(\mathcal{C}) \geqslant \mathsf{P}(\mathbf{\Omega} = \varnothing) = \prod(1-p_i) \geqslant \prod 4^{-p_i} = 4^{-\omega}$ (мы воспользовались тем, что $p_i \leqslant 1/2$). При $k \geqslant 2\omega$, так же в силу неравенства (1.26), имеем

$$\mathsf{P}(\#\mathbf{\Omega} = k) \leqslant \mathsf{P}(|\#\mathbf{\Omega} - \omega| \geqslant k/2) \leqslant 2\exp(-ck).$$

Итого, величина (2.16) не превосходит

$$4^{\omega}\sum_{k>A\omega} k\mathsf{P}(\#\mathbf{\Omega} = k) \leqslant 4^{\omega}\cdot 2\sum_{k>A\omega} k\exp(-ck) \leqslant \frac{\delta}{2},$$

если A и ω достаточно велики. Суммируя, получаем $\mathsf{E}(\#\mathbf{\Omega}|\mathcal{C}) \leqslant (1-\delta)\omega + (\delta/2) < (1-\delta/2)\omega < \mathsf{E}(\#\mathbf{\Omega}|\mathcal{R})$. Лемма, а вместе с ней и утверждение 2.2, доказаны. □

Дискретизация для классов Бесова. Хорошо известно, что пространства Бесова изоморфны пространствам последовательностей (Лизоркин, Трибель, Мейер) — норма функции эквивалентна норме последовательности коэффициентов по подходящему базису. Это позволит нам дискретизовать задачу о поперечниках этих классов.

Наиболее разработана теория пространств Бесова на прямой и в $\mathbb{R}^d$; часто пространства Бесова на отрезке определяют как ограничение на этот отрезок функций из пространства $\mathcal{B}^r_{p,\theta}(\mathbb{R})$, см. [144, Definition 1.95]. Этот подход эквивалентен определению через модуль гладкости, данному в §1.1.

Теорема 2.A ([144], §1.11, Theorem 1.118)**.** *Норма пространства $\mathcal{B}^r_{p,\theta}[a,b]$ эквивалентна норме*

$$\|g\| := \inf\{\|f\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})}\colon\ f|_{[a,b]} \equiv g\}.$$

Нам также потребуется одно техническое утверждение.

Лемма 2.A ([143], Theorem 4.2.2, [8], лемма 3)**.** *Пусть η — бесконечно-гладкая функция на прямой с компактным носителем. Тогда справедливо неравенство*

$$\|\eta f\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})} \leqslant C(r,p,\theta,\eta)\|f\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})}, \quad f \in \mathcal{B}^r_{p,\theta}(\mathbb{R}).$$

Для дискретизации удобнее всего было бы использовать функции Хаара, как в §2.1, однако они не образуют безусловного базиса в рассматриваемых нами пространствах (см. [102]). Поэтому будем использовать базисы (регулярных) всплесков.

Теория всплесков изложена в ряде монографий и книг, см., например, [153, 60]. Напомним, что функция $\psi \in L_2(\mathbb{R})$ называется всплеском, если система $\{\psi_{k,j}(x) := 2^{k/2}\psi(2^k x - j)\}_{k,j\in\mathbb{Z}}$ образует ортонормированный базис в $L_2(\mathbb{R})$. Стандартным способом построения всплесков является кратномасштабный анализ (КМА), где помимо всплеск-функции возникает масштабирующая функция; подробнее см. [153, Ch.2].

Пусть ψ, φ — всплеск-функция и масштабирующая функция на прямой, соответственно. Полагаем $\psi_{k,j}(x) := 2^{k/2}\psi(2^k x - j)$, $\varphi_{0,j}(x) := \varphi(x-j)$, $k, j \in \mathbb{Z}$. Пространства $W_k := \overline{\operatorname{span}}\{\psi_{k,j} \colon j \in \mathbb{Z}\}$ и $V_0 := \overline{\operatorname{span}}\{\varphi_{0,j} \colon j \in \mathbb{Z}\}$ в $L_2(\mathbb{R})$ обладают свойствами:

$$V_0 = \bigoplus_{k<0} W_k, \quad V_0^\perp = \bigoplus_{k\geqslant 0} W_k. \tag{2.17}$$

Мы будем использовать регулярные всплески Добеши с компактным носителем (см. [153], §3.1, §4.2), которые обладают следующими свойствами:

$$\psi, \varphi \in C^{r_0}(\mathbb{R}), \quad \operatorname{supp}\psi, \varphi \subset [-M, M], \tag{2.18}$$

$$\int_{\mathbb{R}} \psi(x) x^j \, dx = 0, \quad j = 0, 1, \ldots, r_0 - 1, \tag{2.19}$$

где $r_0 \in \mathbb{N}$ может быть выбран любым, M зависит от r_0.

При наличии определённой регулярности всплеска можно доказать его безусловность в L_p (нам потребуется только слабая версия такого результата).

Теорема 2.B ([153], §8.2)**.** *Если ψ — гладкая всплеск-функция с компактным носителем, то $\{\psi_{k,j} \colon k, j \in \mathbb{Z}\}$ является безусловным базисом в $L_p(\mathbb{R})$ при $1 < p < \infty$.*

Рассмотрим пространство $\ell^s_{p,\theta}$ последовательностей $(x_{k,j})_{k\geqslant 0, j\in\mathbb{Z}}$ с нормой

$$\|(x_{k,j})_{k\geqslant 0,j\in\mathbb{Z}}\|_{\ell^s_{p,\theta}} = \begin{cases} \left(\sum_{k\geqslant 0} 2^{k(s-1/p)\theta} \|(x_{k,j})_{j\in\mathbb{Z}}\|^\theta_{\ell_p}\right)^{1/\theta}, & \theta < \infty, \\ \sup_{k\geqslant 0} 2^{k(s-1/p)} \|(x_{k,j})_{j\in\mathbb{Z}}\|_{\ell_p}, & \theta = \infty. \end{cases}$$

В этом подпараграфе мы используем обозначение $\langle f, g\rangle := \int_{\mathbb{R}} f(x)g(x)\,dx$.

Теорема 2.C (См. [142], §1.2, Theorem 1.20)**.** *Пусть $r_0 \in \mathbb{N}$, система всплесков удовлетворяет условиям* (2.18), (2.19), *$0 < r < r_0$; $p, \theta \in [1, \infty]$. Функция $f \in L_p(\mathbb{R})$ принадлежит $\mathcal{B}^r_{p,\theta}(\mathbb{R})$ тогда и только тогда, когда конечна норма*

$$\|f\| := \|(\langle f, \varphi_{0,j}\rangle)_{j\in\mathbb{Z}}\|_{\ell_p} + \|(2^{k/2}\langle f, \psi_{k,j}\rangle)_{k\geqslant 0, j\in\mathbb{Z}}\|_{\ell^r_{p,\theta}}. \tag{2.20}$$

Эта норма эквивалентна норме пространства $\mathcal{B}^r_{p,\theta}(\mathbb{R})$.

(См. также [107], Th.1.13, где используются более компактные обозначения.)

Помимо класса, нужно дискретизовать и норму.

Лемма 2.2. *Пусть $f \in L_q(\mathbb{R}) \cap V_0^\perp$. Тогда при $2 \leqslant q < \infty$ имеем*

$$\|(2^{k/2}\langle f, \psi_{k,j}\rangle)_{k\geqslant 0, j\in\mathbb{Z}}\|_{\ell^0_{q,q}} \lesssim \|f\|_q \lesssim \|(2^{k/2}\langle f, \psi_{k,j}\rangle)_{k\geqslant 0, j\in\mathbb{Z}}\|_{\ell^0_{q,2}}. \tag{2.21}$$

(Константы в "$\lesssim$" зависят от q и от системы всплесков.)

Для полноты изложения мы докажем лемму; такие оценки, вероятно, известны. В контексте поперечников аналогичные неравенства применялись в [50] и нами в §2.1 для всплесков Хаара, см. (2.4).

Доказательство. Мы будем пользоваться следующими фактами.

(i) Поскольку система всплесков является безусловным базисом в $L_q(\mathbb{R})$, то норма функции эквивалентна норме функции Пэли (см., например, [153, §7.3, Corollary 7.11]), откуда

$$\|f\|_q \asymp \|(\sum_{k,j}\langle f, \psi_{k,j}\rangle^2 |\psi_{k,j}|^2)^{1/2}\|_q.$$

(ii) При фиксированном k, в силу [153, §8.1, Proposition 8.3]:

$$\|\sum_{j\in\mathbb{Z}}\langle f, \psi_{k,j}\rangle\psi_{k,j}\|_q \asymp 2^{k(1/2-1/q)}\|(\langle f, \psi_{k,j}\rangle)_{j\in\mathbb{Z}}\|_{\ell_q}.$$

В силу (2.17) и того, что $f \in V_0^\perp$, коэффициенты $\langle f, \psi_{k,j}\rangle = 0$ при $k < 0$.

Воспользуемся нижней p-оценкой (1.1) для $p = q/2$, а также соотношением $\|\psi_{k,j}\|_q \asymp 2^{k(1/2-1/q)}$:

$$\|f\|_q^q \asymp \int |\sum_{k\geqslant 0, j\in\mathbb{Z}}\langle f, \psi_{k,j}\rangle^2 \psi_{k,j}(t)^2|^{q/2}\,dt \geqslant \sum_{k\geqslant 0, j\in\mathbb{Z}}\|\langle f, \psi_{k,j}\rangle\psi_{k,j}\|_q^q \asymp$$
$$\asymp \sum_{k\geqslant 0} 2^{kq(1/2-1/q)}\|(\langle f, \psi_{k,j}\rangle)_{j\in\mathbb{Z}}\|^q_{\ell_q} = \sum_{k\geqslant 0} 2^{kq(-1/q)}\|(2^{k/2}\langle f, \psi_{k,j}\rangle)_{j\in\mathbb{Z}}\|^q_{\ell_q},$$

что даёт нужную оценку нормы снизу.

Для оценки сверху применим (i), положим $g_k := (\sum_{j\in\mathbb{Z}}\langle f,\psi_{k,j}\rangle^2\psi_{k,j}^2)^{1/2}$ и получим:

$$\|f\|_q^2 \asymp \|(\sum_{k\geqslant 0} g_k^2)^{1/2}\|_q^2 = \|\sum_{k\geqslant 0} g_k^2\|_{q/2} \leqslant \sum_{k\geqslant 0}\|g_k^2\|_{q/2} = \sum_{k\geqslant 0}\|g_k\|_q^2.$$

Далее, снова используя (i), а потом (ii),

$$\|g_k\|_q = \|(\sum_{j\in\mathbb{Z}}\langle f,\psi_{k,j}\rangle^2\psi_{k,j}^2)^{1/2}\|_q \asymp \|\sum_{j\in\mathbb{Z}}\langle f,\psi_{k,j}\rangle\psi_{k,j}\|_q \asymp$$
$$\asymp 2^{k(1/2-1/q)}\|(\langle f,\psi_{k,j}\rangle)_{j\in\mathbb{Z}}\|_{\ell_q} \asymp 2^{k(-1/q)}\|(2^{k/2}\langle f,\psi_{k,j}\rangle)_{j\in\mathbb{Z}}\|_{\ell_q}.$$

Возводя в квадрат и суммируя по k, получаем нужную оценку. Лемма доказана. □

Обозначим $T := \{(k,j)\colon k \geqslant 0,\ j = 1,\dots,2^k\}$ и пусть $\ell^s_{p,\theta}(T)$ — аналог пространства $\ell^s_{p,\theta}$ для векторов $x \in \mathbb{R}^T$:

$$\|x\|_{\ell^s_{p,\theta}(T)} := \begin{cases} \left(\sum_{k\geqslant 0} 2^{k(s-1/p)\theta}\|(x_{k,j})_{1\leqslant j\leqslant 2^k}\|_{\ell_p}^\theta\right)^{1/\theta}, & \theta<\infty,\\ \sup_{k\geqslant 0} 2^{k(s-1/p)}\|(x_{k,j})_{1\leqslant j\leqslant 2^k}\|_{\ell_p}, & \theta=\infty, \end{cases}$$

а $b^s_{p,\theta}(T)$ — единичный шар этого пространства.

Лемма 2.3. *Пусть $r>0$, $1\leqslant p,\theta\leqslant\infty$, $2\leqslant q<\infty$. Существует $n_0 = n_0(r)\in\mathbb{N}$, такой что для всех $n\in\mathbb{N}$ справедливы неравенства*

$$d_n(B^r_{p,\theta}[0,1], L_q[0,1]) \geqslant c(r,p,\theta,q)d_n(b^r_{p,\theta}(T), \ell^0_{q,q}(T)\cap\ell^0_{2,2}(T)). \tag{2.22}$$

$$d_{n+n_0}(B^r_{p,\theta}[0,1], L_q[0,1]) \leqslant C(r,p,\theta,q)d_n(b^r_{p,\theta}(T), \ell^0_{q,2}(T)). \tag{2.23}$$

Доказательство. Возьмём $r_0 := \lfloor r\rfloor + 1$ и зафиксируем систему всплесков, удовлетворяющую условиям (2.18), (2.19).

Для функции f мы обозначаем

$$x_{k,j} := 2^{k/2}\langle f,\psi_{k,j}\rangle.$$

Рассмотрим подпространство

$$L_q\langle\Psi_T\rangle := \overline{\operatorname{span}}\{\psi_{k,j}\colon (k,j)\in T\}$$

(замыкание в $L_q(\mathbb{R})$). Функции $\psi_{k,j}$, $(k,j) \in T$, имеют носитель, лежащий в отрезке внутри некоторого интервала (a_1, b_1); то же верно и для $L_q\langle\Psi_T\rangle$. Таким образом,

$$\mathcal{B}^r_{p,\theta}(\mathbb{R}) \cap L_q\langle\Psi_T\rangle \subset \mathcal{B}^r_{p,\theta}[a_1, b_1].$$

В силу базисности, естественный проектор $L_q(\mathbb{R}) \to L_q\langle\Psi_T\rangle$ ограничен, откуда

$$d_n(B^r_{p,\theta}[a_1, b_1], L_q[a_1, b_1]) \gtrsim d_n(B^r_{p,\theta}(\mathbb{R}) \cap L_q\langle\Psi_T\rangle, L_q\langle\Psi_T\rangle). \tag{2.24}$$

Ясно, что линейная замена переменной задаёт изоморфизм пространств Бесова и не влияет на порядок поперечников:

$$d_n(B^r_{p,\theta}[0,1], L_q[0,1]) \asymp d_n(B^r_{p,\theta}[a,b], L_q[a,b]). \tag{2.25}$$

Следовательно, оценка (2.24) справедлива для класса Бесова на $[0,1]$.

Из теоремы 2.C следует, что для $f \in L_q\langle\Psi_T\rangle$ имеем

$$\|f\|_{\mathcal{B}^r_{p,\theta}} \asymp \|x\|_{\ell^r_{p,\theta}(T)}. \tag{2.26}$$

Действительно, в правой части (2.26) записано второе слагаемое из теоремы. Первое слагаемое равно нулю, так как $\psi_{k,j} \perp \varphi_{0,i}$ при $k \geqslant 0$ (см. (2.17)).

В силу леммы 2.2, имеем $\|f\|_q \gtrsim \|x\|_{\ell^0_{q,q}(T)}$. Кроме того, поскольку $q \geqslant 2$, то $\|f\|_q \geqslant \|f\|_2 \geqslant \|x\|_{\ell^0_{2,2}(T)}$. С учётом (2.24) и (2.26), получаем требуемую оценку снизу (2.22).

Оценка сверху получается, в целом, аналогично. Возьмём функцию $f \in B^r_{p,\theta}[0,1]$. Продолжим её с отрезка на прямую, не сильно увеличив норму и так, чтобы носитель лежал в некотором большем отрезке:

$$\|f_1\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})} \lesssim \|f\|_{\mathcal{B}^r_{p,\theta}[0,1]}, \quad \operatorname{supp} f_1 \subset [a_2, b_2], \quad f_1|_{[0,1]} \equiv f.$$

Это всегда можно сделать — сначала продолжим на $\mathbb{R}$ (см. теорему 2.A), потом домножим на гладкую срезку η, см. лемму 2.A. Ясно, что достаточно приблизить в $L_q(\mathbb{R})$ функцию f_1. Линейной заменой — если изначально рассматривать вместо $[0,1]$ отрезок вида $I := [1/2 - \varepsilon, 1/2 + \varepsilon]$ с малым ε — можно свести дело к случаю $[a_2, b_2] \subset [1/3, 2/3]$:

$$\|f_1\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})} \lesssim \|f\|_{\mathcal{B}^r_{p,\theta}(I)}, \quad \operatorname{supp} f_1 \subset [1/3, 2/3], \quad f_1|_I \equiv f.$$

Покажем, что есть лишь конечное число функций $\varphi_{0,j}$, $j \in \mathbb{Z}$, и $\psi_{k,j}$, $k \geqslant 0$, $j \in \mathbb{Z}$, $(k,j) \notin T$, у которых носитель пересекается с $[1/3, 2/3]$. Для $\varphi_{0,j}$ это

понятно; проверим для $\psi_{k,j}$. Имеем $\operatorname{supp}\psi_{k,j} \subset [(-M+j)/2^k, (M+j)/2^k]$ (см. (2.18)), этот отрезок при больших k и $j \notin \{1,\dots,2^k\}$ не пересекается с $[1/3, 2/3]$. Вычтем из f_1 линейную комбинацию всех "исключительных" функций, это элемент фиксированного пространства конечной размерности n_0; получим функцию f_2. При этом, в силу теоремы 2.C, норма в пространстве Бесова (точнее, ей эквивалентная) не вырастет. Кроме того, $f_2 \in V_0^\perp$, поэтому $f_2 \in L_q\langle\Psi_T\rangle$. В итоге, нам достаточно аппроксимировать функции f_2, для которых

$$f_2 \in L_q\langle\Psi_T\rangle, \quad \|f_2\|_{\mathcal{B}^r_{p,\theta}(\mathbb{R})} \lesssim \|f\|_{\mathcal{B}^r_{p,\theta}(I)} \leqslant 1.$$

Отсюда

$$d_{n+n_0}(B^r_{p,\theta}(I), L_q(I)) \lesssim d_n(B^r_{p,\theta}(\mathbb{R}) \cap L_q\langle\Psi_T\rangle, L_q(\mathbb{R})) \leqslant d_n(B^r_{p,\theta}(\mathbb{R}) \cap L_q\langle\Psi_T\rangle, L_q\langle\Psi_T\rangle).$$

Применяя (2.21) и (2.26), получаем требуемую оценку (2.23); в силу (2.25), отрезок I можно заменить обратно на $[0,1]$. □

Оценки поперечников.

Теорема 2.2. *Для любых $q \in (2,\infty)$, $1 \leqslant \theta \leqslant \infty$ и $n \in \mathbb{N}$ имеют место неравенства*

$$c(q,\theta) n^{-1/2} \log^{1/2} n \leqslant d_n(B^1_{1,\theta}[0,1], L_q[0,1]) \leqslant C(q,\theta) n^{-1/2} \log^{1/2} n, \quad 1 \leqslant \theta \leqslant 2,$$

$$c(q,\theta) n^{-1/2} \log^{1-1/\theta} n \leqslant d_n(B^1_{1,\theta}[0,1], L_q[0,1]) \leqslant C(q,\theta) n^{-1/2} \log^{1-1/\theta} n, \quad 2 \leqslant \theta \leqslant \infty.$$

Доказательство. Лемма 2.3 при $r = p = 1$ даёт следующие оценки:

$$d_n(B^1_{1,\theta}[0,1], L_q[0,1]) \gtrsim d_n(b^1_{1,\theta}(T), \ell^0_{q,q}(T) \cap \ell^0_{2,2}(T)), \tag{2.27}$$

$$d_{n+n_0}(B^1_{1,\theta}[0,1], L_q[0,1]) \lesssim d_n(b^1_{1,\theta}(T), \ell^0_{q,2}(T)). \tag{2.28}$$

Шар $b^1_{1,\theta}(T)$ задаётся неравенством:

$$\left\| \left(\sum_{j=1}^{2^k} |x_{k,j}| \right)_{k \geqslant 0} \right\|_{\ell_\theta} \leqslant 1.$$

Это множество представляет собой бесконечномерный октаэдр при $\theta = 1$ и произведение октаэдров размерностей $1, 2, \dots, 2^k, \dots$ при $\theta = \infty$.

Оценка снизу. Оценим поперечник в (2.27), используя следствие 2.1. Нормы $\ell^0_{q,q}(T)$ и $\ell^0_{2,2}(T)$ представляют собой весовые нормы с весами $w_{k,j} = 2^{-k}$, поскольку

$$\|x\|^p_{\ell^0_{p,p}(T)} = \sum_{k=0}^{\infty} 2^{-k} \sum_{j=1}^{2^k} |x_{k,j}|^p.$$

Сначала рассмотрим случай $\theta = 1$; множество $b^1_{1,1}(T)$, как уже отмечалось, является октаэдром. Обозначим через k_0 и k_1, соответственно, минимальный и максимальный k, удовлетворяющий условиям $4n < 2^k < n^{\tilde{q}/2}$, где $\tilde{q} \in (2, q)$, например, $\tilde{q} = 1 + q/2$. Применим следствие 2.1 в следующих условиях:

- от вектора $x \in \mathbb{R}^T$ оставляем $N = \sum_{k=k_0}^{k_1} 2^k$ координат $x_{k,j}$, $k_0 \leqslant k \leqslant k_1$, $j = 1, \dots, 2^k$; от $b^1_{1,1}(T)$ остаётся октаэдр B^N_1;
- веса $w_{k,j} = 2^{-k}$, $k_0 \leqslant k \leqslant k_1$, $j = 1, \dots, 2^k$;
- $m = 1$ (один блок),
- $q_1 = 2$, $q_2 = q$.

Проверим выполнение условий утверждения 2.2. Условие 1):

$$\max w_{k,j} = 2^{-k_0} < (4n)^{-1}, \quad \sum w_{k,j} = (k_1 - k_0 + 1) \asymp \log n.$$

Условие 2) выполнено, так как можно считать, что n достаточно велико (в левой части неравенства единица, в правой — C/n). Следовательно, имеет место оценка

$$d_n \gtrsim \min\{n^{-1/2}(\log n)^{1/2}, 2^{-k_1/q}\} \asymp n^{-1/2} \log^{1/2} n,$$

поскольку $2^{-k_1/q} > n^{-\frac{1}{2}\tilde{q}/q}$.

При $\theta \in (1, 2]$ поперечник можно оценить через $\theta = 1$, так как тело $b^1_{1,\theta}$ расширяется с ростом θ.

При $\theta = \infty$ применим следствие 2.1 к тому же набору координат и весов, однако теперь множество $b^1_{1,\infty}(T)$ на этих координатах представляет собой произведение октаэдров $\prod_{k=k_0}^{k_1} B_1^{2^k}$ и координаты разбиты на $m \asymp \log n$ блоков $\Delta_k := \{(k, j)\}_{j=1}^{2^k}$. Аналогично предыдущему получаем

$$d_n \gtrsim \min\{(n/\log n)^{-1/2} \cdot (\log n)^{1/2}, 2^{-k_1/q}\} \asymp n^{-1/2} \log n.$$

Заметим, что мы доказали нечто большее: если оставить от T только

$$m = (k_1 - k_0 + 1) \asymp \log n$$

уровней, заменив его на

$$T_m := \{(k, j) \colon 0 \leqslant k < m,\ j = 1, \dots, 2^k\},$$

то оценка всё равно справедлива:

$$d_n(b^1_{1,\infty}(T_m), \ell^0_{q,q}(T_m) \cap \ell^0_{2,2}(T_m)) \gtrsim n^{-1/2} \log n. \tag{2.29}$$

Наконец, пусть $\theta \in (2, \infty)$. В неравенстве (2.27) заменим T на T_m, где m — то же, что и в (2.29). Вложим $b^1_{1,\infty}(T_m) \subset m^{1/\theta} b^1_{1,\theta}(T_m)$ и воспользуемся (2.29), получим требуемое неравенство.

Оценка сверху. Вначале покажем, что в (2.28) можно ограничиться $m \asymp \log n$ блоками. Обозначим $x[k] := (x_{k,j})_{j=1}^{2^k}$. Норма $\ell^0_{q,2}$ записывается так:

$$\|x\|_{\ell^0_{q,2}} = (\sum_{k=0}^{\infty} 2^{-2k/q} \|x[k]\|_q^2)^{1/2}.$$

Если $x \in b^1_{1,\theta}(T)$, то $\|x[k]\|_1 \leqslant 1$ при всех k, следовательно, $\|x[k]\|_q \leqslant 1$, и хвост оценивается:

$$\sum_{k \geqslant m} 2^{-2k/q} \|x[k]\|_q^2 \lesssim 2^{-2m/q}.$$

При $m \geqslant C_q \log n$ этот "хвост" мал, и он не повлияет на оценку поперечника сверху:

$$d_n(b^1_{1,\theta}(T), \ell^0_{q,2}(T)) \leqslant d_n(b^1_{1,\theta}(T_m), \ell^0_{q,2}(T_m)) + C_q n^{-1}.$$

Следовательно, достаточно оценить поперечник для T_m, $m \asymp \log n$.

Рассмотрим случай $\theta = 2$. Множество координат разбито на блоки $\Delta_k = \{(k, j)\}_{j=1}^{2^k}$. Для блоков размера $2^k < n/4$ возьмём всё пространство $\mathbb{R}^{\Delta_k}$, на это мы "потратим" размерность не более $n/2$. В каждом из остальных блоков возьмём пространство размерности $l \approx n/(2m)$, экстремальное для поперечника $d_l(B_1^{2^k}, \ell_q^{2^k})$. В соответствии с (1.15), мы получим аппроксимацию:

$$\|x[k] - y[k]\|_q \lesssim l^{-1/2} 2^{k/q} \|x[k]\|_1 \asymp (n/m)^{-1/2} 2^{k/q} \|x[k]\|_1,$$

откуда

$$2^{-k/q} \|x[k] - y[k]\|_q \lesssim (n/m)^{-1/2} \|x[k]\|_1,$$

суммируя по k:

$$\|x-y\|^2_{\ell^0_{q,2}(T_m)} \lesssim (n/m)^{-1} \sum_{k=0}^{m-1} \|x[k]\|_1^2 \leqslant (n/m)^{-1},$$

поскольку $x \in b^1_{1,2}(T_m)$. Итого, погрешность приближения $\lesssim n^{-1/2} \log^{1/2} n$.

Случай $\theta \in [1,2)$ вытекает из уже рассмотренного, так как тело $b^1_{1,\theta}$ расширяется с ростом θ.

При $\theta \geqslant 2$ воспользуемся тем, что $b^1_{1,\theta}(T_m) \subset m^{1/2-1/\theta} b^1_{1,2}(T_m)$,

$$d_n(b^1_{1,\theta}(T_m), \ell^0_{q,2}(T_m)) \leqslant m^{1/2-1/\theta} d_n(b^1_{1,2}(T_m), \ell^0_{q,2}(T_m)) \lesssim n^{-1/2} \log^{1-1/\theta} n.$$

□

Глава 3

Поперечники классов функций высокой гладкости

3.1 Относительные поперечники

В этом параграфе мы работаем с функциями на $\mathbb{T}$ и для краткости зачастую пишем L_p, W_p^r, $\mathcal{W}_p^r$ вместо $L_p(\mathbb{T})$, $W_p^r(\mathbb{T})$ и $\mathcal{W}_p^r(\mathbb{T})$, соответственно. Напомним, что $W_p^r = \{f\colon \|f^{(r)}\|_p \leqslant 1\}$, $\mathcal{W}_p^r = \{f\colon f^{(r)} \in L_p\}$, см. §1.1.

Оценка расстояния. Пространства L_p и $L_{p'}$ находятся в двойственности; будем использовать обозначение

$$\langle f, g\rangle = \int_{\mathbb{T}} f(x)g(x)\,dx, \quad f \in L_p,\ g \in L_{p'}.$$

(Как обычно, p' — это сопряжённый показатель.)

Для функций на $\mathbb{T}$ с нулевым средним определён периодический интеграл:

$$I_{\mathbb{T}}g := Ig - \frac{1}{2\pi}\int_{\mathbb{T}} Ig, \quad \text{где } Ig(x) := \textstyle\int_0^x g(t)\,dt,\ x \in [0, 2\pi].$$

Лемма 3.1. *Пусть $r \in \mathbb{N}$, $1 \leqslant p \leqslant \infty$, $M > 0$, Q_n — n-мерное линейное подпространство $\mathcal{W}_p^r$. Справедлива оценка*

$$\rho(W_p^r, MW_p^r \cap Q_n)_p \geqslant \sup_{G \in W_{p'}^r} \left(E_0(G)_{p'} - M \sup_{u \in Q_n^{(r)}, \|u\|_p \leqslant 1} \langle u, G\rangle \right),$$

где $E_0(G)_{p'}$ — это величина наилучшего приближения функции G константами в метрике $L_{p'}$.

Доказательство. Применим формулу двойственности (1.19) для расстояния:

$$\rho(W_p^r, MW_p^r \cap Q_n)_p \geqslant \sup_{\|g\|_{p'}\leqslant 1} \left(\sup_{f\in W_p^r} \langle f, g\rangle - \sup_{h\in MW_p^r\cap Q_n} \langle h, g\rangle \right). \qquad (3.1)$$

(В случае $X = L_\infty$ сопряжённое пространство шире L_1, поэтому мы довольствуемся оценкой снизу.)

Можно считать, что Q_n содержит пространство констант, иначе расстояние бесконечно. Следовательно, внешний супремум достаточно брать по функциям g с нулевым средним. Интегрируя по частям r раз, мы получим, что $\langle f, g\rangle = (-1)^r\langle f^{(r)}, G\rangle$, $\langle h, g\rangle = (-1)^r\langle h^{(r)}, G\rangle$, где через G обозначен r-кратный периодический интеграл от g (т.е. $G := I_{\mathbb{T}}^r g$). Пусть $z := f^{(r)}$, $u := h^{(r)}$, $Q_n^{(r)} := \{h^{(r)} \colon h \in Q_n\}$, тогда выражение в правой части (3.1) равно

$$\sup_{\|G^{(r)}\|_{p'}\leqslant 1} \left(\sup_{\|z\|_p\leqslant 1, \int_{\mathbb{T}} z=0} \langle z, G\rangle - M \sup_{u\in Q_n^{(r)}, \|u\|_p\leqslant 1} \langle u, G\rangle \right).$$

В силу двойственности (1.20) (при $p = 1$ нужно воспользоваться (1.21)), число $\sup\langle z, G\rangle$ равно наилучшему приближению G в $L_{p'}$ константами, и мы получаем требуемую оценку. □

Случай $M = 1 + o(1)$.

Теорема 3.1. *Пусть $r, n \in \mathbb{N}$, $r \geqslant 3$. Для всех $\varepsilon \in [n^{-2}, 1]$ справедливо неравенство:*

$$d_n(W_\infty^r(\mathbb{T}), (1+\varepsilon)W_\infty^r(\mathbb{T}), L_\infty(\mathbb{T})) \geqslant c(r)n^{-r}\varepsilon^{1-r/2}.$$

Пусть $m \in \mathbb{N}$. Построим функцию на $\mathbb{T}$ следующим образом: разобьём $\mathbb{T}$ на m равных отрезков; на каждом из них положим её равной некоторому знаку на левой половине и противоположному знаку на правой:

$$s_\sigma(x) := \sum_{k=0}^{m-1} \sigma_k \Big(\mathbf{1}_{(2\pi\frac{k}{m}, 2\pi\frac{k+1/2}{m})}(x) - \mathbf{1}_{(2\pi\frac{k+1/2}{m}, 2\pi\frac{k+1}{m})}(x) \Big),$$

где $\sigma = (\sigma_0, \ldots, \sigma_{m-1}) \in \{-1, 1\}^m$.

Лемма 3.2. *Пусть $m, n \in \mathbb{N}$, $m \leqslant n$. Тогда*

$$d_n(\{s_\sigma(x-\theta) \colon \sigma \in \{-1,1\}^m,\ \theta \in \mathbb{T}\}, L_2) \geqslant c(m/n)^{1/2}.$$

Доказательство. Пусть $\mathbf{s}$ — случайная функция $x \mapsto s_\sigma(x-\theta)$, где все $\sigma_k \in \{-1,1\}$ и $\theta \in \mathbb{T}$ выбраны случайно, равномерно и независимо друг от друга. Нетрудно видеть, что за счёт усреднения по θ при любых $j \neq l$ имеем $\mathsf{E}\,\widehat{\mathbf{s}}(j)\overline{\widehat{\mathbf{s}}(l)} = 0$. Воспользуемся утверждением 1.4:

$$d_n\big(\{s_\sigma(\cdot-\theta)\}_{\sigma,\theta}, L_2\big) \geqslant d_n^{\mathrm{avg}}(\mathbf{s}, L_2)_2 = (2\pi\sum_{j>n}\lambda_j)^{1/2}, \tag{3.2}$$

где $\lambda_1 \geqslant \lambda_2 \geqslant \ldots$ — невозрастающая перестановка чисел $\mathsf{E}\,|\widehat{\mathbf{s}}(j)|^2$, $j \in \mathbb{Z}$.

Вычислим λ_j, положив для краткости $f_k := \mathbf{1}_{(2\pi\frac{k}{m}, 2\pi\frac{k+1/2}{m})} - \mathbf{1}_{(2\pi\frac{k+1/2}{m}, 2\pi\frac{k+1}{m})}$:

$$\mathsf{E}\,|\widehat{\mathbf{s}}(j)|^2 = \mathsf{E}\,|\widehat{s_\sigma}(j)|^2 = \mathsf{E}\sum_{k_1=0}^{m-1}\sigma_{k_1}\widehat{f_{k_1}}(j)\sum_{k_2=0}^{m-1}\sigma_{k_2}\overline{\widehat{f_{k_2}}(j)} = \sum_{k=0}^{m-1}|\widehat{f_k}(j)|^2 = m|\widehat{f_0}(j)|^2,$$

$$2\pi|\hat{f}_0(j)| = \left|\int_0^{\frac{\pi}{m}} e^{-ijx}\,dx - \int_{\frac{\pi}{m}}^{\frac{2\pi}{m}} e^{-ijx}\,dx\right| = j^{-1}|(e^{-i\frac{\pi j}{m}}-1) - (e^{-i\frac{2\pi j}{m}} - e^{-i\frac{\pi j}{m}})| =$$

$$= j^{-1}|e^{-\pi ij/m}-1|^2 = 4j^{-1}\sin^2\frac{\pi j}{2m}.$$

Кроме того, $\widehat{\mathbf{s}}(0) \equiv 0$.

Задача свелась к следующей: из последовательности

$$\left\{mj^{-2}\sin^4\frac{\pi j}{2m}\colon j \in \mathbb{Z}\setminus\{0\}\right\}$$

убрали n наибольших чисел, нужно оценить снизу сумму оставшихся. Разобьём множество индексов на пачки размера $4m$: $\{j\colon |j| = 2sm+k,\ 1 \leqslant k \leqslant 2m\}$, для $s = 0, 1, \ldots$ Так как всего убрали n чисел, есть не более $n/(2m)$ пачек, из которых убрали больше половины. Нетрудно видеть, что в каждой из оставшихся пачек сумма чисел не меньше $c(s+1)^{-2}$. Следовательно, оцениваемая сумма не меньше $c\sum_{s>\frac{n}{2m}} s^{-2} \asymp m/n$. □

Перейдём к доказательству теоремы 3.1.

Доказательство. Пусть Q_n — некоторое приближающее пространство. Без ограничения общности можно считать, что $Q_n \subset \mathcal{W}_p^r$; положим $Q_n^{(r)} := \{f^{(r)}\colon f \in Q_n\}$. В силу леммы 3.1,

$$K := \rho(W_\infty^r, Q_n \cap (1+\varepsilon)W_\infty^r)_\infty = \sup_{G\in W_1^r}\left(E_0(G)_1 - (1+\varepsilon)\sup_{u\in Q_n^{(r)},\|u\|_\infty\leqslant 1}\langle u, G\rangle\right).$$

Положим $m := \lfloor An\varepsilon^{1/2} \rfloor$, где A — положительная абсолютная постоянная, которую подберем позже. Мы построим функцию $G \in \mathcal{W}_1^r$ со свойствами:

(i) $E_0(G)_1 = \|G\|_1$.

(ii) $\|G^{(r)}\|_1 \leqslant C_r \|G\|_1 m^r$.

(iii) $\langle u, G\rangle \leqslant (1-2\varepsilon)\|G\|_1$ для $u \in Q_n^{(r)}$, $\|u\|_\infty \leqslant 1$.

Из свойств (i), (ii) и (iii) вытекает нужная нам оценка: действительно, нормируем G условием $\|G^{(r)}\|_1 = 1$ и получим

$$K \geqslant E_0(G)_1 - (1+\varepsilon)\sup\langle u, G\rangle \geqslant \|G\|_1 - (1+\varepsilon)(1-2\varepsilon)\|G\|_1 \geqslant \\ \geqslant \varepsilon\|G\|_1 \geqslant \varepsilon C_r^{-1} m^{-r} \asymp n^{-r}\varepsilon^{1-r/2}.$$

В силу леммы 3.2, существует набор знаков σ и сдвиг $\theta \in \mathbb{T}$ такие, что

$$\inf_{u \in Q_n^{(r)}} \|s_\sigma(x-\theta) - u(x)\|_2 \geqslant c(m/n)^{1/2}. \tag{3.3}$$

Для простоты обозначений считаем, что неравенство выполнено с $\theta = 0$. Построим функцию G, $\operatorname{sign} G = s_\sigma$, удовлетворяющую условиям (i)–(iii).

Зафиксируем бесконечно-гладкую функцию $\varphi\colon [0,1] \to [-1,1]$ такую, что

$$\varphi(t) = \begin{cases} t, & 0 \leqslant t \leqslant 0.1, \\ 1, & 0.2 \leqslant t \leqslant 0.3, \\ 0.5 - t, & 0.4 \leqslant t \leqslant 0.6, \\ -1, & 0.7 \leqslant t \leqslant 0.8, \\ t - 1, & 0.9 \leqslant t \leqslant 1, \end{cases}$$

и $1/10 \leqslant \varphi(t) \leqslant 1$ на $[0.1, 0.4]$, $-1 \leqslant \varphi(t) \leqslant -1/10$ на $[0.6, 0.9]$.

Рассмотрим функцию

$$\widetilde{G}(x) := \sum_{k=0}^{m-1} \sigma_k \varphi\Big(\frac{mx}{2\pi} - k\Big)\mathbf{1}_{(2\pi\frac{k}{m}, 2\pi\frac{k+1}{m})}(x).$$

Ясно, что $\operatorname{sign}\widetilde{G} = s_\sigma$, однако в точках $2\pi(k+1)/m$, в которых $\sigma_k\sigma_{k+1} = -1$, происходит разрыв производной; исправим это. Мы имеем $\widetilde{G}(x) = -\sigma_k$ при $mx/(2\pi) \in [k+0.7, k+0.8]$ и $\widetilde{G}(x) = \sigma_{k+1} = -\sigma_k$ при $mx/(2\pi) \in [k+$

$1.2, k + 1.3]$. Сделаем функцию константой на проблемном участке: $G(x) := -\sigma_k$ при $mx/(2\pi) \in [k + 0.7, k + 1.3]$. Проделав это для всех проблемных k, получим искомую функцию G (в остальных точках $G(x) = \widetilde{G}(x)$). Схематично "исправление" функции $\widetilde{G}$ показано на рис. 3.1. Проверим выполнение условий.

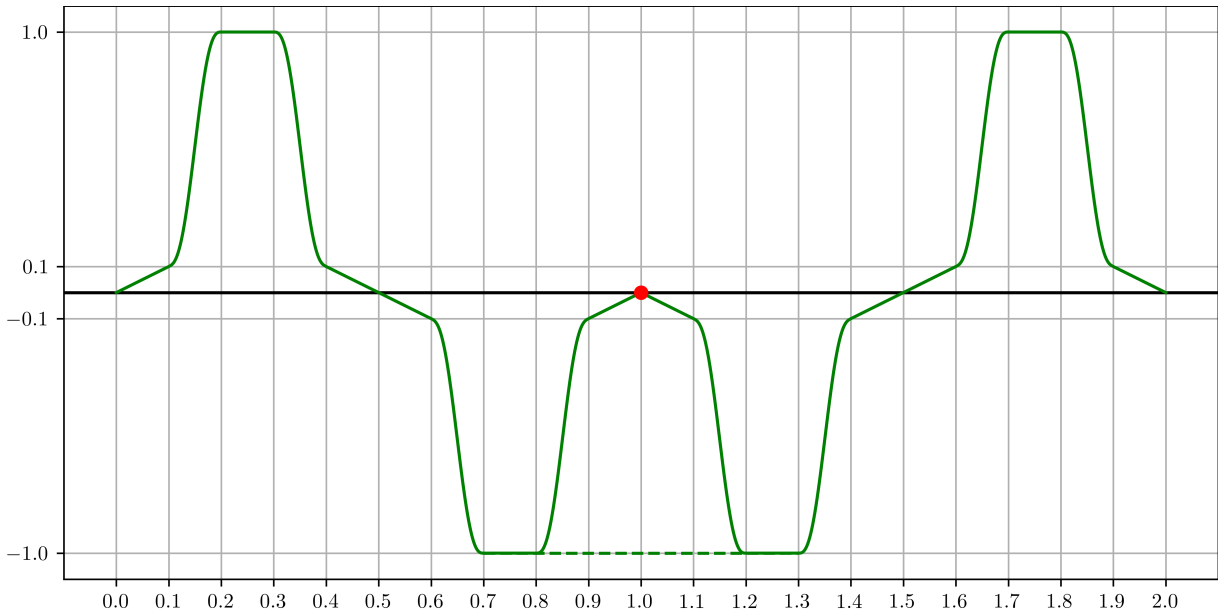


Рис. 3.1: Исправление функции $\widetilde{G}$.

Условие (i) следует из того, что $\operatorname{sign} G = s_\sigma \perp \{\text{const}\}$ (см., например, [95, Ch.3, §10]).

Условие (ii) следует из того, что в каждой точке x либо $G^{(r)}(x) = 0$, либо $G^{(r)}(x) = \pm m^r/(2\pi)^r \varphi^{(r)}(x')$ для некоторой точки x', поэтому $\|G^{(r)}\|_1 \leqslant \|G^{(r)}\|_\infty \leqslant C_r m^r$. В то же время $\|G\|_1 \geqslant 0.2 \cdot 2\pi$.

Для дальнейшего нам понадобятся множества $U_\gamma := \{x\colon |G(x)| \leqslant \gamma\}$. Заметим, что G имеет не более $2m$ нулей, $|G(x)| < 1/10$ только в окрестности этих нулей, причём $|G'| = m/(2\pi)$ в этих окрестностях. Следовательно, при $\gamma < 1/10$ имеем $\mu(U_\gamma) \leqslant 2m \cdot 4\pi\gamma/m = 8\pi\gamma < 30\gamma$.

Проверим (iii). Пусть $u \in Q_n^{(r)}$, $\|u\|_\infty \leqslant 1$. Тогда

$$\|G\|_1 - \langle u, G\rangle = \int_{\mathbb{T}} G(\operatorname{sign} G - u) = \int_{\mathbb{T}} |G|\,|\operatorname{sign} G - u| = \int_{\mathbb{T}} |G|\psi,$$

где $\psi := |s_\sigma - u|$. Заметим, что $0 \leqslant \psi \leqslant 2$, и в силу (3.3):

$$\int_{\mathbb{T}} \psi \geqslant \frac{1}{2}\int_{\mathbb{T}} \psi^2 \geqslant \delta := c_1 m/n.$$

Оценим снизу интеграл $\int_{\mathbb{T}} |G|\psi$. Рассмотрим задачу минимизации

$$\int_{\mathbb{T}} |G|\zeta \to \min, \quad 0 \leqslant \zeta \leqslant 2, \ \int_{\mathbb{T}} \zeta \geqslant \delta.$$

Экстремальная функция ζ^* имеет следующий вид: $\zeta^*(x) = 2$ на подходящем множестве U меры $\delta/2$ (точки, в которых значения $|G|$ малы) и $\zeta^*(x) = 0$ в остальных точках. Так как $\delta \leqslant \int_{\mathbb{T}} \psi \leqslant \pi$, то $\delta/120 < 1/10$ и мы имеем

$$\mu(U_{\delta/120}) \leqslant \frac{1}{4}\delta \leqslant \frac{1}{4}\int_{\mathbb{T}} \zeta^* = \frac{1}{2}\mu(U).$$

Следовательно, $|G| > \delta/120$ на половине носителя ζ^*, откуда

$$\|G\|_1 - \langle u, G\rangle = \int_{\mathbb{T}} |G|\psi \geqslant \int_{\mathbb{T}} |G|\zeta^* \geqslant \Big(\frac{\delta}{120}\Big)\mu(U) \geqslant \frac{\delta^2}{240}.$$

При достаточно большом A (см. определение m) имеем $\delta \geqslant 100\varepsilon^{1/2}$, откуда $\delta^2/240 \geqslant 4\pi\varepsilon \geqslant 2\varepsilon\|G\|_1$. Свойство (iii) установлено. Теорема доказана. □

Оценки при фиксированном M.

Теорема 3.2. *Пусть $n, r \in \mathbb{N}$, $r \geqslant 3$, $M \geqslant 2$; $p = 1$ или $p = \infty$.*

1. Для $\delta_1 := \exp(-c_1 M)$ и всех $n \geqslant \delta_1^{-1}$ имеет место неравенство

$$d_{2n}(W_p^r(\mathbb{T}), MW_p^r(\mathbb{T}), L_p(\mathbb{T})) \geqslant \frac{1}{8}\left(\frac{1+\delta_1}{n}\right)^r.$$

2. Для $\delta_2 := \exp(-c_2 M)$ и всех $n \geqslant C_1\delta_2^{-1}$ справедливо неравенство

$$d_{2n-1}(W_p^r(\mathbb{T}), MW_p^r(\mathbb{T}), L_p(\mathbb{T})) \leqslant \mathcal{K}_r\left(\frac{1+C_2\delta_2}{n}\right)^r.$$

Мы будем использовать ядра Валле Пуссена (см. §1.1). Напомним, что $V_{m,l} \in \mathcal{T}_{l-1}$ и $\widehat{V_{m,l}}(j) = 1$ при $|j| \leqslant m$.

Ключевым элементом доказательства оценки снизу является теорема Ал.А. Привалова. Напомним, что проектор U — это линейный ограниченный оператор, тождественный на образе. Мы будем использовать обозначение $U(f, x) := (Uf)(x)$.

Теорема 3.A ([61])**.** *Пусть $m, n \in \mathbb{N}$, $m \leqslant 2n$, U — проектор из $C(\mathbb{T})$ на $\mathcal{T}_n$. Тогда в любом подпространстве $C(\mathbb{T})$ коразмерности m найдется функция f, для которой $\|Uf\|_\infty/\|f\|_\infty > c\log\frac{2n+1}{m}$.*

Аналог теоремы Привалова справедлив для $p = 1$; доказательство проходит аналогичным образом (и даже несколько проще). Мы формулируем утверждение в чуть более сильном виде (достаточно существования проектора на $\mathcal{T}_{3n}$), но это не будет играть роли в дальнейшем.

Утверждение 3.1. *Пусть $m, n \in \mathbb{N}$, $m \leqslant 2n$, U — проектор из $\mathcal{T}_{3n}$ на $\mathcal{T}_n$. Тогда в любом подпространстве $\mathcal{T}_{3n}$ коразмерности m найдется функция f, для которой $\|Uf\|_1/\|f\|_1 > c \log \frac{2n+1}{m}$.*

Доказательство. Можно считать, что n/m достаточно велико, поскольку на $\mathcal{T}_n$ оцениваемое отношение норм равно единице.

Введем параметр $k := \lfloor (n/m)^{1/3} \rfloor$, а также $m_1 := \lfloor n/k \rfloor$ и $s := 2mk$. Параметр k подобран таким образом, чтобы выполнялись следующие неравенства:

$$\log k > c \log(n/m),$$

$$2km_1 + s < 3n, \tag{3.4}$$

$$(k-1)m_1 < n - s, \tag{3.5}$$

$$(k+1)m_1 > n + 2s. \tag{3.6}$$

Проверим неравенство (3.5). Поскольку n/m и, следовательно, n, достаточно велики, то

$$\frac{1}{2}(n/m)^{1/3} \leqslant k \leqslant (n/m)^{1/3}, \quad n/2 < n - k < km_1 \leqslant n.$$

Имеем: $(k-1)m_1 = km_1 - m_1 \leqslant n - m_1$; требуется показать, что $m_1 > s$:

$$\frac{m_1}{s} > \frac{n}{2k \cdot 2mk} \geqslant \frac{n}{4m(n/m)^{2/3}} = \frac{1}{4}(n/m)^{1/3} > 1.$$

Неравенства (3.4), (3.6) проверяются аналогично.

Возьмем функцию $\tilde{f} = \tilde{f}_1 + \tilde{f}_2$ такую, что

- $\tilde{f} \in \mathcal{T}_{2k}$, $\hat{\tilde{f}}(\pm k) = 0$, $\|\tilde{f}\|_1 \leqslant 1$,
- $\tilde{f}_1 \in \mathcal{T}_{k-1}$, $\|\tilde{f}_1\|_1 > c \log k$,
- $\tilde{f}_2 \perp \mathcal{T}_k$.

Например, $\tilde{f} = c_1(V_{k,2k} - 2\cos(kx))$.

Далее полагаем $f_1(x) := \tilde{f}_1(m_1 x)$, $f_2(x) := \tilde{f}_2(m_1 x)$, $f := f_1 + f_2$.

Обозначим заданное подпространство коразмерности m через L^m. Подберем нетривиальный полином $Q \in \mathcal{T}_s$ такой, что

$$T(m_1 x)Q(x) \in L^m, \quad \text{если } T \in \mathcal{T}_{2k} \text{ и } \hat{T}(\pm k) = 0. \tag{3.7}$$

(Отметим, что, в силу (3.4), мы не выходим из пространства $\mathcal{T}_{3n}$.) Проверим, что это возможно. Пусть L^m задаётся некоторыми функционалами $\xi_1, \ldots, \xi_m$. Тогда (3.7) задаётся $m(4k-1)$ условиями, а именно, $\langle \xi_j, Q \rangle = 0$,

$$\langle \xi_j, \cos(q m_1 t)Q(t) \rangle = \langle \xi_j, \sin(q m_1 t)Q(t) \rangle = 0, \quad q \in \{1, 2, \ldots, 2k\} \setminus \{k\},$$

$j = 1, \ldots, m$. Поскольку $\dim(\mathcal{T}_s) = 2s + 1 = 4mk + 1 > m(4k-1)$, существует нетривиальный полином $Q \in \mathcal{T}_s$, для которого эти условия выполняются. Нормируем Q: $\|Q\|_1 = 1$.

Положим $g_\tau(t) := f(t+\tau)Q(t)$. В силу (3.4), при всех $\tau \in \mathbb{T}$ полином $g_\tau \in \mathcal{T}_{3n}$ и по построению (см. (3.7)) лежит в подпространстве L^m. Докажем, что $\|U g_\tau\|_1 / \|g_\tau\|_1 > c \log k$ при некотором τ. Пусть

$$F(x, \tau) := U(g_\tau(\cdot), x - \tau) = F_1(x, \tau) + F_2(x, \tau),$$

$$F_1(x, \tau) := U(f_1(\cdot + \tau)Q(\cdot), x - \tau), \quad F_2(x, \tau) := U(f_2(\cdot + \tau)Q(\cdot), x - \tau).$$

В силу (3.5), при фиксированном τ функция $f_1(\cdot + \tau)Q(\cdot)$ есть полином степени $\leqslant n$, поэтому U оставляет его на месте: $F_1(x, \tau) = f_1(x)Q(x - \tau)$. Рассмотрим функцию F_2:

$$f_2(\cdot + \tau)Q(\cdot) = \sum_{k+1 \leqslant |\nu| \leqslant 2k} c_\nu e^{i\nu m_1 \tau} h_\nu(\cdot) \overset{U}{\longmapsto} \sum_{k+1 \leqslant |\nu| \leqslant 2k} c_\nu e^{i\nu m_1 \tau} T_\nu(\cdot), \quad T_\nu \in \mathcal{T}_n.$$

Следовательно, $F_2(x, \tau)$ при каждом x является полиномом по τ со спектром внутри множества $\{\nu + \mu \colon |\nu| \geqslant (k+1)m_1, |\mu| \leqslant n\}$. В силу (3.6), $F_2(x, \cdot) \perp \mathcal{T}_{2s}$. Функция $F_1(x, \tau) = f_1(x)Q(x - \tau)$ при фиксированном x это полином по τ из пространства $\mathcal{T}_s$. Значит, $F(x, \cdot) * V_{s,2s}(\cdot) = F_1(x, \cdot)$ при фиксированном x, и мы можем оценить норму:

$$\|F(x, \cdot)\|_1 \geqslant \frac{1}{3}\|F(x, \cdot) * V_{s,2s}(\cdot)\|_1 = \frac{1}{3}\|F_1(x, \cdot)\|_1.$$

Проинтегрируем полученное неравенство по $x \in [0, 2\pi]$:

$$3\int_0^{2\pi}\int_0^{2\pi} |F(x,\tau)|\,dx\,d\tau \geqslant \int_0^{2\pi}\int_0^{2\pi} |F_1(x,\tau)|\,dx\,d\tau = \\ = \int_0^{2\pi}\int_0^{2\pi} |f_1(x)Q(x-\tau)|\,dx\,d\tau = \|f_1\|_1 \cdot \|Q\|_1 > c\log k.$$

С другой стороны, пусть $A = \sup_\tau \|Ug_\tau\|_1/\|g_\tau\|_1$, тогда

$$\int_0^{2\pi}\int_0^{2\pi} |F(x,\tau)|\,dxd\tau = \int_0^{2\pi} \|Ug_\tau\|_1\,d\tau \leqslant A\int_0^{2\pi} \|g_\tau\|_1\,d\tau = \\ = A\int_0^{2\pi}\int_0^{2\pi} |f(x+\tau)Q(x)|\,d\tau dx = A\cdot\|f\|_1\cdot\|Q\|_1 \leqslant A.$$

Отсюда $A > c\log k$, что и требовалось доказать. □

Лемма 3.3. *Пусть $p = 1$ или $p = \infty$, $n \geqslant 2$, X_{2n+1} — подпространство в L_p размерности $2n+1$ и $1/n \leqslant \delta \leqslant 1/2$. Тогда существует тригонометрический полином $t \in \mathcal{T}_{n_1}$, $n_1 := \lfloor n(1-\delta)\rfloor$ такой, что*

$$\sup_{u\in X_{2n+1}, \|u\|_p\leqslant 1} \langle u, t\rangle < \frac{C}{\log(1/\delta)}\|t\|_{p'}.$$

Доказательство. Без ограничения общности можно считать, что билинейная форма $\langle u, f\rangle$ невырождена на $X_{2n+1}\times\mathcal{T}_n$ и, следовательно, существуют сопряжённые базисы $u_1,\dots,u_{2n+1} \in X_{2n+1}$ и $t_1,\dots,t_{2n+1} \in \mathcal{T}_n$: $\langle u_i, t_j\rangle = \delta_{i,j}$. (Если это не так, немного "пошевелим" пространство X_{2n+1}.) Положим $Uf := \sum_{k=1}^{2n+1}\langle u_k, f\rangle t_k$; ясно, что это проектор $L_{p'}\to\mathcal{T}_n$, и при этом

$$\langle u, Uf\rangle = \langle u, f\rangle, \quad u\in X_{2n+1},\ f\in L_{p'}.$$

Подпространство, заданное условием $Uf \in \mathcal{T}_{n_1}$ (эквивалентно, $\widehat{Uf}(j) = 0$, $n_1 < |j| \leqslant n$) имеет коразмерность $m = 2(n-n_1) \leqslant 2(\delta n+1)$. По теореме Привалова (см. теорему 3.A при $p=1$ и утверждение 3.1 при $p=\infty$) существует функция f_0, для которой $Uf_0 \in \mathcal{T}_{n_1}$, $\|f_0\|_{p'} = 1$, $\|Uf_0\|_{p'} > c\log\frac{2n+1}{m} \geqslant c_1\log(1/\delta)$. Возьмём в качестве искомого полинома $t = Uf_0$. Тогда $\|t\|_{p'} > c\log(1/\delta)$, но

$$\forall u\in X_{2n+1}\quad \langle u,t\rangle = \langle u, Uf_0\rangle = \langle u, f_0\rangle \leqslant \|u\|_p\|f_0\|_{p'} \leqslant \|u\|_p,$$

что и требовалось. Лемма доказана. □

Доказательство теоремы 3.2. Оценка снизу (п.1). Пусть Q_{2n} — подпространство в L_p размерности $2n$, требуется оценить величину $K := \rho(W_p^r, MW_p^r \cap Q_{2n})_p$. Как и раньше, мы можем считать, что $Q_{2n} \subset \mathcal{W}_p^r$ и воспользоваться леммой 3.1:

$$K \geqslant \sup_{G \in W_{p'}^r} \left(E_0(G)_{p'} - M \sup_{u \in Q_{2n}^{(r)}, \|u\|_p \leqslant 1} \langle u, G \rangle \right).$$

Пусть X_{2n+1} — пространство размерности $2n+1$, содержащее $Q_{2n}^{(r)}$ и пространство констант. Построим функцию G, для которой

$$\sup_{u \in X_{2n+1}, \|u\|_p \leqslant 1} \langle u, G \rangle \leqslant \frac{1}{4M} \|G\|_{p'}. \tag{3.8}$$

Возьмём в качестве функции G полином t из леммы 3.3; чтобы обеспечить неравенство (3.8), достаточно выбрать $\delta = \exp(-c_1 M)$ с достаточно большим c_1. Нормируем G условием $\|G^{(r)}\|_{p'} = 1$.

Положим $\iota(x) \equiv (2\pi)^{-1/p}$, $\|\iota\|_p = 1$. В силу (3.8), $|\langle \iota, G \rangle| \leqslant \frac{1}{4M} \|G\|_{p'} \leqslant \frac{1}{4} \|G\|_{p'}$; выведем отсюда, что $E_0(G)_{p'} \geqslant \frac{3}{8} \|G\|_{p'}$. Действительно, в противном случае

$$\frac{3}{8}\|G\|_{p'} > E_0(G)_{p'} = \|G - c\|_{p'} \geqslant |\langle \iota, G \rangle - \langle \iota, c \rangle| \geqslant |\langle \iota, c \rangle| - \frac{1}{4}\|G\|_{p'} = \|c\|_{p'} - \frac{1}{4}\|G\|_{p'},$$

откуда $\|c\|_{p'} < (5/8)\|G\|_{p'}$ и $E_0(G)_{p'} \geqslant \|G\|_{p'} - \|c\|_{p'} > (3/8)\|G\|_{p'}$, противоречие.

Далее из (3.8) выводим:

$$M \sup_{u \in Q_{2n}^{(r)}, \|u\|_p \leqslant 1} \langle u, G \rangle \leqslant M \cdot \frac{1}{4M} \|G\|_{p'} \leqslant \frac{2}{3} E_0(G)_{p'},$$

откуда интересующее нас уклонение $K \geqslant \frac{1}{3} E_0(G)_{p'} \geqslant \frac{1}{8} \|G\|_{p'}$.

Наконец, по неравенству Бернштейна (1.5),

$$8K \geqslant \|G\|_{p'} \geqslant n_1^{-r} \|G^{(r)}\|_{p'} \geqslant (n(1-\delta))^{-r} > \left(\frac{1+\delta}{n} \right)^r.$$

2. Оценка сверху. При $M > (4/\pi^2) \ln r + C_0$ поперечник совпадает с колмогоровским и равен $\mathcal{K}_r n^{-r}$ (см. (0.10) и (0.12) и обсуждение ниже); требуемое неравенство выполнено. Далее считаем, что $2 \leqslant M \leqslant (4/\pi^2) \ln r + C_0$.

Для аппроксимации используем суммы Валле Пуссена при подходящем m и неравенство Фавара (1.6):

$$\|f - V_{m,n} * f\|_p \leqslant (v_{m,n} + 1) \rho(f, \mathcal{T}_{m-1})_p \leqslant (v_{m,n} + 1) \frac{\mathcal{K}_r \|f^{(r)}\|_p}{m^r}.$$

При этом $\|(V_{m,n} * f)^{(r)}\|_p = \|V_{m,n} * f^{(r)}\|_p \leqslant v_{m,n}\|f^{(r)}\|_p$. Отсюда по определению

$$d_{2n-1}(W_p^r, v_{m,n}W_p^r, L_p) \leqslant \rho(W_p^r, v_{m,n}W_p^r \cap \mathcal{T}_{n-1})_p \leqslant \mathcal{K}_r(v_{m,n}+1)m^{-r}, \qquad (3.9)$$

и нужно только правильно оценить $v_{m,n}$. Дальше важно будет правильно подобрать константы c_2, C_1, C_2 — те самые, что выписаны в формулировке теоремы.

Положим $m := \lceil n(1-\delta_2)\rceil$, где $\delta_2 = \exp(-c_2 M)$. Покажем, что при достаточно малом c_2 будет $v_{m,n} \leqslant M$ и, следовательно, (3.9) даёт оценку исходного относительного поперечника. Используем оценку нормы ядра Валле Пуссена (1.9). Из этой оценки получаем (т.к. $n\delta_2 \geqslant C_1$ достаточно велико):

$$v_{m,n} \leqslant \frac{4}{\pi^2}\ln\frac{n+m}{n-m} + C \leqslant \frac{4}{\pi^2}\ln\frac{2n}{n\delta_2 - 1} + C \leqslant \frac{4}{\pi^2}\ln\frac{3}{\delta_2} + C =$$
$$= \frac{4}{\pi^2}c_2 M + C' \leqslant M/2 + C' \leqslant M,$$

если $c_2 \leqslant 1$ и M достаточно велико: $M \geqslant M_0$. При $M < M_0$ имеем $\delta_2 \geqslant \exp(-c_2 M_0) \geqslant 3/4$, если c_2 мало; $m = \lceil n(1-\delta_2)\rceil \leqslant n/4 + 1$, откуда $v_{m,n} \leqslant (n+m)/(n-m) \leqslant (5n+4)/(3n-4) \leqslant 2$, поскольку n достаточно велико.

Для завершения доказательства нужно оценить правую часть (3.9):

$$\mathcal{K}_r(v_{m,n}+1)m^{-r} \leqslant \mathcal{K}_r(M+1)m^{-r} \leqslant \mathcal{K}_r(\frac{4}{\pi^2}\ln r + C_0 + 1)n^{-r}(1-\delta_2)^{-r}.$$

Докажем, что эта величина не превосходит $\mathcal{K}_r n^{-r}(1+C_2\delta_2)^r$. Таким образом мы свели дело к неравенству

$$(1+C_2\delta_2)^r(1-\delta_2)^r \geqslant \frac{4}{\pi^2}\ln r + C_0 + 1. \qquad (3.10)$$

Выберем c_2 достаточно малым так, что

$$\delta_2 = \exp(-c_2 M) \geqslant \exp(-c_2(\frac{4}{\pi^2}\ln r + C_0)) \geqslant \frac{1}{2}r^{-1/2}.$$

Теперь мы можем зафиксировать c_2. Поскольку $M \geqslant 2$, параметр δ_2 удовлетворяет ограничениям $(1/2)r^{-1/2} \leqslant \delta_2 \leqslant -e^{-2c_2}$.

Покажем, что (3.10) справедливо, если $C_2 \geqslant 4$ и r достаточно велико. Действительно, если $\delta_2 \geqslant 1/2$, то $(1+C_2\delta_2)(1-\delta_2) \geqslant 3(1-e^{-2c_2})$; при $\delta_2 < 1/2$, как нетрудно видеть, $(1+C_2\delta_2)(1-\delta_2) \geqslant 1+\delta_2$. При достаточно больших $r \geqslant r_0$ имеем: $3^r(1-e^{-2c_2})^r > (4/\pi^2)\ln r + C_0 + 1$, и $(1+\delta_2)^r \geqslant (1+\frac{1}{2}r^{-1/2})^r > (4/\pi^2)\ln r + C_0 + 1$, поэтому (3.10) выполнено. При $r \leqslant r_0$ параметр δ_2 отделён от нуля и единицы, и неравенство (3.10) можно удовлетворить за счёт достаточно большого C_2.

Теорема доказана. □

3.2 Классы W_∞^r на отрезке и чебышёвские сплайны

Этот параграф посвящён чебышёвским сплайнам — совершенным сплайнам с заданной степенью и количеством узлов и минимальной равномерной нормой на отрезке. О связи сплайнов и поперечников написано во введении.

В этом параграфе $\|\cdot\|$ обозначает норму в пространстве $C[-1,1]$.

Чебышёвские сплайны. Пусть $r \in \mathbb{N}$, $n \in \mathbb{Z}_+$. Функция $x \in W_\infty^r[-1,1]$ называется совершенным сплайном порядка r с n узлами, если $x^{(r)}$ по модулю равна единице на $[-1,1]$ и имеет n точек перемен знака (узлов) внутри $(-1,1)$. Рассмотрим задачу о наименьшем уклонении от нуля, $\|x\| \to \min$, на множестве совершенных сплайнов порядка r с не более чем n узлами. В.М. Тихомиров [78] доказал, что решение задачи существует и единственно (с точностью до знака), а экстремальный сплайн $x_{n,r}$ обладает следующими свойствами:

(Ч1) $x_{n,r}$ имеет ровно n узлов внутри $(-1,1)$;

(Ч2) $x_{n,r}$ имеет чебышёвский $(n+r+1)$-альтернанс (включающий концы отрезка), то есть принимает значения $\pm\|x_{n,r}\|$, чередуясь, в $n+r+1$ точке.

В.Н. Малоземов и А.Б. Певный [57] доказали, что совершенный сплайн порядка r с n узлами, обладающий свойствами (Ч1) и (Ч2), единственен и, следовательно, совпадает с $x_{n,r}$. Зафиксируем также свойство наименьшего уклонения от нуля:

(Ч3) норма $\|x_{n,r}\|$ минимальна среди всех совершенных сплайнов порядка r с $\leqslant n$ узлами.

Пусть $n = 0$. В силу свойства (Ч1), $x_{0,r} = \pm\frac{t^r}{r!} + P_{r-1}$, $P_{r-1} \in \mathcal{P}_{r-1}$, откуда, в силу минимальности нормы,

$$x_{0,r} = \pm\frac{T_r}{r!2^{r-1}}.$$

Исходя из этого равенства, функции $x_{n,r}$ стали называть *чебышёвскими сплайнами*. Приведем ещё несколько фактов о чебышёвских сплайнах, следуя [78, 79].

График функции $x_{n,1}$ представляет собой ломаную с узлами $\{-1+\frac{2k}{n+1}\}_{k=1}^{n}$. В альтернанс $x_{n,1}$ входят узлы, а также концы отрезка $[-1,1]$. Сплайн $x_{n,2}$ также допускает явное описание — он "склеивается" из парабол, узлы равноотстоящие.

В силу единственности, сплайн $x_{n,r}$ симметричен относительно нуля: $x_{n,r}(-t) = \pm x_{n,r}(t)$; легко понять, что знак равен $(-1)^{n+r}$. Следовательно, узлы $x_{n,r}$ также симметричны относительно нуля. При четном $n = 2n'$ имеем n' положительных узлов $\{\tau_j\}_{j=1}^{n'}$ и n' отрицательных $\{-\tau_j\}_{j=1}^{n'}$; при нечетном $n = 2n'+1$ имеем n' положительных узлов, n' отрицательных и узел $\tau = 0$.

Теорема 3.3. *Пусть $n \in \mathbb{N}$ и $n' := \lfloor n/2 \rfloor$. При $n \geqslant 2$ положительные узлы сплайнов $x_{n,r}$ удовлетворяют асимптотике:*

$$\lim_{r\to\infty} \tau_k(x_{n,r}) r^{1/2} = \varkappa_k^{(n)}, \quad k = 1, \ldots, n',$$

где числа $0 < \varkappa_1^{(n)} < \ldots < \varkappa_{n'}^{(n)} < \infty$ образуют канонический набор для системы $\{x^{n-2q}e^{-x^2}\}_{q=1}^{n'}$ в $L(0,\infty)$.

Имеет место асимптотика при $r \to \infty$:

$$\lim_{r\to\infty} \frac{\|x_{n,r}\|_{C[-1,1]}}{\|x_{0,r}\|_{C[-1,1]}} r^{n/2} = A_n,$$

где числа $A_n > 0$ определяются следующим образом: $A_1 = 1/\sqrt{\pi}$, а при $n \geqslant 2$:

$$A_n = (-1)^{n'-1} \frac{2}{\sqrt{\pi}} \sum_{k=1}^{n'} (-1)^{k-1} \varkappa_k^{n-1} e^{-\varkappa_k^2}, \quad \varkappa_k := \varkappa_k^{(n)}.$$

В этом параграфе используются некоторые дополнительные обозначения: T_m — многочлен Чебышёва: $T_m(\cos\theta) = \cos m\theta$; $E_m(f)$ — наилучшее равномерное приближение функции f многочленами из $\mathcal{P}_m$ на $[-1,1]$.

Запись $A \sim B$ обозначает асимптотическое равенство: $A = B(1+o(1))$; при этом мы считаем n фиксированным, а $r \to \infty$.

Численные расчёты. Теоретические результаты согласуются с компьютерными вычислениями. Код (программа) для вычисления чебышёвских сплайнов (с небольшим числом узлов) находится в свободном доступе на GitHub по адресу:

`https://github.com/malykhin-yuri/cheb-spline`

При $n = 2$ число $\varkappa_1^{(2)}$ задаёт канонический набор для $\{e^{-x^2}\}$ на $(0, \infty)$ и определяется уравнением

$$\int_0^{\varkappa_1^{(2)}} e^{-x^2}\,dx = \frac{1}{2}\int_0^{\infty} e^{-x^2}\,dx = \frac{\sqrt{\pi}}{4}.$$

При $n = 3$ функция меняется на xe^{-x^2} и возникает уравнение

$$\int_0^{\varkappa_1^{(3)}} xe^{-x^2}\,dx = \frac{1}{4}, \quad \varkappa_1^{(3)} = \sqrt{\ln 2}.$$

При $n = 4$ возникает уже система уравнений

$$\begin{cases} \int_{\varkappa_1}^{\varkappa_2} e^{-x^2}\,dx = \frac{\sqrt{\pi}}{4}, \\ \varkappa_1 e^{-\varkappa_1^2} = \varkappa_2 e^{-\varkappa_2^2}. \end{cases}$$

Приведём результаты вычислений для небольших n и их сравнение с асимптотикой (округление до трёх значащих цифр).

n	r	$\frac{\|x_{n,r}\|}{\|x_{0,r}\|}$	A_n	$A_n r^{-n/2}$	$\tau_1(x_{n,r})$	$\tau_1 \cdot r^{1/2}$	$\varkappa_1^{(n)}$
1	5	0.237	0.564	0.252	-	-	-
1	10	0.172	0.564	0.178	-	-	-
2	5	0.0783	0.429	0.0857	0.202	0.453	0.477
2	10	0.0407	0.429	0.0429	0.147	0.466	0.477
3	5	0.0317	0.391	0.0350	0.335	0.748	0.833
3	10	0.0117	0.391	0.0124	0.251	0.792	0.833

Пример чебышёвского сплайна показан на рисунке 3.2.

В дальнейшем мы не указываем индекс n и пишем просто $\varkappa_1, \ldots, \varkappa_{n'}$, для упрощения обозначений.

Канонические наборы. Напомним определение канонического набора для системы функций $\{\psi_j\}_{j=1}^M$ в $L(a, b)$. (Допускается случай $b = \infty$.) Это последовательность точек $a < x_1 < \ldots < x_K < b$ такая, что функция σ, равная ± 1 и меняющая знак в точках x_k, ортогональна всем функциям ψ_j (или, что то же самое, натянутому на них подпространству).

Предположим, система $\{\psi_j\}_{j=1}^M$ является чебышёвской в (a, b), т.е. она состоит из непрерывных функций и любая нетривиальная линейная комбинация $\psi = \sum_{j=1}^M \alpha_j \psi_j$ имеет не более M нулей в (a, b). Тогда в каноническом наборе

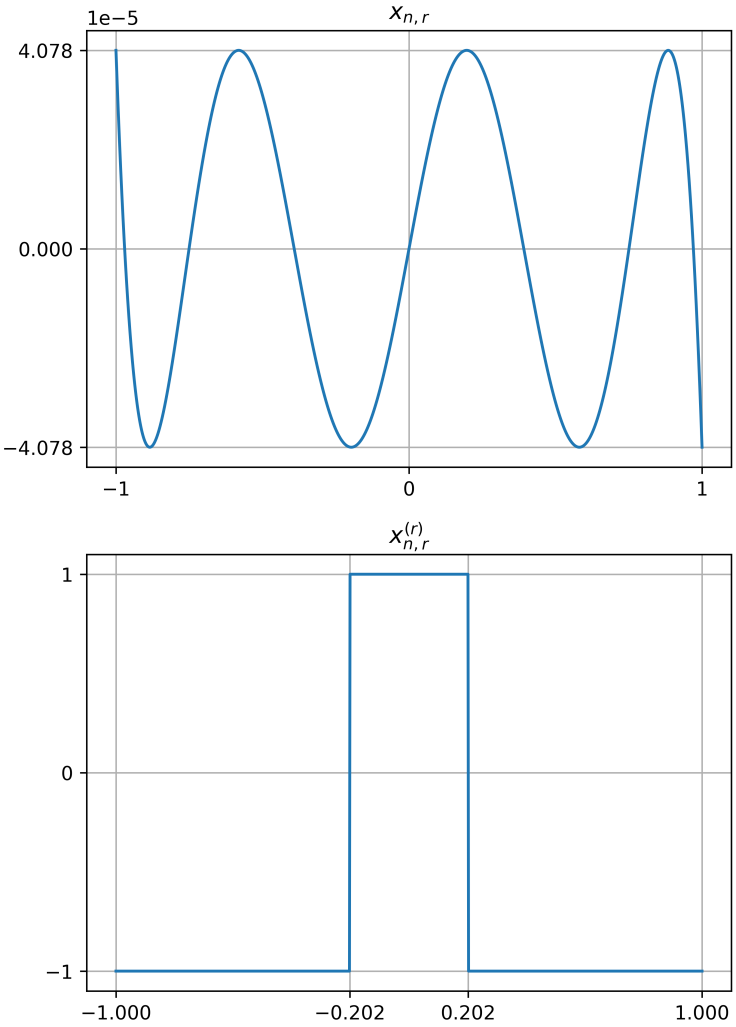


Рис. 3.2: Чебышёвский сплайн $x_{n,r}$ при $n = 2$, $r = 5$.

не может быть меньше M точек, иначе можно было бы построить функцию ψ, которая меняет в знак точках канонического набора; очевидно, что для неё ортогональность нарушается. Канонический набор из M точек, если существует, то единственен: если функции σ_1, σ_2 меняют знак в M точках, ортогональны $\{\psi_j\}$, и $\sigma_i(a{+}0) = 1$, то разность $\sigma_1 - \sigma_2$ тоже ортогональна $\{\psi_j\}$, но меняет знак в меньшем числе точек, и мы аналогичным образом приходим к противоречию.

Существование канонических наборов — общий факт, установленный Хобби и Райсом; позже короткое доказательство привел А. Пинкус [126].

Теорема 3.B ([110]). *Пусть $\{\psi_j\}_{j=1}^{M}$ — вещественные функции в $L_1([0,1],\mu)$, где мера μ неатомарна и конечна. Тогда существуют точки $\{x_j\}_{j=1}^{K}$, $K \leqslant M$,*

$0 =: x_0 < x_1 < \ldots < x_K < x_{K+1} := 1$ *такие, что*

$$\sum_{i=1}^{K+1}(-1)^i \int_{x_{i-1}}^{x_i} \psi_j(x)\, d\mu(x) = 0, \quad j = 1, \ldots, M.$$

В нашей теореме возникают системы $\{x^{2q}e^{-x^2}\}_{q=0}^{n'-1}$ при чётном $n = 2n'$ и $\{x^{2q+1}e^{-x^2}\}_{q=0}^{n'-1}$ при нечётном $n = 2n' + 1$. Теорема Хобби–Райса гарантирует существование канонических наборов из не более чем n' точек (полупрямую можно свести к отрезку, скажем, с помощью замены $z = x/(x+1)$). Эти системы являются чебышёвскими на $(0, +\infty)$, что следует, например, из правила знаков Декарта. Следовательно, для них существует единственный канонический набор из n' точек. Итак, мы обосновали корректность определения $\varkappa_j$ в теореме.

Вспомогательные понятия. Оставшаяся часть параграфа посвящена доказательству теоремы. Здесь и далее $n' = \lfloor n/2 \rfloor$.

Выберем знак $x_{n,r}$ так, что $x_{n,r}^{(r)}(+0) = 1$. Выпишем функцию, r-я производная которой равна по модулю $r!$ и меняет знак в n симметричных относительно нуля узлах:

$$\varphi^+_{n,r,\{\tau_j\}}(t) := t_+^r + 2\sum_{k=1}^{n'}(-1)^k(t-\tau_k)_+^r,$$
$$\varphi_{n,r,\{\tau_j\}}(t) := \varphi^+_{n,r,\{\tau_j\}}(t) + (-1)^{n+r}\varphi^+_{n,r,\{\tau_j\}}(-t).$$

(При чётном n, как нетрудно видеть, узла в нуле не возникает.)

В силу свойства (Ч3), сплайн $x_{n,r}$ является решением следующей экстремальной задачи:

$$r!\|x_{n,r}\| = \min_{\{\tau_j\}, P_{r-1}\in\mathcal{P}_{r-1}} \|\varphi_{n,r,\{\tau_j\}} - P_{r-1}\|.$$

Через $\varphi_{n,r}$ мы обозначим функцию $\varphi_{n,r,\{\tau_j\}}$, соответствующую оптимальным узлам, и аналогично для $\varphi^+_{n,r}$; таким образом, $r!x_{n,r} = \varphi_{n,r} - \hat{P}_{r-1}$,

$$r!\|x_{n,r}\| = E_{r-1}(\varphi_{n,r}). \tag{3.11}$$

В наших рассуждениях ключевую роль будет играть следующий функционал на пространстве $C[-1,1]$:

$$\Lambda_m(f) := \frac{1}{m}\left(\frac{1}{2}f(1) + \sum_{k=1}^{m-1}(-1)^k f\left(\cos\frac{\pi k}{m}\right) + (-1)^m\frac{1}{2}f(-1)\right).$$

Свойства этого функционала зафиксируем в следующей лемме.

Лемма 3.4. *Пусть $f \in C[-1,1]$, $m \in \mathbb{N}$. Имеют место следующие свойства:*

1. $E_{m-1}(f) \geqslant |\Lambda_m(f)|$;
2. $|\Lambda_m(f)| \leqslant \|f\|$;
3. $\Lambda_m(P) = 0$ *для* $P \in \mathcal{P}_{m-1}$;
4. $\Lambda_m(T_m) = 2^{m-1}\Lambda_m(t^m) = 1$;
5. $\Lambda_m(f(t)) = (-1)^m \Lambda_m(f(-t))$.

Доказательство. Для п.2) достаточно заметить, что сумма модулей коэффициентов, входящих в определение Λ_m, равна 1. Для доказательства пп.3), 4) вычислим значение $\Lambda_m(T_s)$, пользуясь тем, что $T_s(\cos(\pi k/m)) = \cos(\pi sk/m)$:

$$2m\Lambda_m(T_s) = 1 + 2\sum_{k=1}^{m-1}(-1)^k \cos\frac{\pi sk}{m} + (-1)^m \cos(\pi s) = D_{m-1}(\pi + \frac{\pi s}{m}) + (-1)^{m+s}. \tag{3.12}$$

При $s < m$, используя (1.7), получаем

$$2m\Lambda_m(T_s) = \frac{\sin\frac{(m-1/2)\pi(s+m)}{m}}{\sin\frac{\pi(s+m)}{2m}} + (-1)^{m+s} = 0.$$

При $s = m$ получаем $2m\Lambda_m(T_m) = D_{m-1}(0) + 1 = 2m$, что и требовалось.

Из пп.2) и 3) сразу следует п.1): для любого $P \in \mathcal{P}_{m-1}$ имеем $\|f - P\| \geqslant |\Lambda_m(f - P)| = |\Lambda_m(f)|$.

Наконец, п.5) проверяется непосредственно по определению. □

Лемма мотивирована следующими соображениями. Можно доказать (мы не будем это ни доказывать, ни использовать в доказательствах), что нормированный чебышёвский сплайн близок к многочлену Чебышёва:

$$\lim_{r\to\infty} \left\| \frac{x_{n,r}}{x_{n,r}(1)} - T_{n+r} \right\| = 0.$$

Поэтому должна хорошо работать оценка снизу через Λ_m, основанная на точках альтернанса многочлена Чебышёва T_m.

Дальнейший план доказательства таков: сначала мы изучим поведение функционалов Λ_m для функций $(t-\tau)_+^r$, входящих в $\varphi_{n,r}$. Потом получим оценки снизу на $E_{r-1}(\varphi_{n,r})$ через Λ_m и выведем из них асимптотику узлов и нижнюю оценку нормы $x_{n,r}$. Далее получим оценку нормы сверху.

Асимптотика Λ_m. Цель этого подпараграфа — получить асимптотическую формулу для величины $\Lambda_m(f)$, где $f(t) = (t-\tau)_+^{m-\nu}$, при фиксированном $\nu \in \mathbb{Z}_+$ и $m \to \infty$.

Начнем с вычислений для $\tau = 0$.

Лемма 3.5. *Пусть $m \in \mathbb{N}$, $\nu < m$. Имеем $\Lambda_m(t_+^m) = 2^{-m}$, $\Lambda_m(t_+^{m-\nu}) = 0$ для чётных $\nu \geqslant 2$. Для нечётных $\nu = 2\nu' + 1$ справедливо равенство*

$$\Lambda_m(t_+^{m-\nu}) = (-1)^{\nu'} m^{-1} 2^{-m+\nu-1} \Delta_{\frac{2}{m}}^{m-\nu} F\left(-1 + \frac{\nu}{m}\right), \tag{3.13}$$

где $F(x) = \operatorname{tg}(\pi x/2)$ при четном m, $F(x) = 1/\cos(\pi x/2)$ при нечетном m.

Напомним, что через Δ_h^l обозначена конечная разность

$$\Delta_h^l F(x) := \sum_{k=0}^{l} (-1)^k \binom{l}{k} F(x + (l-k)h).$$

Доказательство. Воспользуемся леммой 3.4. Имеем $t^m = t_+^m + (-1)^m (-t)_+^m$, откуда $2^{1-m} = \Lambda_m(t^m) = \Lambda_m(t_+^m) + (-1)^m \Lambda_m((-t)_+^m) = 2\Lambda_m(t_+^m)$, и случай $\nu = 0$ разобран. В случае чётного $\nu \geqslant 2$ всё аналогично, только $\Lambda_m(t^{m-\nu}) = 0$, и мы получаем 0.

Рассмотрим теперь случай нечетного $\nu = 2\nu' + 1$.

Пусть $m = 2m'$. Обозначим $\zeta = \exp(i\pi/m)$. Выпишем $\Lambda_m(t_+^{m-\nu})$, пользуясь определением, после чего преобразуем сумму:

$$2m\Lambda_m(t_+^{m-\nu}) = 1 + 2\sum_{k=1}^{m'-1} (-1)^k \left(\cos\frac{\pi k}{m}\right)^{m-\nu} =$$

$$= \sum_{k=-m'}^{m'-1} (-1)^k \left(\cos\frac{\pi k}{m}\right)^{m-\nu} = \sum_{k=-m'}^{m'-1} (-1)^k \left(\frac{\zeta^k + \zeta^{-k}}{2}\right)^{m-\nu};$$

$$2^{1+m-\nu} m\Lambda_m(t_+^{m-\nu}) = \sum_{k=-m'}^{m'-1} (-1)^k (\zeta^k + \zeta^{-k})^{m-\nu} =$$

$$= \sum_{k=-m'}^{m'-1} (-1)^k \sum_{l=0}^{m-\nu} \binom{m-\nu}{l} \zeta^{kl} \zeta^{-k(m-\nu-l)} = \sum_{l=0}^{m-\nu} \binom{m-\nu}{l} \sum_{k=-m'}^{m'-1} (-1)^k \zeta^{k(2l+\nu-m)}.$$

Пользуясь тем, что $\zeta^m = -1$, имеем $\zeta^{-mk} = (-1)^k$, откуда внутренняя сумма в полученном выражении равна

$$\sum_{k=-m'}^{m'-1} \zeta^{k(2l+\nu)} = \frac{1}{\zeta^{2l+\nu}-1} \cdot (\zeta^{m'(2l+\nu)} - \zeta^{-m'(2l+\nu)}) =$$

$$= \frac{\zeta^{-l-\nu/2}}{\zeta^{l+\nu/2} - \zeta^{-l-\nu/2}} \cdot (\zeta^{m'(2l+\nu)} - \zeta^{-m'(2l+\nu)}) =$$

$$= \left(\frac{1}{2i} \operatorname{ctg} \frac{\pi}{m}(l+\nu/2) - \frac{1}{2}\right) \cdot (2i \sin \frac{\pi}{m} m'(2l+\nu)) =$$

$$= (\operatorname{ctg} \frac{\pi}{m}(l+\nu/2) - i) \cdot (-1)^{l+\nu'}.$$

Поскольку исходное выражение было вещественным, мы можем суммировать только вещественные части сумм:

$$2^{1+m-\nu} m \Lambda_m(t_+^{m-\nu}) = \sum_{l=0}^{m-\nu} \binom{m-\nu}{l} (-1)^{l+\nu'} \operatorname{ctg} \frac{\pi}{m}(l+\nu/2) =$$

$$= (-1)^{\nu'} \sum_{l=0}^{m-\nu} \binom{m-\nu}{l} (-1)^l \operatorname{ctg} \frac{\pi}{m}(l+\nu/2) =$$

$$= (-1)^{\nu'} \sum_{l=0}^{m-\nu} \binom{m-\nu}{l} (-1)^l \operatorname{tg} \frac{\pi}{2m}(m-\nu-2l) = (-1)^{\nu'} \Delta_{\frac{2}{m}}^{m-\nu} \operatorname{tg}(\pi x/2)\Big|_{x=-1+\frac{\nu}{m}}.$$

Случай нечетного m разбирается аналогично. □

Выражение для $\Lambda_m(t_+^{m-\nu})$ через конечную разность не проще, чем изначальное определение. Однако следующее наблюдение позволяет получить простое выражение для его асимптотики при $m \to \infty$ и $\nu = O(1)$.

Утверждение 3.2. *Пусть функции G, H разлагаются в ряды*

$$G(x) = \sum_{k=1}^{\infty} G_k x^{2k-1}, \quad |x| < 1,$$

$$H(x) = \sum_{k=1}^{\infty} H_k x^{2k-1}, \quad |x| < 1,$$

причем коэффициенты $G_k, H_k > 0$. Тогда при всех $l \in \mathbb{N}$, $x \in (-1,1)$ и $h > 0$, при которых $x, x + (2l-1)h \in (-1,1)$, величины $\Delta_h^{2l-1} G(x)$ и $\Delta_h^{2l-1} H(x)$ положительны и выполнено неравенство

$$\frac{\Delta_h^{2l-1} G(x)}{\Delta_h^{2l-1} H(x)} \leqslant \sup_{k \geqslant l} \frac{G_k}{H_k}. \tag{3.14}$$

Следовательно, если $G_k \sim H_k$ *при* $k \to \infty$, *то*

$$\Delta_h^{2l-1} G(x) \sim \Delta_h^{2l-1} H(x) \quad \textit{при } l \to \infty. \tag{3.15}$$

Точнее, при указанных условиях на x *и* h *отношение левой и правой частей (3.15) равномерно по* x, h *стремится к единице при* $l \to \infty$.

Аналогичное утверждение верно для рядов вида $G(x) = \sum_{k=0}^{\infty} G_k x^{2k}$, *и величин* Δ_h^{2l}.

Доказательство. Разложим $G^{(2l-1)}$ в ряд с неотрицательными слагаемыми:

$$G^{(2l-1)}(x) = \sum_{k=l}^{\infty} G_k (2k-1)\dots(2k-2l+1) x^{2k-2l}$$

и сравним это разложение с аналогичным разложением для $H^{(2l-1)}$. Отсюда получаем $\sup G^{(2l-1)}(x)/H^{(2l-1)}(x) \leqslant \sup_{k \geqslant l} G_k/H_k$. Для доказательства неравенства (3.14) осталось применить (известную и легко доказываемую по индукции) формулу

$$\Delta_h^l F(x) := \int_0^h \cdots \int_0^h F^{(l)}(x + u_1 + \ldots + u_l)\, du_1\, du_2 \ldots du_l$$

к функциям G и H. □

Разложение тригонометрических функций в ряды и асимптотика коэффициентов хорошо известна.

Утверждение 3.A ([81], §4.3, §23.1)**.** *Справедливы следующие разложения в ряды в интервале* $|x| < \pi/2$ *(где* B_{2k} *и* E_{2k} *— числа Бернулли и Эйлера, соответственно):*

$$\operatorname{tg}(x) = \sum_{k=1}^{\infty} \frac{|B_{2k}|}{(2k)!} 2^{2k}(2^{2k}-1) x^{2k-1}, \quad |B_{2k}| = (-1)^{k-1} B_{2k} \sim 2\frac{(2k)!}{(2\pi)^{2k}},$$

$$\frac{1}{\cos x} = \sum_{k=0}^{\infty} \frac{|E_{2k}|}{(2k)!} x^{2k}, \quad |E_{2k}| = (-1)^k E_{2k} \sim \frac{(2k)! 4^{k+1}}{\pi^{2k+1}}.$$

Это позволяет сделать следующий шаг к нахождению асимптотики. Мы используем обозначение:

$$k!! = \prod_{j \in \mathbb{Z}_+ :\, k-2j>0} (k-2j), \quad \text{где } k \in \mathbb{Z};$$

в частности, $k!! = 1$ при $k \leqslant 1$.

Лемма 3.6. *При фиксированном нечётном $\nu = 2\nu' + 1$ и $m \to \infty$ имеет место асимптотика*

$$\Lambda_m(t_+^{m-\nu}) \sim (-1)^{\nu'} \frac{2}{\sqrt{\pi}} 2^{-m+\nu'} m^{-\nu'-1/2} (\nu-2)!! \tag{3.16}$$

Кроме того, для всех $m \in \mathbb{N}$, $\nu = 2\nu' + 1 \leqslant m/2$ справедливо неравенство

$$|\Lambda_m(t_+^{m-\nu})| \leqslant C 2^{-m+\nu} m^{-\nu'-1/2} (\nu-2)!!. \tag{3.17}$$

Доказательство. Установим (3.16). Из утверждения 3.A нетрудно вывести, что функция $G(x) := \operatorname{tg}(\pi x/2)$, возникающая в случае чётного m, "эквивалентна" в смысле утверждения 3.2 (т.е. $G_k \sim H_k$) функции $H(x) := \frac{4}{\pi}\frac{x}{1-x^2}$ (для которой $H_k = 4/\pi$). В случае нечетного m аналогичным образом функция $1/\cos(\pi x/2)$ "эквивалентна" дроби $\frac{4}{\pi}(1-x^2)^{-1}$.

В силу (3.15), мы можем заменить функцию на эквивалентную, откуда получаем при четных $m \to \infty$:

$$\Delta_{\frac{2}{m}}^{m-\nu} \operatorname{tg} x \Big|_{x=-1+\frac{\nu}{m}} \sim \frac{4}{\pi} \Delta_{\frac{2}{m}}^{m-\nu} \frac{x}{1-x^2} \Big|_{x=-1+\frac{\nu}{m}}.$$

Функцию $x/(1-x^2)$ можно заменить на $1/(1-x)$, поскольку $2x/(1-x^2) = 1/(1-x) - 1/(1+x)$, и всё симметрично относительно нуля. При нечетном m функция $1/\cos(\pi x/2)$ заменяется на $\frac{4}{\pi}(1-x^2)^{-1}$, а функцию $1/(1-x^2)$ можно заменить на $1/(1-x)$. Получаем, что для любых номеров m (чётных или нечётных) при $m \to \infty$:

$$\Delta_{\frac{2}{m}}^{m-\nu} F(-1+\frac{\nu}{m}) \sim \frac{4}{\pi} \Delta_{\frac{2}{m}}^{m-\nu} \frac{1}{1-x} \Big|_{x=-1+\frac{\nu}{m}}, \tag{3.18}$$

где функция F та же, что в (3.13).

Конечную разность функции $1/(1-x)$ легко вычислить по индукции:

$$\Delta_h^l \frac{1}{1-x} = \frac{l! h^l}{(1-x)(1-x-h)\dots(1-x-lh)}.$$

Подставляя l, h и x, получим:

$$\Delta_{\frac{2}{m}}^{m-\nu} \frac{1}{1-x} \Big|_{x=-1+\frac{\nu}{m}} = \frac{(m-\nu)!(2/m)^{m-\nu}}{(2-\frac{\nu}{m})(2-\frac{\nu+2}{m})\cdots(2-\frac{\nu+2(m-\nu)}{m})} =$$
$$= \frac{(m-\nu)! 2^{m-\nu} m}{(2m-\nu)(2m-\nu-2)\cdots\nu} = \frac{(m-\nu)! 2^{m-\nu} m (\nu-2)!!}{(2m-\nu)!!}.$$

Преобразуем знаменатель:

$$(2m-\nu)!! = \frac{(2m-2\nu')!}{(2m-2\nu')!!} = \frac{(2m-2\nu')!}{2^{m-\nu'}(m-\nu')!} = \frac{\binom{2m-2\nu'}{m-\nu'}(m-\nu')!}{2^{m-\nu'}}.$$

Отсюда после простых преобразований получим:

$$\Delta_{\frac{2}{m}}^{m-\nu}\frac{1}{1-x}\bigg|_{x=-1+\frac{\nu}{m}} = \frac{2^{2m-2\nu'}}{\binom{2m-2\nu'}{m-\nu'}}\frac{2^{-\nu'-1}m(\nu-2)!!}{(m-\nu')(m-\nu'-1)\dots(m-2\nu')} \tag{3.19}$$

Воспользовавшись соотношением $\binom{2k}{k} \sim 2^{2k}/\sqrt{\pi k}$ $(k \to \infty)$ и равенствами (3.13), (3.18), (3.19), получаем асимптотику (3.16).

Оценка сверху на $\Lambda(t_+^{m-\nu})$ получается аналогично, только вместо (3.15) нужно использовать неравенство (3.14), и оценить дробь (3.19) сверху. □

Перейдём к общему случаю $\tau \in [-1,1]$. Введём следующее обозначение, чтобы подчеркнуть зависимость от τ:

$$\lambda_{m,\nu}(\tau) := \Lambda_m((t-\tau)_+^{m-\nu}).$$

По определению,

$$\lambda_{m,\nu}(\tau) = \frac{1}{m}\left(\frac{1}{2}(1-\tau)^{m-\nu} + \sum_{k=1}^{m-1}(-1)^k(\cos\frac{\pi k}{m} - \tau)_+^{m-\nu}\right). \tag{3.20}$$

Имеет место симметрия относительно нуля (рассуждаем аналогично вычислению $\Lambda_m(t_+^m)$ в лемме 3.5):

$$\lambda_{m,0}(-\tau) = 2^{1-m} - \lambda_{m,0}(\tau), \tag{3.21}$$

$$\lambda_{m,\nu}(-\tau) = (-1)^{\nu-1}\lambda_{m,\nu}(\tau), \quad \nu \geqslant 1. \tag{3.22}$$

Далее, $\lambda_{m,\nu}(1) = 0$ при всех ν, поскольку $(t-1)_+ \equiv 0$ на $[-1,1]$. Отсюда и $\lambda_{m,\nu}(-1) = 0$ при $\nu \geqslant 1$.

Функция $\lambda_{m,\nu}$ является линейной комбинацией слагаемых вида $(t_k-\tau)_+^{m-\nu}$ и поэтому лежит в пространстве $C^{m-\nu-1}[-1,1]$. Нетрудно видеть, что на самом деле $\lambda_{m,\nu}$ пропорционально совершенному сплайну порядка $m-\nu$ с $(m-1)$ узлами, см. (3.20).

Исследуемые нами функции получаются друг из друга дифференцированием — нетрудно проверить, что

$$\lambda_{m,0}^{(s)}(\tau) = (-1)^s m(m-1)\cdots(m-s+1)\lambda_{m,s}(\tau). \tag{3.23}$$

Функции $\lambda_{m,\nu}$ мы будем контролировать "локально" (через значения функции и её производных в нуле) за счёт леммы 3.6, а также "глобально" с помощью подсчёта количества нулей в следующей лемме.

Лемма 3.7. *При $\nu \geqslant 1$ функция $\lambda_{m,\nu}$ имеет не более $\nu - 1$ нуля в интервале $(-1, 1)$.*

Доказательство. Для краткости обозначим $g(\tau) = \lambda_{m,\nu}(\tau)$. В силу (3.23), имеем $g(\pm 1) = g'(\pm 1) = \ldots = g^{(m-\nu-1)}(\pm 1) = 0$.

Предположим, g имеет не менее ν нулей. По теореме Ролля, функция $g'(\tau)$ имеет на $(-1, 1)$ не менее $\nu + 1$ нуля. Продолжая рассуждения, получим, что $g^{(m-\nu-1)}$ имеет на $(-1, 1)$ не менее $m - 1$ нуля. Но это невозможно, поскольку $g^{(m-\nu-1)}$ представляет собой кусочно-линейную функцию с ненулевой производной $\pm\mathrm{const}$, узлами в $\{\cos \pi k/m\}_{k=1}^{m-1}$, и обращающуюся в ноль в ± 1. □

Напомним определение многочленов Эрмита:

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

Хорошо известны следующие их свойства: H_n имеет n различных вещественных нулей; $H_n(-x) \equiv (-1)^n H_n(x)$; старший моном равен $(2x)^n$. Для единообразия формул введём также дополнительное обозначение H_{-1} для функции, задаваемой равенством

$$H_{-1}(x) e^{-x^2} = \int_x^\infty e^{-u^2}\, du.$$

Нам потребуется известное равенство $\int_{-\infty}^{\infty} e^{-u^2}\, du = \sqrt{\pi}$.

Теперь мы можем сформулировать и доказать основную лемму этого подпараграфа.

Лемма 3.8. *При $m \to \infty$ и фиксированном $\nu \in \mathbb{Z}_+$ имеет место асимптотическое равенство:*

$$\lambda_{m,\nu}(\tau) = \Lambda_m((t-\tau)_+^{m-\nu}) = 2^{-m} m^{-\nu/2} \left\{ \frac{2}{\sqrt{\pi}} H_{\nu-1}(\tau\sqrt{m}) e^{-m\tau^2} + o(1) \right\}, \quad \tau \in [-1, 1].$$

Величины $o(1)$ стремятся к нулю равномерно по $\tau \in [-1, 1]$.

Доказательство. Сначала докажем асимптотическую формулу в области $\tau = O(m^{-1/2})$. Точнее, докажем, что для любых фиксированных $\nu \in \mathbb{Z}_+$, $A > 0$ и $\varepsilon > 0$, при всех $m \geqslant m(\nu, A, \varepsilon)$ справедливо неравенство

$$\max_{|\tau| \leqslant Am^{-1/2}} |2^m m^{\nu/2} \lambda_{m,\nu}(\tau) - \frac{2}{\sqrt{\pi}} H_{\nu-1}(\tau\sqrt{m}) e^{-m\tau^2}| \leqslant \varepsilon. \tag{3.24}$$

Разберём случай $\nu = 0$. Выпишем разложение Тейлора для функции $\lambda_{m,0}(\tau)$ в точке $\tau = 0$, используя формулу (3.23). В силу леммы 3.5, слагаемые ряда Тейлора для чётных производных зануляются; для нечётных будем использовать асимптотику из леммы 3.6. Сумму разобьём на части (параметры N_1 и N_2 выберем позже):

$$\lambda_{m,0}(\tau) = S_1 + S_2 + R,$$

$$S_1 := \lambda_{m,0}(0) + \sum_{s=0}^{N_1-1} \lambda_{m,0}^{(2s+1)}(0) \frac{\tau^{2s+1}}{(2s+1)!}, \quad S_2 := \sum_{s=N_1}^{N_2-1} \lambda_{m,0}^{(2s+1)}(0) \frac{\tau^{2s+1}}{(2s+1)!},$$

$$R = \lambda_{m,0}^{(2N_2+1)}(\xi) \frac{\tau^{2N_2+1}}{(2N_2+1)!}.$$

Подставим $\tau = \varkappa m^{-1/2}$, $|\varkappa| \leqslant A$. Тогда при $m \to \infty$ и фиксированном N_1 имеем

$$2^m S_1 = 1 - 2^m \sum_{s=0}^{N_1-1} m(m-1)\dots(m-2s) \lambda_{m,2s+1}(0) \frac{\varkappa^{2s+1} m^{-s-1/2}}{(2s+1)!} =$$

$$= 1 - 2^m \sum_{s=0}^{N_1-1} m^{s+1/2} (-1)^s \frac{2}{\sqrt{\pi}} 2^{-m+s} m^{-s-1/2} (2s-1)!! \frac{\varkappa^{2s+1}}{(2s+1)!} (1 + o(1)) =$$

$$= 1 + \frac{2}{\sqrt{\pi}} \sum_{s=0}^{N_1-1} (-1)^{s-1} 2^s (2s-1)!! \frac{\varkappa^{2s+1}}{(2s+1)!} + o(1) =$$

$$= 1 + \frac{2}{\sqrt{\pi}} \sum_{s=0}^{N_1-1} (-1)^{s-1} \frac{\varkappa^{2s+1}}{(2s+1) \cdot s!} + o(1).$$

Полученная сумма совпадает с начальным отрезком разложения Тейлора функции $\frac{2}{\sqrt{\pi}} \int_\varkappa^\infty e^{-x^2}\, dx$. Следовательно,

$$\max_{|\varkappa| \leqslant A} |2^m S_1 - \frac{2}{\sqrt{\pi}} \int_\varkappa^\infty e^{-u^2}\, du| \leqslant \varepsilon/3$$

при достаточно больших $N_1 = N_1(A, \varepsilon)$ и $m \geqslant m(A, \varepsilon)$.

Оценим теперь сумму

$$2^m S_2 = 2^m \sum_{s=N_1}^{N_2-1} h^{(2s+1)}(0) \frac{(\varkappa m^{-1/2})^{2s+1}}{(2s+1)!}.$$

Пользуясь (3.17), получаем

$$|2^m S_2| \leqslant \sum_{s=N_1}^{N_2-1} 2^m m^{s+1/2} \, |\lambda_{m,2s+1}(0)| \, \frac{A^{2s+1}}{(2s+1)!} \leqslant$$

$$\leqslant C \sum_{s=N_1}^{N_2-1} 2^{2s}(2s-1)!! \frac{A^{2s+1}}{(2s+1)!} = C \sum_{s=N_1}^{N_2-1} 2^s \frac{A^{2s+1}}{(2s+1)s!}.$$

Ясно, что эта величина меньше $\varepsilon/3$ при достаточно большом $N_1 = N_1(\varepsilon, A)$ и произвольном $N_2 > N_1$.

Оценим остаточный член, воспользовавшись при $\tau \geqslant 0$ тривиальным неравенством $|\lambda_{m,\nu}(\tau)| = |\Lambda_m((t-\tau)_+^{m-\nu})| \leqslant \|(t-\tau)_+^{m-\nu}\| \leqslant 1$:

$$2^m |R| \leqslant 2^m m^{2N_2+1} \left|\Lambda_m((t-\xi)_+^{m-2N_2-1})\right| \frac{(\varkappa m^{-1/2})^{2N_2+1}}{(2N_2+1)!} \leqslant$$

$$\leqslant 2^m m^{N_2+1/2} \frac{A^{2N_2+1}}{(2N_2+1)!}.$$

При $N_2 = \lfloor m/3 \rfloor$ эта величина стремится к нулю и, следовательно, не превосходит $\varepsilon/3$ при $m \geqslant m(\varepsilon)$. Неравенство (3.24) получено для $\tau \geqslant 0$. Случай $\tau < 0$ сводится к рассмотренному: в силу (3.21), $2^m \lambda_{m,0}(-\tau) = 2 - 2^m \lambda_{m,0}(\tau)$ и

$$\frac{2}{\sqrt{\pi}} \int_{\varkappa}^{\infty} e^{-u^2} \, du = \frac{2}{\sqrt{\pi}} \Big(\int_{-\infty}^{\infty} e^{-u^2} \, du - \int_{-\infty}^{\varkappa} e^{-u^2} \, du \Big) = 2 - \frac{2}{\sqrt{\pi}} \int_{-\varkappa}^{\infty} e^{-u^2} \, du.$$

Аналогично разбирается случай $\nu \geqslant 1$.

Чтобы охватить весь отрезок $\tau \in [-1, 1]$, воспользуемся соображениями монотонности. При $\nu \geqslant 1$ многочлен $H_{\nu-1}(x)$ имеет $\nu - 1$ нуль на прямой, все эти нули простые и лежат в некотором отрезке $[-B_\nu, B_\nu]$. Воспользуемся доказанной нами асимптотикой (3.24) в области $|\tau| \leqslant 2B_\nu m^{-1/2}$ и леммой 3.7. Следовательно, все нули $\lambda_{m,\nu}$ в интервале $(-1, 1)$ при больших m находятся рядом с нулями $H_{\nu-1}$ и лежат на отрезке $|\tau| \leqslant 2B_\nu m^{-1/2}$. Поэтому функция $\lambda_{m,\nu}$ монотонна при $\tau m^{1/2} > 2B_{\nu+1}$ (в том числе при $\nu = 0$). Поскольку $\lambda_{m,\nu}(1) = 0$, и $H_{\nu-1}(\varkappa)e^{-\varkappa^2}$ стремится к нулю при $\varkappa \to +\infty$, то и $2^m m^{\nu/2} |\lambda_{m,\nu}(\tau)| \leqslant \varepsilon/2$ при

$\tau m^{1/2} > A(\nu,\varepsilon)$ и больших m, т.е. обе сравниваемые величины малы. Асимптотика получена для положительных τ. При $\tau < 0$ можно воспользоваться (3.21) и (3.22). Лемма полностью доказана. $\square$

Асимптотика узлов и оценка нормы сплайнов снизу. Перейдем непосредственно к доказательству теоремы. Оно проводится индукцией по количеству узлов. При индуктивном переходе от $n-1$ к n будем использовать, что $r!\|x_{n-1,r}\| = O(2^{-r}r^{-(n-1)/2})$. (Для $x_{0,r}$ эта оценка очевидна.) В силу неравенства $\|x_{n,r}\| \leqslant \|x_{n-1,r}\|$ (которое следует, например, из свойства (Ч3)), та же оценка справедлива для $x_{n,r}$.

Как обычно, $n' := \lfloor n/2 \rfloor$. Напомним (см. (3.11)), что $r!\|x_{n,r}\| = E_{r-1}(\varphi_{n,r})$, где $\varphi_{n,r}(t) = \varphi^+_{n,r}(t) + (-1)^{n+r}\varphi^+_{n,r}(-t)$,

$$\varphi^+_{n,r}(t) = t^r_+ + 2\sum_{k=1}^{n'}(-1)^k(t-\tau_k)^r_+.$$

В силу леммы 3.4, для любого $m \geqslant r$ справедлива оценка

$$r!\|x_{n,r}\| = E_{r-1}(\varphi_{n,r}) \geqslant E_{m-1}(\varphi_{n,r}) \geqslant |\Lambda_m(\varphi_{n,r})|.$$

Применим эту оценку для $m = n+r-2q$, $q = 0,1,\ldots,n'$. Лемма 3.8 позволяет записать асимптотику ($\nu = n-2q$):

$$\begin{aligned}
\Lambda_{n+r-2q}(\varphi_{n,r}) &= 2\Lambda_{n+r-2q}(\varphi^+_{n,r}) = \\
&= 2\left\{\Lambda_{n+r-2q}(t^r_+) + 2\sum_{k=1}^{n'}(-1)^k\Lambda_{n+r-2q}((t-\tau_k)^r_+)\right\} = \\
&= 2^{1-(n+r-2q)}(n+r-2q)^{-\frac{n-2q}{2}}\cdot\frac{2}{\sqrt{\pi}}\times \\
&\times\left\{H_{n-2q-1}(0) + 2\sum_{k=1}^{n'}(-1)^k H_{n-2q-1}(\tau_k r^{1/2})e^{-\tau_k^2 r} + o(1)\right\}.
\end{aligned}$$

Обозначим $\varkappa_{(r)} := (\tau_1 r^{1/r},\ldots,\tau_{n'}r^{1/2})$ и положим

$$F_q(x_1,\ldots,x_{n'}) := H_{n-2q-1}(0) + 2\sum_{k=1}^{n'}(-1)^k H_{n-2q-1}(x_k)e^{-x_k^2}.$$

Мы установили, что

$$\Lambda_{n+r-2q}(\varphi_{n,r}) = 2^{1-(n+r-2q)}r^{-\frac{n-2q}{2}}\frac{2}{\sqrt{\pi}}(F_q(\varkappa^{(r)}) + o(1)). \tag{3.25}$$

Равенство при $q = 0$ мы используем чуть позже; при $q \geqslant 1$ из оценки

$$|\Lambda_{n+r-2q}(\varphi_{n,r})| \leqslant r!\|x_{n,r}\| = O(2^{-r}r^{-(n-1)/2})$$

следует, что $\lim_{r\to\infty} F_q(\varkappa_{(r)}) = 0$. Предположим, $\varkappa_{(r)} \to \varkappa_{(\infty)}$, причём в $\varkappa_{(\infty)}$ мы допускаем как нулевые, так и бесконечные значения (сходимость понимается покоординатно). Тогда вектор $\varkappa_{(\infty)}$ удовлетворяет системе уравнений

$$F_q(x_1, \ldots, x_{n'}) = 0,\ q = 1, \ldots, n'. \tag{3.26}$$

Мы покажем, что у этой системы есть единственное решение $x = \varkappa$. Отсюда будет следовать, что $\varkappa_{(r)} \to \varkappa$ при $r \to \infty$, что и даст нужную нам асимптотику узлов.

Покажем, что система уравнений (3.26) эквивалентна тому, что точки $\{x_k\}_{k=1}^{n'}$ являются каноническим набором для системы функций $\{x^{n-2q}e^{-x^2}\}_{q=1}^{n'}$ на $(0, +\infty)$. Обозначим через σ функцию, равную по модулю единице, которая меняет знак в $\{x_k\}_{k=1}^{n'}$ и для которой $\sigma(0) = 1$. Пусть $F(x) \to 0$ при $x \to +\infty$ и $F' = f$, тогда (полагаем $x_0 := 0, x_{n'+1} := +\infty$):

$$\int_0^\infty f(x)\sigma(x)\,dx = \sum_{k=0}^{n'}(-1)^k \int_{x_k}^{x_{k+1}} f(x)\,dx =$$

$$= \sum_{k=0}^{n'}(-1)^k(F(x_{k+1}) - F(x_k)) = -(F(0) + 2\sum_{k=1}^{n'}(-1)^k F(x_k)). \tag{3.27}$$

В силу линейности, система (3.26) означает равенство нулю суммы (3.27) для всех $F \in \operatorname{span}\{H_{n-2q-1}(x)e^{-x^2}\}_{q=1}^{n'}$. Следовательно, интеграл в (3.27) равен нулю для всех $f = F'$; множество производных функций из указанного пространства есть $\operatorname{span}\{x^{n-2q}e^{-x^2}\}_{q=1}^{n'}$. Как мы уже обсуждали в начале параграфа, у такой системы не может быть канонического набора менее чем из n' точек; следовательно, все x_k отличны от 0 и ∞ и различны между собой. Кроме того, канонический набор определён однозначно.

Итак, асимптотика узлов получена.

Выпишем теперь оценку нормы $x_{n,r}$ снизу, пользуясь найденной асимптотикой узлов $\tau_k \sim \varkappa_k/\sqrt{r}$. Применим соотношение (3.25) при $q = 0$:

$$r!\|x_{n,r}\| \geqslant |\Lambda_{n+r}(\varphi_{n,r})|,$$

$$\Lambda_{n+r}(\varphi_{n,r}) = 2^{1-(n+r)}r^{-n/2}\frac{2}{\sqrt{\pi}}(F_0(\varkappa_1, \ldots, \varkappa_{n'}) + o(1)). \tag{3.28}$$

Мы докажем, что число

$$F_0(\varkappa_1, \ldots, \varkappa_{n'}) = H_{n-1}(0) + 2\sum_{k=1}^{n'}(-1)^k H_{n-1}(\varkappa_k)e^{-\varkappa_k^2} \tag{3.29}$$

не равно нулю и найдем его знак при $n \geqslant 2$ (в случае $n = 1$ дальнейшее очевидно). Возьмем многочлен $Q \in \mathcal{P}_n$, четный при четном n и нечетный при нечетном n, который меняет знак в точках $\varkappa_1, \ldots, \varkappa_{n'}$, причем $Q(+0) > 0$ и, следовательно, $\operatorname{sign} Q(+\infty) = (-1)^{n'}$. Тогда $\int_0^\infty Q(x)\sigma(x)e^{-x^2}\,dx > 0$, где функция σ определена так же, как в (3.27). Подставим в (3.27) функцию $F(x) = H_{n-1}(x)e^{-x^2}$, получим, что сумма (3.29) равна $-\int_0^\infty f(x)\sigma(x)\,dx$ для $f = F'$. Поскольку $H_{n-1}(x) = (2x)^{n-1} + P_1$, то $f(x) = (-(2x)^n + P_2(x))e^{-x^2}$. Поскольку $f(x) < 0$ при $x \to +\infty$, то $f(x) = (-1)^{n'+1}(cQ + P_3)e^{-x^2}$, где $c > 0$ (остатки P_1, P_2, P_3 — суть некоторые многочлены меньшей степени той же чётности). Отсюда, в силу каноничности $\{\varkappa_k\}$,

$$\operatorname{sign}\int_0^\infty f(x)\sigma(x)\,dx = \operatorname{sign}\int_0^\infty (-1)^{n'+1}cQ(x)\sigma(x)e^{-x^2}\,dx = (-1)^{n'+1}.$$

Итого, знак суммы (3.29) есть $(-1)^{n'}$.

В силу (3.26) и равенства $H_{n-1}(x) = (2x)^{n-1} + P_1$,

$$H_{n-1}(0) + 2\sum_{k=1}^{n'}(-1)^k H_{n-1}(\varkappa_k)e^{-\varkappa_k^2} = \\ = 2\sum_{k=1}^{n'}(-1)^k(2\varkappa_k)^{n-1}e^{-\varkappa_k^2} = -2^n\sum_{k=1}^{n'}(-1)^{k-1}\varkappa_k^{n-1}e^{-\varkappa_k^2}.$$

Отсюда получаем нижнюю оценку

$$r!\|x_{n,r}\| \geqslant |\Lambda_{n+r}(\varphi_{n,r})| \sim 2^{1-r}r^{-n/2}\frac{2}{\sqrt{\pi}}\left|\sum_{k=1}^{n'}(-1)^{k-1}\varkappa_k^{n-1}e^{-\varkappa_k^2}\right| = \\ = 2^{1-r}r^{-n/2}(-1)^{n'-1}\frac{2}{\sqrt{\pi}}\sum_{k=1}^{n'}(-1)^{k-1}\varkappa_k^{n-1}e^{-\varkappa_k^2} = A_n 2^{1-r}r^{-n/2}. \tag{3.30}$$

Оценка нормы сплайна сверху. В оценке снизу мы пользовались тем, что $r!\|x_{n,r}\| = E_{r-1}(\varphi_{n,r})$. На самом деле, имеет место и равенство

$$r!\|x_{n,r}\| = E_{n+r-1}(\varphi_{n,r}). \tag{3.31}$$

Действительно, по свойству (Ч2), сплайн $x_{n,r}$ обладает $(n+r+1)$-альтернансом, поэтому $\|x_{n,r}\| = E_{n+r-1}(x_{n,r})$. Поскольку $r!x_{n,r} - \varphi_{n,r} \in \mathcal{P}_{r-1}$, это равенство можно продолжить: $E_{n+r-1}(r!x_{n,r}) = E_{n+r-1}(\varphi_{n,r})$, и мы приходим к (3.31).

Для оценки нормы сверху нужно приблизить $\varphi_{n,r}$ многочленом степени $n+r-1$. Мы будем использовать интерполяцию по точкам альтернанса многочлена Чебышёва.

Лемма 3.9. *Пусть $f \in C[-1,1]$, $m \in \mathbb{N}$. Тогда многочлен*

$$P_m(f) := \Lambda_m(f)T_m + 2\sum_{\nu=1}^{m-1} \Lambda_m(fT_\nu)T_{m-\nu} + \Lambda_m(fT_m)$$

интерполирует f в точках $\{\cos\frac{\pi k}{m}\}_{k=0}^m$.

Доказательство. Введем на $C[-1,1]$ скалярное произведение

$$\langle f,g\rangle_m := \frac{1}{m}\left\{\frac{1}{2}f(1)g(1) + \sum_{k=1}^{m-1} f(\cos\frac{\pi k}{m})g(\cos\frac{\pi k}{m}) + \frac{1}{2}f(-1)g(-1)\right\}.$$

Система многочленов $\{T_k\}_{k=0}^m$ ортогональна относительно этого скалярного произведения:

$$\langle T_0,T_0\rangle_m = \langle T_m,T_m\rangle_m = 1, \quad \langle T_k,T_k\rangle = \frac{1}{2},\ 1 \leqslant k \leqslant m-1,$$
$$\langle T_i,T_j\rangle_m = 0, \quad 0 \leqslant i < j \leqslant m.$$

Эти равенства проверяются аналогично вычислению $\Lambda_m(T_s)$ в лемме 3.4. Следовательно, любую функцию можно разложить по системе $\{T_k\}_{k=0}^m$:

$$f = \sum_{k=0}^m b_k T_k \quad \text{на сетке } \{\cos\tfrac{\pi k}{m}\}_{k=0}^m,$$

$$b_k = \begin{cases} 2\langle f,T_k\rangle_m, & 1 \leqslant k \leqslant m-1, \\ \langle f,T_k\rangle_m, & k=0,\ k=m. \end{cases}$$

Многочлен $\sum_{k=0}^m b_k T_k$ и будет интерполяционным многочленом $P_m(f)$; осталось выразить его коэффициенты через f. Для любого $0 \leqslant \nu \leqslant m$ имеем

$$\langle f,T_{m-\nu}\rangle_m = \frac{1}{m}\left\{\frac{1}{2}f(1) + \sum_{k=1}^{m-1} f(\cos\frac{\pi k}{m})\cos\frac{\pi k(m-\nu)}{m} + \frac{1}{2}f(-1)(-1)^{m-\nu}\right\} =$$
$$= \frac{1}{m}\left\{\frac{1}{2}f(1) + \sum_{k=1}^{m-1}(-1)^k f(\cos\frac{\pi k}{m})\cos\frac{\pi k\nu}{m} + (-1)^m\frac{1}{2}f(-1)(-1)^\nu\right\} = \Lambda_m(fT_\nu).$$

Отсюда и из формулы для b_k получаем нужное нам выражение для $P_m(f)$. $\square$

Итак, приступим к оценке $\|x_{n,r}\|$ сверху. Возьмем функцию $f = \varphi_{n,r}$ и построим её интерполяционный многочлен $P_{n,r} := P_{n+r}(\varphi_{n,r})$. Обозначим через $L = \Lambda_{n+r}(\varphi_{n,r})$ старший коэффициент в разложении этого многочлена по системе многочленов Чебышёва:

$$P_{n,r} = LT_{n+r} + 2\sum_{\nu=1}^{n+r-1} \Lambda_{n+r}(\varphi_{n,r}T_\nu)T_{n+r-\nu} + \Lambda_{n+r}(\varphi_{n,r}T_{n+r}). \tag{3.32}$$

Имеем $P_{n,r} - LT_{n+r} \in \mathcal{P}_{n+r-1}$, откуда

$$E_{n+r-1}(\varphi_{n,r}) \leqslant \|\varphi_{n,r} - (P_{n,r} - LT_{n+r})\| \leqslant |L| + \|\varphi_{n,r} - P_{n,r}\|.$$

Асимптотику величины $L = \Lambda_{n+r}(\varphi_{n,r})$, являющейся оценкой снизу на $r!\|x_{n,r}\|$, мы уже вычислили в предыдущем подпараграфе, см. (3.30): $|L| \sim A_n 2^{1-r} r^{-n/2}$. Для получения той же асимптотики в оценке сверху и завершения доказательства теоремы нам достаточно доказать, что

$$\|\varphi_{n,r} - P_{n,r}\| = O(2^{-r} r^{-n/2-1/2}). \tag{3.33}$$

Сначала оценим старшие коэффициенты интерполяционного многочлена.

Лемма 3.10. *Для всех r, n и $0 \leqslant \nu \leqslant n+1$ справедливо неравенство*

$$|\Lambda_{n+r}(\varphi_{n,r}T_\nu)| \leqslant C(n) 2^{-r} r^{-(n-\nu-1)/2}.$$

Доказательство. Пусть сначала $\nu \leqslant n$. Для любого τ можно разложить многочлен $T_\nu(t)$ по степеням $(t-\tau)^k$: $T_\nu(t) = \sum_{k=0}^{\nu} g_{k,\nu}(\tau)(t-\tau)^k$. В силу леммы 3.8, имеем при $k \leqslant n$: $|\Lambda_{n+r}((t-\tau)_+^{r+k})| \lesssim 2^{-r} r^{-(n-k)/2}$. Тогда (полагаем $\tau_0 := 0$):

$$|\Lambda_{n+r}(\varphi_{n,r}T_\nu)| \leqslant 2|\Lambda_{n+r}(\varphi_{n,r}^+ T_\nu)| \leqslant 4\sum_{j=0}^{n'} |\Lambda_{n+r}((t-\tau_j)_+^r T_\nu(t))| \leqslant$$

$$\leqslant 4\sum_{j=0}^{n'}\sum_{k=0}^{\nu} |g_{k,\nu}(\tau_j)| \cdot |\Lambda_{n+r}((t-\tau_j)_+^{r+k})| \lesssim 2^{-r} r^{-(n-\nu)/2} \lesssim 2^{-r} r^{-(n-\nu-1)/2}.$$

При $\nu = n+1$ есть нюанс, заключающийся в том, что лемма 3.8 не применима: возникает величина $\Lambda_{n+r}((t-\tau)_+^{n+r+1})$. (Здесь в оригинальной публикации был пробел!) Обозначим $h(\tau) := \Lambda_m((t-\tau)_+^{m+1})$. Аналогично (3.23), можно продифференцировать эту функцию: $h'(\tau) = -(m+1)\Lambda_m((t-\tau)_+)^m = O(m2^{-m})$. Поскольку $h(1) = 0$, то $h(\tau) = -\int_\tau^1 h'(t)\,dt = O(m2^{-m})$. Следовательно, в случае $\nu = n+1$ также получается нужная оценка. □

Мы будем использовать, что многочлен $P_{n,r}$ интерполирует $\varphi_{n,r}$, и оценим погрешность стандартным методом оценки погрешности интерполяции. Нам потребуется взять производную порядка $r-1$ и проанализировать полученные функции. Для этого понадобится одна лемма (но, наконец, последняя в этом параграфе!) о многочленах.

Удобно использовать следующую норму в пространстве алгебраических многочленов $\mathcal{P}_\nu$:

$$\|\sum_{k=0}^{\nu} c_k x^k\|_{\mathcal{P}_\nu} := \max|c_k|.$$

Если степень не важна, пишем просто $\|\cdot\|_{\mathcal{P}}$. Отметим простую оценку

$$\|Q(x) - Q((1+\theta)x)\|_{\mathcal{P}_\nu} \leqslant C(\nu)|\theta| \cdot \|Q\|_{\mathcal{P}_\nu}, \quad Q \in \mathcal{P}_\nu,\ |\theta| < 1. \tag{3.34}$$

Лемма 3.11. *Для $\nu \leqslant m$ рассмотрим оператор $D_m^\nu\colon \mathcal{P}_m \to \mathcal{P}_\nu$, сопоставляющий многочлену Q многочлен*

$$D_m^\nu(Q;x) = \frac{m^{\nu/2}}{m!} Q^{(m-\nu)}(xm^{-1/2}).$$

1. При фиксированном ν и $m \to \infty$ имеет место сходимость $D_m^\nu 2^{1-m} T_m \to V_\nu$ в пространстве $\mathcal{P}_\nu$ к некоторому многочлену V_ν.

2. Для $m \geqslant 4$ определим многочлены $\omega_m \in \mathcal{P}_m$ равенством

$$\omega_m(t) := t \prod_{k=0}^{m-2} \left(t - \cos\frac{\pi k}{m-2}\right).$$

Справедливы соотношение $\omega_m = 2^{1-m}(T_m - T_{m-4})$ и оценка $\|\omega_m\| \leqslant 2^{2-m}$.

3. При фиксированном ν и $m \to \infty$ имеет место сходимость $D_m^\nu \omega_m \to V_\nu$ в пространстве $\mathcal{P}_\nu$.

4. Многочлен V_ν имеет степень ровно ν, и все его нули вещественные и простые (имеют кратность 1).

Доказательство. 1. Коэффициенты при старших мономах у многочлена $2^{1-m}T_m$ устроены следующим образом: $2^{1-m}T_m(t) = t^m + a_{1,m}t^{m-2} + a_{2,m}t^{m-4} + \ldots$, причем $a_{q,m} = a_q m^q(1+o(1))$ при фиксированном q и $m \to \infty$ (см. далее).

Рассмотрим действие D_m^ν на одно слагаемое:

$$\frac{m^{\nu/2}}{m!}\frac{d^{m-\nu}}{dt^{m-\nu}}a_{q,m}t^{m-2q}\Big|_{t=xm^{-1/2}} =$$
$$= \frac{m^{\nu/2}}{m!}(m-2q)(m-2q-1)\cdots(\nu-2q+1)\cdot a_{q,m}t^{\nu-2q}\Big|_{t=xm^{-1/2}} =$$
$$= m^{\nu/2}\frac{(m-2q)!}{m!(\nu-2q)!}a_q m^q(1+o(1))x^{\nu-2q}m^{-\frac{1}{2}(\nu-2q)} = \frac{a_q}{(\nu-2q)!}(1+o(1))x^{\nu-2q}.$$

Отсюда ясно, что

$$D_m^\nu 2^{1-m}T_m \to V_\nu(x) := \sum_{q=0}^{\lfloor\nu/2\rfloor} a_q\frac{x^{\nu-2q}}{(\nu-2q)!}. \tag{3.35}$$

Нам будет важно, что $a_q \neq 0$ при каждом q. Можно получить явный вид этих коэффициентов, воспользовавшись известными формулами (см., например, [81, §22.3]):

$$a_{q,m} = a_q\frac{m(m-q-1)!}{(m-2q)!}, \quad a_q = \frac{(-1)^q}{2^{2q}q!}.$$

2. Понятно, что $\omega_m(t) = t(t^2-1)2^{3-m}U_{m-3}(t)$, где U_{m-3} — многочлен Чебышёва второго рода. Подставим $t = \cos\theta$:

$$\omega(\cos\theta) = 2^{3-m}\cos\theta(-\sin^2\theta)U_{m-3}(\cos\theta) = -2^{3-m}\cos\theta\sin\theta\sin(m-2)\theta =$$
$$= -2^{2-m}\sin 2\theta\sin(m-2)\theta = 2^{1-m}(\cos m\theta - \cos(m-4)\theta),$$

что доказывает равенство $\omega_m = 2^{1-m}(T_m - T_{m-4})$. Оценка на $\|\omega_m\|$ следует мгновенно.

3. Нетрудно видеть, что $D_m^\nu 2^{-m}T_{m-4} \to 0$, так как коэффициенты имеют меньший порядок по m, чем нужно. Значит, $\lim\limits_{m\to\infty} D_m^\nu\omega_m = \lim\limits_{m\to\infty} D_m^\nu 2^{1-m}T_m = V_\nu$.

4. Из формул выше ясно, что $V_\nu \in \mathcal{P}_\nu \setminus \mathcal{P}_{\nu-1}$ и нужно лишь доказать невырожденность нулей.

Многочлен T_m имеет m вещественных нулей, поэтому $D_m^\nu T_m$ имеет ν нулей. Следовательно, V_ν также имеет ν вещественных нулей, но, возможно, с кратностями. Предположим, при некотором ν_0 один из нулей V_{ν_0} кратный: $V_{\nu_0}(x_0) = 0$, $V'_{\nu_0}(x_0) = 0$. Поскольку каждый второй коэффициент у V_{ν_0} ненулевой, то $x_0 \neq 0$. По построению (см. также (3.35)), $V'_\nu = V_{\nu-1}$. При переходе от V_ν к производной кратность всех нулей уменьшается на единицу и добавляются новые не

кратные нули (кратных добавиться не может — нулей станет слишком много). Следовательно, x_0 является нулем кратности 3 для V_{ν_0+1}, кратности 4 для V_{ν_0+2}, и т.д. При больших ν этот нуль имеет кратность $\nu - \nu_0 + 2 > \nu/2$, что невозможно, поскольку $-x_0$ имеет ту же самую кратность по соображениям чётности. □

Теперь мы можем завершить доказательство теоремы. Напомним, что нам осталось доказать оценку (3.33). Обозначим $R = \varphi_{n,r} - P_{n,r}$ и рассмотрим многочлен

$$\omega(t) := \omega_{n+r+2}(t) = t\prod_{k=0}^{n+r}(t - \cos\frac{\pi k}{n+r}).$$

Ясно, что R обращается в ноль в точках $\{\cos\frac{\pi k}{n+r}\}_{k=0}^{n+r}$. Если $n+r$ четно, то функции $\varphi_{n,r}, P_{n,r}, R, \omega$ четны, и нуль $t=0$ является двойным нулем для R и ω; если же $n+r$ нечетно, то указанные функции нечетны, и $t=0$ также является нулем R и ω. Оценим $R(t_0)$ в некоторой точке $t_0 \notin \{0, \cos\frac{\pi k}{n+r}\}$. Положим $K := R(t_0)/\omega(t_0)$ и $S := R - K\omega$. Функция S имеет на $[-1,1]$ не менее $n+r+4$ нулей за счет $n+r+2$ нулей ω и $\pm t_0$. По теореме Ролля, $S^{(r-1)}$ имеет не менее $n+5$ нулей. Мы покажем, что это невозможно при больших K и выведем отсюда оценку на R.

Итак, проанализируем $S^{(r-1)}$, записав эту функцию в удобном для нас виде:

$$\frac{1}{r!}S^{(r-1)}(x/r^{1/2}) = \frac{1}{r!}(\varphi_{n,r}^{(r-1)}(x/r^{1/2}) - P_{n,r}^{(r-1)}(x/r^{1/2}) - K\omega^{(r-1)}(x/r^{1/2})).$$

Рассмотрим каждое из слагаемых отдельно.

Функция $(1/r!)\varphi_{n,r}^{(r-1)}(t)$ является кусочно-линейной с производной ± 1, причём в нуле равна нулю. Значит, функция $\psi_{n,r}(x) := (1/r!)\varphi_{n,r}^{(r-1)}(x/r^{1/2})$ удовлетворяет условиям

$$|\psi_{n,r}(x_1) - \psi_{n,r}(x_2)| \leqslant r^{-1/2}|x_1 - x_2|, \quad \psi_{n,r}(0) = 0. \tag{3.36}$$

Второе слагаемое, $Q_{n,r}(x) := (1/r!)P_{n,r}^{(r-1)}(x/r^{1/2})$, получается дифференцированием старших членов в (3.32) (достаточно рассматривать $\nu \leqslant n+1$). Нам не важен явный вид получаемого многочлена, только его ограниченность:

$$\|Q_{n,r}\|_{\mathcal{P}_{n+1}} = O(1). \tag{3.37}$$

То, что многочлен имеет степень не выше $n+1$, очевидно. Оценим его $\mathcal{P}$-норму. Рассмотрим вклад слагаемых $\Lambda_{n+r}(\varphi_{n,r}T_\nu)T_{n+r-\nu}$, $0 \leqslant \nu \leqslant n+1$. В силу п.1

леммы 3.11, следующие нормы ограничены: $\|2^{-m}D_m^\nu T_m\|_{\mathcal{P}} = O(1)$, откуда (мы используем также (3.34)):

$$\frac{1}{r!}\|T^{(r-1)}_{n+r-\nu}(x/r^{1/2})\|_{\mathcal{P}} \lesssim \frac{1}{r!}\|T^{(r-1)}_{n+r-\nu}(x/(n+r-\nu)^{1/2})\|_{\mathcal{P}} =$$
$$= \frac{1}{r!}\frac{(n+r-\nu)!}{(n+r-\nu)^{(n-\nu+1)/2}}\|D^{n-\nu+1}_{n+r-\nu}T_{n+r-\nu}\|_{\mathcal{P}} \lesssim$$
$$\lesssim \frac{(n+r-\nu)!}{r!}(n+r-\nu)^{-(n-\nu+1)/2}2^{n+r-\nu} \asymp 2^r r^{n-\nu}r^{-(n-\nu+1)/2} = 2^r r^{(n-\nu-1)/2}.$$

Лемма 3.10 даёт оценку коэффициента: $\Lambda_{n+r}(\varphi_{n,r}T_\nu) = O(2^{-r}r^{-(n-\nu-1)/2})$. Умножая оценку для нормы на оценку коэффициента, получаем норму $O(1)$.

Наконец, последнее слагаемое представим в следующем виде:

$$\frac{1}{r!}K\omega^{(r-1)}(x/r^{1/2}) = r^{n/2+1/2}KW_{n,r}(x),$$

где $W_{n,r}$ лежит в $\mathcal{P}_{n+3}$ и, в силу п.3 леммы 3.11, сходится в этом пространстве:

$$W_{n,r}(x) = r^{-n/2-1/2}\frac{1}{r!}\omega^{(r-1)}(x/r^{1/2}) =$$
$$= r^{-n/2-1/2}\frac{1}{r!}\omega^{(r-1)}(x/(n+r+2)^{1/2})(1+o(1)) =$$
$$= r^{-n/2-1/2}\frac{(n+r+2)!}{r!(n+r+2)^{(n+3)/2}}(D^{n+3}_{n+r+2}\omega_{n+r+2})(1+o(1)) \to V_{n+3}.$$

Таким образом, третье слагаемое равно $\widetilde{K}W_{n,r}$, где

$$\widetilde{K} = r^{n/2+1/2}K, \quad \|W_{n,r} - V_{n+3}\|_{\mathcal{P}_{n+3}} = o(1) \text{ при } r \to \infty. \tag{3.38}$$

Итак, функция

$$\frac{1}{r!}S^{(r-1)}(x/r^{1/2}) = \psi_{n,r}(x) - Q_{n,r}(x) - \widetilde{K}W_{n,r}(x) \tag{3.39}$$

имеет по крайней мере $n+5$ нулей, причём слагаемые удовлетворяют условиям (3.36), (3.37), (3.38).

Докажем, что при больших $\widetilde{K}$ и r полученное уравнение имеет не более $n+3$ решений (что приведёт к противоречию).

Поскольку степень V_{n+3} выше, чем у второго слагаемого, то при, скажем, $|\widetilde{K}| \geqslant 1$ и больших r нули функции (3.39) сосредоточены на некотором отрезке $[-x_n, x_n]$. Поскольку на $[-x_n, x_n]$ первое и второе слагаемое, а также их производные равномерно ограничены, то при больших $\widetilde{K}$ и r решения (3.39) будут

близки к нулям V_{n+3} и, в силу невырожденности этих нулей, решений будет не более $n+3$.

Итак, $\widetilde{K} = O(1)$, $K = O(r^{-n/2-1/2})$. Вспомним, что $R(t_0) = K\omega(t_0)$. В силу п.4 леммы 3.11, $|R(t_0)| \leqslant |K|2^{-n-r} = O(2^{-r}r^{-n/2-1/2})$, что и требовалось доказать.

Глава 4

Полилинейные приближения

4.1 Векторы знаков для полиномиального отображения

Пусть $p_1, \ldots, p_N$ — вещественные алгебраические многочлены от K переменных, степени не выше d. Они определяют множество

$$\Omega := \mathbb{R}^K \setminus \bigcup_{i=1}^{N} \{x \in \mathbb{R}^K : p_i(x) = 0\}.$$

Теорема 4.A ([152]). *При $N \geqslant K$ количество компонент связности Ω не превосходит $(4edN/K)^K$.*

Рассмотрим векторы $p(x) = (p_1(x), \ldots, p_N(x))$ и их знаки:

$$\operatorname{sign} p(x) := (\operatorname{sign} p_1(x), \ldots, \operatorname{sign} p_N(x)).$$

Для определённости, будем рассматривать только векторы с ненулевыми элементами, то есть брать $x \in \Omega$.

Итак, каждая точка $x \in \Omega$ даёт сигнум-вектор $\operatorname{sign} p(x) \in \{-1, 1\}^N$. По соображениям непрерывности, количество таких сигнум-векторов не превосходит количества компонент связности Ω. Отсюда и из теоремы 4.A сразу следует

Теорема 4.B ([152]). *При $N \geqslant K$ количество различных сигнум-векторов $\operatorname{sign} p(x)$, $x \in \Omega$, не превосходит $(4edN/K)^K$.*

Уоррен применил это следствие для получения оценок снизу погрешности аппроксимации некоторых функциональных классов специальными множествами, обобщающими линейные пространства. Мы применяем его подход в

другой ситуации. Отметим, что Уоррен развивал технику Шапиро; в его работе [134] рассматриваются линейные функции (т.е. $d = 1$).

Утверждение 4.1. *Пусть $d \geqslant 2$. Если $N \geqslant AK \log_2 d$, то существует более чем $2^N - 2^{N(1-a)}$ сигнум-векторов $\sigma \in \{-1, 1\}^N$, которые отличаются от каждого вектора вида $\operatorname{sign} p(x)$, $x \in \Omega$, не менее чем в bN позициях. Можно взять $A = 18$, $a = 1/6$, $b = 1/18$.*

Это утверждение, по существу, также содержится в работе Уоррена; отличие в том, что ему нужно было лишь существование хотя бы одного такого вектора σ. Для полноты изложения мы приведём доказательство (для указанных в утверждении значений параметров; хотя нам эти конкретные значения не будут важны).

Доказательство. Покажем, что количество различных векторов $\operatorname{sign} p(x)$, $x \in \Omega$, меньше $2^{N/2}$. В силу теоремы 4.B, достаточно установить неравенство $(4edN/K)^K < 2^{N/2}$. Положим $t := N/K$ и прологарифмируем: $\log_2(4et) + \log_2 d < t/2$. Поскольку $t \geqslant 18 \log_2 d$, достаточно установить следующее неравенство:

$$\log_2(4et) < t/2 - t/18, \quad \text{при } t \geqslant 18.$$

Оно проверяется непосредственно при $t = 18$; дальше достаточно сравнить производные левой и правой частей.

Количество сигнум-векторов, отличающихся от фиксированного $\operatorname{sign} p(x)$ не более чем в m позициях, равно $1 + \binom{N}{1} + \ldots + \binom{N}{m} \leqslant (eN/m)^m$. Покажем, что $(eN/m)^m < 2^{N/3}$, если $m \leqslant N/18$. Положим $s := N/m$ и прологарифмируем: $es < 2^{s/3}$, что верно при $s \geqslant 18$.

Итак, количество всех векторов, близких к хотя бы одному $\operatorname{sign} p(x)$, меньше $2^{N/2} \cdot 2^{N/3} = 2^{5N/6}$, что и требовалось доказать. □

Можно интерпретировать этот результат так: пусть $M = \{p(x) \colon x \in \mathbb{R}^K\}$ — множество, заданное полиномиальным образом с использованием K параметров. Оно имеет топологическую размерность K и, конечно, не может заполнить всё $\mathbb{R}^N$. При достаточно большом N/K множество M мало и в следующем "комбинаторном" смысле: в нём "реализуется" экспоненциально малая доля знаков.

4.2 Жёсткость и сигнум ранг тензоров

Мы применим метод Шапиро–Уоррена к тензорам. Пусть $d \in \mathbb{N}$, $d \geqslant 2$, $m_1, \ldots, m_d \in \mathbb{N}$ и $N := m_1 m_2 \cdots m_d$. Пространство $\mathbb{R}^{m_1 \times m_2 \times \cdots m_d}$ состоит из тензоров, т.е. наборов чисел вида

$$T = (T_{i_1,\ldots,i_d}),\ 1 \leqslant i_j \leqslant m_j,\ j = 1, \ldots, d.$$

Тензор имеет ранг 1, если он имеет вид $T = u^1 \otimes u^2 \otimes \cdots \otimes u^d$, то есть

$$T_{i_1,\ldots,i_d} = u^1_{i_1} u^2_{i_2} \cdots u^d_{i_d}, \quad 1 \leqslant i_j \leqslant m_j,\ j = 1, \ldots, d,$$

для некоторых векторов $u^1 \in \mathbb{R}^{m_1}, \ldots, u^d \in \mathbb{R}^{m_d}$. (Формально, нулевой тензор имеет ранг 0.) Тензор имеет ранг не выше n, если он представляется в виде суммы не более чем n одноранговых тензоров. Отметим, что ранг любого тензора не превосходит $N/\max\{m_j\}$: действительно, пусть $\max\{m_j\} = m_1$, тогда любой тензор представляется в виде

$$T = \sum_{I=(i_2,\ldots,i_d)} T^I, \quad T^I_{j_1,j_2,\ldots,j_d} := T_{j_1,I} \cdot \mathbf{1}\{j_2 = i_2\} \cdots \mathbf{1}\{j_d = i_d\}, \tag{4.1}$$

ясно, что $\operatorname{rank} T^I \leqslant 1$.

Для сигнум-тензора, т.е. тензора с элементами ± 1, определено понятие *сигнум-ранга*. Это минимальный ранг тензоров с теми же знаками элементов, но произвольными абсолютными значениями. Формально, если $S \in \{-1, 1\}^{m_1 \times \cdots \times m_d}$, то

$$\operatorname{rank}_{\pm} S := \min\{\operatorname{rank} T \colon T \in \mathbb{R}^{m_1 \times \cdots \times m_d},\ \operatorname{sign} T_{i_1,\ldots,i_d} \equiv S_{i_1,\ldots,i_d}\}.$$

Здесь предполагается, что все элементы тензора T ненулевые.

Метод Шапиро–Уоррена применили Н. Алон, П. Франкл и В. Рёдль [85] для доказательства существования $m \times m$ сигнум матриц с большим сигнум-рангом: $\operatorname{rank}_{\pm} S \geqslant cm$. Этот подход работает и для тензоров.

Утверждение 4.2. *Пусть $N = m_1 \cdots m_d$, $d \geqslant 2$, $n < N$.*

1) Если $K := n(m_1 + \ldots + m_d) \leqslant N$, то количество сигнум-тензоров сигнум-ранга $\leqslant n$ не превосходит $(4edN/K)^K$.

2) Предположим, выполнено неравенство

$$N \geqslant An(m_1 + \ldots + m_d) \log_2 d. \tag{4.2}$$

Тогда найдётся множество сигнум-тензоров $\mathcal{S} \subset \{-1,1\}^{m_1\times\cdots\times m_d}$, $|\mathcal{S}| > 2^N - 2^{N(1-a)}$ *такое, что для любого* $S \in \mathcal{S}$ *и любого тензора* $T \in \mathbb{R}^{m_1\times\cdots\times m_d}$, $\operatorname{rank} T < n$, *найдётся не менее* bN *индексов* $(i_1,\ldots,i_d)$, *для которых* $S_{i_1,\ldots,i_d}T_{i_1,\ldots,i_d} \leqslant 0$.

3) В условиях п.2) для любого $S \in \mathcal{S}$ *имеем* $\operatorname{rank}_\pm S > n$.

4) В условиях п.2) для любого тензора R *такого, что* $\min|R_{i_1,\ldots,i_d}| \geqslant 1$ *и* $\operatorname{sign} R \in \mathcal{S}$, *имеем*

$$\min_{\operatorname{rank} T<n} \|R-T\|_{L_0^N} \geqslant b,$$
$$\min_{\operatorname{rank} T<n} \|R-T\|_{\ell_1^N} \geqslant bN. \tag{4.3}$$

(Числа A, a, b *те же, что в утверждении 4.1.)*

Доказательство. 1) Множество тензоров ранга не выше n — это подмножество $\mathbb{R}^N$, представимое в виде полиномиального образа $\mathbb{R}^K$. Действительно, такой тензор имеет вид

$$T_{i_1,\ldots,i_d} = \sum_{k=1}^{n} u_{i_1}^{1,k} u_{i_2}^{2,k} \cdots u_{i_d}^{d,k},$$

для некоторых $u_i^{j,k}$. Получается $K = n(m_1+\ldots+m_d)$ переменных $(u_i^{j,k})$ и N многочленов $p_{i_1,\ldots,i_d}$ степени d от этих переменных:

$$p_{i_1,\ldots,i_d}(u) = \sum_{k=1}^{n} u_{i_1}^{1,k} u_{i_2}^{2,k} \cdots u_{i_d}^{d,k}.$$

Значит, применим метод Шапиро–Уоррена и теорема 4.B об оценке количества знаков.

2) Применим утверждение 4.1 для тензоров ранга не выше n и получим множество "далёких" сигнум-тензоров $\mathcal{S}$, $|\mathcal{S}| \geqslant 2^N - 2^{N(1-a)}$. Далее однако, есть нюанс: тензор T может иметь нулевые координаты, и конструкция к нему не применима. Однако если $\operatorname{rank} T < n$, то для достаточно малых $\varepsilon > 0$ тензор $T' = T + \varepsilon\mathbf{1}$ (т.е. ко всем элементам прибавили ε) имеет ранг не выше n, ненулевые координаты и $\operatorname{sign} T' = \operatorname{sign} T$ там, где T ненулевой. Для bN индексов выполнено $S_{i_1,\ldots,i_d} \neq \operatorname{sign} T'_{i_1,\ldots,i_d}$. При этом либо элемент $T_{i_1,\ldots,i_d} = 0$, либо он ненулевой и того же знака, что и T'. В любом случае $S_{i_1,\ldots,i_d}T_{i_1,\ldots,i_d} \leqslant 0$, что и требовалось.

3) Если $\operatorname{rank}_\pm S \leqslant n$, то $S = \operatorname{sign} T$, $\operatorname{rank} T \leqslant n$. Но по построению для bN индексов знаки S и T отличаются, что невозможно.

4) Напомним, что если не менее bN элементов тензора $R - T$ по модулю $\geqslant b$, то $\|R - T\|_{L_0} \geqslant b$. Но в нашем случае есть не менее bN элементов, для которых $R_{i_1,\dots,i_d} T_{i_1,\dots,i_d} \leqslant 0$ и, значит, $|R_{i_1,\dots,i_d} - T_{i_1,\dots,i_d}| \geqslant |R_{i_1,\dots,i_d}| \geqslant 1$.

Отсюда получается и оценка в ℓ_1^N. □

Прежде чем перейти к полилинейным приближениям, дадим одно следствие для поперечников.

Жёсткость случайных систем функций. В главе 1 было показано, что независимые системы функций и близкие к ним являются жёсткими в L_1 и даже в L_0. Типичным примером является система Радемахера $r_1, \dots, r_N$. Заметим, что вектор $(r_1(t), \dots, r_N(t))$ принимает 2^N различных значений, и функции с тем же распределением не могут быть определены на пространстве с мерой, состоящем из меньшего числа атомов. Возможно ли существование жёстких ортонормированных систем с много меньшим числом "атомов"? В этом подпараграфе даётся положительный ответ на этот вопрос.

Обозначим через $D_N(0,1)$ множество кусочно–постоянных на $(0,1)$ функций с N интервалами постоянства $((j-1)/N, j/N)$, $j = 1, \dots, N$.

Утверждение 4.3. *Для любого $N \in \mathbb{N}$ существует ортонормированная система $\varphi_1, \dots, \varphi_N \in D_N(0,1)$ такая, что*

$$d_n(\{\varphi_1, \dots, \varphi_N\}, L_0(0,1)) \geqslant c, \quad \text{при } n \leqslant cN.$$

Лемма 4.1. *Пусть $\mathcal{E}$ — случайная $N \times N$ матрица с независимыми ± 1 элементами. Тогда*

$$\mathsf{P}\Big(\min_{\operatorname{rank} B \leqslant cN} \|\mathcal{E} - B\|_{L_0^{N\times N}} \leqslant c\Big) \leqslant \exp(-cN^2).$$

Доказательство. Применим утверждение 4.2 к $d = 2$ и $N \times N$ матрицам (теперь общее число элементов это N^2, а не N, как было в утверждении). В силу пункта 4), при $n \leqslant N/A$ для всех матриц $S \in \mathcal{S}$ выполнена требуемая оценка на аппроксимацию: $\|S - B\|_{L_0} \geqslant b$, если $\operatorname{rank} B < n$. При этом $|S| > 2^{N^2} - 2^{N^2(1-a)}$, то есть $\mathsf{P}(\mathcal{E} \notin \mathcal{S}) < 2^{-aN^2}$, что и требовалось. □

Утверждение 4.4. *Пусть $\varphi_1, \dots, \varphi_N$ — случайная система функций в $D_N(0,1)$, где $\varphi_{k,j} := \varphi_k|_{(\frac{j-1}{N}, \frac{j}{N})}$ — суть независимые случайные ± 1 величины. Тогда с вероятностью не менее $1 - 2\exp(-cN^2)$ мы имеем*

$$d_n(\{\varphi_1, \dots, \varphi_N\}, L_0) \geqslant c, \quad \text{если } n \leqslant cN.$$

Отметим, что, хотя для построения φ_k используются независимые случайные величины, сами функции $\{\varphi_k\}_1^N$ не образуют независимую систему.

Доказательство. Предположим, есть хорошая аппроксимация нашей системы в L_0 функциями из n-мерного подпространства: $\varphi_k \approx g_k$. Мы воспользуемся предыдущей леммой про матрицы, однако здесь есть небольшая техническая трудность: мы не можем сразу считать, что $g_k \in D_N$ (усреднить по отрезку нельзя — этот оператор не определен в L_0).

Проведём рассуждения в духе утверждения 1.12. Если $\|\varphi_k - g_k\|_{L_0} < c$ при всех k, то $\mu\{t\colon |\varphi_k(t) - g_k(t)| > c\} \leqslant c$. Рассмотрим векторы $\varphi(t) := (\varphi_1(t), \dots, \varphi_N(t))$ и $g(t) := (g_1(t), \dots, g_N(t))$. Положим $\Delta_c(u) := \#\{k\colon |u_k| > c\}$. Тогда

$$\int_0^1 \Delta_c(\varphi(t) - g(t))\,dt = \sum_{k=1}^N \mu\{t\colon |\varphi_k(t) - g_k(t)| > c\} \leqslant cN.$$

Векторы $g(t)$ лежат в некотором n-мерном подпространстве $\mathbb{R}^N$ при всех t. В каждом интервале $((j-1)/N, j/N)$ заменим $g(t)$ на какой-то вектор v^j из этого же подпространства, минимизирующий $\Delta_c(\varphi(t) - v^j)$. Полученная система $\widetilde{g}_1(t), \dots, \widetilde{g}_N(t)$ также лежит в n-мерном подпространстве, но ещё и в $D_N(0,1)$. При этом

$$\int_0^1 \Delta_c(\varphi(t) - \widetilde{g}(t))\,dt \leqslant \int_0^1 \Delta_c(\varphi(t) - g(t))\,dt \leqslant cN.$$

Значит, для матриц $\Phi = (\varphi_{k,j})$ и $\widetilde{G}$, образованных значениями систем $\{\varphi_k\}$ и $\{\widetilde{g}_k\}$ на интервалах $((j-1)/N, j/N)$, справедливо неравенство $\|\Phi - \widetilde{G}\|_{L_0^{N\times N}} \leqslant c$. А вероятность этого экспоненциально мала, в силу предыдущей леммы. □

Доказательство утверждения 4.3. Мы построим систему $\varphi_1, \dots, \varphi_N$ в пространстве $D_{2N}(0,2)$. Определим $\varphi_1, \dots, \varphi_N$ на $(0,1)$, используя случайные значения $\varphi_i|_{(\frac{j-1}{N}, \frac{j}{N})} = \pm\delta$, где δ выберем позже. С большой вероятностью система будет жёсткой в $L_0(0,1)$.

Продолжим систему на $(0,2)$, сделав её ортонормированной. Критерий существования такого продолжения хорошо известен (см., например, [43, Гл.8, §1]):

$$\max_{|a|=1} \|\sum_{k=1}^N a_k\varphi_k\|_{L_2(0,1)} \leqslant 1.$$

Если δ достаточно мало, то это условие выполнено с большой вероятностью, поскольку спектральная норма случайной ± 1 матрицы есть $O(N^{1/2})$, см., например, [151, Гл.4].

Ясно также, что можно считать функции $\varphi_k|_{(1,2)}$ лежащими в произвольном N-мерном подпространстве (важны только величины $\int_1^2 \varphi_k \varphi_l$), в частности, $D_N(1,2)$. $\square$

4.3 Приближение по словарю полилинейных функций

Пусть (X,μ) — пространство с мерой, $\mathcal{F}$ — некоторое множество суммируемых функций $X \to \mathbb{R}$, $\mathcal{A}$ — конечное семейство измеримых подмножеств X, и $\gamma > 0$. Скажем, что семейство $\mathcal{A}$ является γ-раздробленным классом функций $\mathcal{F}$ (по мере μ), если для любого выбора знаков $\sigma\colon \mathcal{A} \to \{-1,1\}$ найдётся функция $f \in \mathcal{F}$ такая, что

$$\sigma(A) \cdot \int_A f(x)\, d\mu \geqslant \gamma, \quad \forall A \in \mathcal{A}.$$

Это определение является интегральным аналогом комбинаторной размерности vc (см. §1.1).

Заметим, что если множества в семействе $\mathcal{A} = \{A_1, \dots, A_N\}$ дизъюнктны (точнее, $\mu(A_i \cap A_j) = 0$ при всех $i \neq j$), то оператор

$$\mathcal{I}_{\mathcal{A}}\colon L_1(X,\mu) \to \mathbb{R}^N, \quad f \mapsto (\int_{A_i} f\, d\mu)_{i=1}^N,$$

удовлетворяет неравенству

$$\|f\|_{L_1(X,\mu)} \geqslant \|\mathcal{I}_{\mathcal{A}} f\|_{\ell_1^N}. \tag{4.4}$$

При этом $\mathcal{I}_{\mathcal{A}}$ позволяет "дискретизовать" задачу, т.е. перейти от функций к векторам.

Рассмотрим пространство-произведение $X = X_1 \times \cdots \times X_d$ с мерой $\mu = \mu_1 \times \cdots \times \mu_d$. Полилинейные приближения в X — это приближения линейными комбинациями функций вида $u^1(x_1)u^2(x_2)\cdots u^d(x_d)$. Формально, определим величину наилучшего n-членного полилинейного приближения функции $f\colon X \to \mathbb{R}$ в $L_q(X,\mu)$:

$$\Theta_n(f, L_q) := \inf_{u^{i,k} \in L_q(X_i,\mu_i)} \|f - \sum_{k=1}^n u^{1,k}(x_1)\cdots u^{d,k}(x_d)\|_{L_q(X,\mu)} \tag{4.5}$$

и $\Theta_n(\mathcal{F}, L_q) := \sup_{f\in\mathcal{F}} \Theta_n(f, L_q)$.

Скажем, что семейство $\mathcal{A}$ имеет структуру типа $m_1 \times \ldots \times m_d$, если оно имеет вид

$$\mathcal{A} = \{A_{i_1}^{(1)} \times \cdots \times A_{i_d}^{(d)} \colon i_1 \in [m_1], \ldots, i_d \in [m_d]\}$$

для некоторых множеств $A_{i_k}^{(k)} \subset X_k$. Ясно, что $|\mathcal{A}| = m_1 \cdots m_d$.

Утверждение 4.5. *Предположим, что класс $\mathcal{F} \subset L_1(X, \mu)$ γ-раздробляет (по мере μ) дизъюнктное семейство $\mathcal{A}$ типа $m_1 \times \ldots \times m_d$. Тогда при $n \leqslant c(d)N/\sum m_j$, где $N := m_1 \cdots m_d$, имеет место неравенство*

$$\Theta_n(\mathcal{F}, L_1(X, \mu)) \geqslant c\gamma N.$$

Доказательство. Можно считать, что $N/\sum m_j$ достаточно велико.

Рассмотрим оператор $\mathcal{I}_\mathcal{A} \colon L_1(X, \mu) \to \mathbb{R}^N$. В силу наличия $m_1 \times \cdots \times m_d$-структуры, можно считать, что $\mathcal{I}_\mathcal{A}$ принимает значения в пространстве тензоров $\mathbb{R}^{m_1\times\ldots\times m_d}$. При этом он отображает функции вида $u^1(x_1)\ldots u^d(x_d)$ в тензоры ранга 1:

$$\mathcal{I}_\mathcal{A}(u^1(x_1)\cdots u^d(x_d))_{i_1,\ldots,i_d} = \int_{A_{i_1}^{(1)}\times\cdots\times A_{i_d}^{(d)}} u^1(x_1)\cdots u^d(x_d)\,d\mu = \prod_{j=1}^{d}\int_{A_{i_j}^{(j)}} u^j(x_j)\,d\mu_j.$$

В силу свойства раздробления, для любого сигнум-тензора $\sigma \in \mathbb{R}^{m_1\times\cdots\times m_d}$ найдётся функция $f_\sigma \in \mathcal{F}$ такая, что тензор $T(\sigma) := \mathcal{I}_\mathcal{A} f_\sigma$ имеет знаки $\operatorname{sign} T(\sigma) = \sigma$ и $\min_I |T(\sigma)_I| \geqslant \gamma$. В силу условия на n, выполнено неравенство

$$n + 1 \leqslant \frac{1}{A\log_2 d}\frac{N}{\sum m_j},$$

которое соответствует условию (4.2) утверждения 4.2. Возникает множество "плохо приближаемых" тензоров $\mathcal{S}$. Возьмём произвольный $\sigma^* \in \mathcal{S}$, применим неравенства (4.4) и (4.3):

$$\|f_{\sigma^*} - \sum_{k=1}^{n} u^{1,k}(x_1)\cdots u^{d,k}(x_d)\|_{L_1} \geqslant$$

$$\geqslant \|\mathcal{I}_\mathcal{A}(f_{\sigma^*} - \sum_{k=1}^{n} u^{1,k}(x_1)\cdots u^{d,k}(x_d))\|_{\ell_1^N} = \|T(\sigma^*) - T_n\|_{\ell_1^N} \geqslant bN\gamma.$$

□

Приведём теперь следствия для конкретных классов функций d переменных. Рассмотрим классическое пространство $C^r[0,1]^d$ функций $f\colon [0,1]^d \to \mathbb{R}$ гладкости r и единичный шар этого пространства:

$$U(C^r[0,1]^d) = \{f \in C^r[0,1]^d \colon \|f\|_{C^r} := \max_{\|\alpha\|_1 \leqslant r} \|D^\alpha f\|_\infty \leqslant 1\}.$$

Утверждение 4.6. *Пусть $r, d \in \mathbb{N}$, $1 \leqslant q \leqslant \infty$. Тогда справедливы неравенства*

$$c_1(r,d) n^{-\frac{r}{d-1}} \leqslant \Theta_n(U(C^r[0,1]^d), L_q) \leqslant c_2(r,d) n^{-\frac{r}{d-1}}.$$

Доказательство. Оценку снизу достаточно получить в метрике L_1. Возьмём некоторое число m и разделим куб $[0,1]^d$ на m^d кубиков:

$$Q_I = \prod_{k=1}^{d} \left[\frac{i_k - 1}{m}, \frac{i_k}{m}\right], \quad I = (i_1, \dots, i_d) \in [m]^d; \quad \mathcal{Q} := \{Q_I\}_{I \in [m]^d}.$$

Стандартным образом доказывается, что класс $U(C^r)$ раздробляет семейство кубиков $\mathcal{Q}$ с $\gamma = c(r,d) m^{-(r+d)}$. Действительно, возьмём "шапочку" — бесконечно гладкую функцию $\varphi(x)$ одной переменной с носителем в $(-\frac{1}{3}, \frac{1}{3})$ и $\int_{\mathbb{R}} \varphi(x)\,dx = 1$. Тогда $\max_{k \leqslant r} \|\varphi^{(k)}\|_\infty \leqslant c(r)$. Положим $\varphi^{\otimes d}(x_1, \dots, x_d) = \varphi(x_1) \cdots \varphi(x_d)$. Для любого выбора знаков $\sigma\colon [m]^d \to \{-1, 1\}$ мы берём функцию

$$f_\sigma(x) := \sum_{I \in [m]^d} \sigma(I) \varphi^{\otimes d}(m(x - x_I)),$$

где x_I — это центр куба Q_I. Поскольку носители слагаемых в определении f_σ не пересекаются, имеем

$$\|f_\sigma\|_{C^r} \leqslant c(r,d) m^r, \quad \int_{Q_I} f_\sigma(x)\,dx = \sigma(I) m^{-d}.$$

Раздробление доказано. Применяя утверждение 4.5, получаем нижнюю оценку $\Theta_n \gtrsim m^{-r}$ при $n \lesssim m^{d-1}$, что и требовалось.

Верняя оценка проводится в метрике L_∞. Хорошо известно, что по всякой функции $f \in C^r$ можно построить функцию g, равную многочлену (Тейлора) степени $< r$ на каждом кубике Q_I и такую, что $\|f - g\|_\infty \lesssim m^{-r}$. (См., например, [154, §8.4].) Запишем g в виде суммы $\sum_{\|\alpha\|_1 < r} c_\alpha x^\alpha$, где $c_\alpha = c_\alpha(I)$ для $x \in Q_I$. Для каждого $\alpha = (\alpha_1, \dots, \alpha_d)$ набор коэффициентов $c_\alpha(I)$ представляет собой тензор из $\mathbb{R}^{m \times m \times \dots \times m}$. Его ранг не превосходит m^{d-1}. Следовательно,

g представляется в виде суммы $O(m^{d-1})$ полилинейных функций вида

$$c^1_{i_1} c^2_{i_2} \cdots c^d_{i_d} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}, \quad \text{при } x \in Q_I.$$

Отсюда $\Theta_n \lesssim m^{-r}$ для $n \gtrsim m^{d-1}$. □

Для оценок более общих классов, в том числе нецелой гладкости, удобно использовать тригонометрические полиномы. Для $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$ обозначим через $\mathcal{T}_\mathbf{m}$ множество вещественных тригонометрических полиномов вида

$$t(x) = \sum_{k \in \mathbb{Z}^d:\ |k_j| \leqslant m_j} c_k e^{i(k_1 x_1 + \ldots + k_d x_d)}, \quad c_{-k} = \overline{c_k}.$$

Через B_∞ обозначаем единичный шар $L_\infty[-\pi, \pi]^d$.

Теорема 4.1. *Для любых $d \in \mathbb{N}$, $d \geqslant 2$, $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$, $N := m_1 \cdots m_d$ справедливо неравенство*

$$\Theta_n(\mathcal{T}_\mathbf{m} \cap B_\infty, L_1(\mathbb{T}^d)) \geqslant c_2(d), \quad \textit{при } n \leqslant c_1(d) \tfrac{N}{\sum_{j=1}^d m_j}. \tag{4.6}$$

Эта оценка является точной по порядку при фиксированном d, поскольку любой полином $t \in \mathcal{T}_\mathbf{m}$ может быть записан в виде суммы $\asymp N/\max\{m_j\}$ полилинейных функций, аналогично (4.1). Скажем, если $m_1 = \max\{m_j\}$:

$$t(x) = \sum_{(k_2, \dots, k_d)} \widetilde{t}_{k_2, \dots, k_d}(x_1) e^{ik_2 x_2} \cdots e^{ik_d x_d}.$$

Рассмотрим равномерную сетку ($\nu \in \mathbb{N}^d$)

$$\mathcal{X}_\nu := \{(\frac{\pi k_1}{\nu_1}, \dots, \frac{\pi k_d}{\nu_d}):\ -\nu_i < k_i \leqslant \nu_i\}$$

и пространство $\ell_p(\mathcal{X}_\nu)$ векторов значений функций на этой сетке. Теорема дискретизации Марцинкевича утверждает, что для достаточно плотной сетки интегральные и дискретные нормы полиномов эквивалентны: для любых $\mathbf{m} \in \mathbb{N}^d$, $N := \prod m_j$, $t \in \mathcal{T}_\mathbf{m}$ и $1 \leqslant p \leqslant \infty$ имеем (см. [139, §3.3]):

$$c(d) N^{-1/p} \|t\|_{\ell_p(\mathcal{X}_{2\mathbf{m}})} \leqslant \|t\|_p \leqslant C(d) N^{-1/p} \|t\|_{\ell_p(\mathcal{X}_{2\mathbf{m}})}. \tag{4.7}$$

Перейдём к доказательству теоремы 4.1.

Доказательство. Мы можем считать, что все m_j достаточно большие: $m_j \geqslant m(d)$. Действительно, для любой подпоследовательности $m_{i_1}, \dots, m_{i_k}$ в $\mathbf{m}$ имеем:

$$\Theta_n(\mathcal{T}_{\mathbf{m}} \cap B_\infty, L_1(\mathbb{T}^d)) \geqslant \Theta_n(\mathcal{T}_{(m_{i_1},\dots,m_{i_k})} \cap B_\infty, L_1(\mathbb{T}^k)),$$

поэтому можно убрать малые m_j.

Так же, как и в предыдущем доказательстве, мы построим семейство множеств, раздробленное нашим классом тригонометрических полиномов. Возьмём достаточно малые $\alpha, \delta > 0$. Рассмотрим сетку

$$\mathcal{X}^\alpha := \mathcal{X}_{\mathbf{m}'}, \quad \text{где } (m'_1, \dots, m'_d) := (\lfloor \alpha m_1 \rfloor, \dots, \lfloor \alpha m_d \rfloor)$$

и семейство параллелепипедов

$$Q_x^\delta := x + \delta \prod_{j=1}^d [-m_j^{-1}, m_j^{-1}], \quad x \in \mathcal{X}^\alpha.$$

Параллелепипеды не пересекаются, если δ достаточно мало. Мы докажем, что класс $\mathcal{T}_{\mathbf{m}}$ раздробляет семейство $\{Q_x^\delta\}_{x \in \mathcal{X}^\alpha}$, если α и δ достаточно малы.

Вместо "шапочки" φ мы используем ядро Фейера. Напомним свойства этого ядра (см. §1.1):

$$\frac{1}{m} K_m(0) = 1, \quad \frac{1}{m} |K_m(x)| \leqslant \min\left\{1, \frac{\pi^2}{m^2 x^2}\right\}, \ x \in [-\pi, \pi].$$

Положим

$$\psi(x) := \frac{K_{m_1}(x_1)}{m_1} \frac{K_{m_2}(x_2)}{m_2} \cdots \frac{K_{m_d}(x_d)}{m_d}.$$

Ясно, что $\psi \in \mathcal{T}_{\mathbf{m}}$. Набору знаков $\sigma \colon \mathcal{X}^\alpha \to \{-1, 1\}$ сопоставим тригонометрический полином

$$t_\sigma(y) := \sum_{x \in \mathcal{X}^\alpha} \sigma(x) \psi(y - x).$$

Оценим значения этого полинома на сетке $\mathcal{X}^\alpha$. Возьмём некоторый $x^\circ \in \mathcal{X}^\alpha$. Главный член в выражении для $t_\sigma(x^\circ)$ — это $\sigma(x^\circ)\psi(0) = \sigma(x^\circ)$. Вклад остальных слагаемых оценивается:

$$\Big| \sum_{x \in \mathcal{X}^\alpha \setminus \{x^\circ\}} \sigma(x) \psi(x^\circ - x) \Big| \leqslant \sum_{x \in \mathcal{X}^\alpha \setminus \{0\}} |\psi(x)| \leqslant$$

$$\leqslant \sum_{\substack{\mathbf{k}=(k_1,\dots,k_d) \neq 0 \\ |k_i| \leqslant m'_i}} \frac{1}{m_1} K_{m_1}\Big(\frac{\pi k_1}{m'_1}\Big) \cdots \frac{1}{m_d} K_{m_d}\Big(\frac{\pi k_d}{m'_d}\Big) \leqslant \sum_{\mathbf{k} \in \mathbb{Z}^d \setminus \{0\}} \prod_{j=1}^d \min\{1, \frac{c\alpha^2}{k_j^2}\}.$$

Покажем, что последнее выражение не превосходит $1/2$ для достаточно малых $\alpha = \alpha(d)$. Действительно, сумма по ненулевым k_j оценивается:

$$\sum_{\mathbf{k}\in(\mathbb{Z}\setminus\{0\})^d}\prod_{j=1}^{d}\min\{1,\frac{c\alpha^2}{k_j^2}\} = \sum_{\mathbf{k}\in(\mathbb{Z}\setminus\{0\})^d}\prod_{j=1}^{d}\frac{c\alpha^2}{k_j^2} = (c\alpha^2\cdot\pi^2/3)^d.$$

Суммы, где какие-то индексы k_i, $i\in I$, $I\subset[d]$, обращаются в ноль, сводятся к аналогичным суммам для меньшего d. Итак, $\sigma(x^\circ)t_\sigma(x^\circ)\geqslant 1/2$.

Аналогично можно показать, что t_σ ограничен величиной $C(d)$ на сетке $\mathcal{X}_{2\mathbf{m}}$; следовательно, по теореме Марцинкевича о дискретизации (4.7), мы получаем $\|t\|_\infty\leqslant C_1(d)$.

Из неравенство Бернштейна (1.5) при $p=\infty$ следует, что $\|\frac{\partial}{\partial x_j}t\|_\infty\lesssim m_j$. Отсюда мы видим, что $\sigma(x^\circ)t_\sigma(x)\geqslant 1/3$ при $x\in Q^\delta_{x^\circ}$, если $\delta=\delta(d)$ достаточно мало. Из сказанного следует, что семейство $\{Q^\delta_x\}_{x\in\mathcal{X}^\alpha}$ является γ-раздробленным классом $\mathcal{T}_\mathbf{m}\cap B_\infty$ с $\gamma\asymp\int_{Q^\delta_{x^\circ}}1\asymp N^{-1}$. Остаётся воспользоваться утверждением 4.5. □

Определим теперь классы Соболева смешанной гладкости $W_p^\mathbf{r}(\mathbb{T}^d)$, где $d\in\mathbb{N}$, $\mathbf{r}=(r_1,\dots,r_d)$, $r_j>0$. Рассмотрим одномерное ядро Бернулли

$$F_r(x) := 1+\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{e^{ikx}}{(ik)^r} = 1+2\sum_{k=1}^{\infty}k^{-r}\cos(kx-r\pi/2)$$

и соответствующий многомерный аналог $F_\mathbf{r}(x):=\prod_{j=1}^d F_{r_j}(x_j)$. Класс $W_p^\mathbf{r}(\mathbb{T}^d)$ определяется (см. [98, §3.1]) как множество свёрток

$$f(x)=(\varphi*F_\mathbf{r})(x)=\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d}F_\mathbf{r}(x)\varphi(y-x)\,dy,\quad \|\varphi\|_p\leqslant 1.$$

Отметим, что в монографии [73] класс определяется свёртками с приведённым ядром Бернулли, $\prod_{j=1}^d(F_{r_j}(x)-1)$; в результате получается подмножество пространства L_p^0 функций с нулевыми средними: $\int_\mathbb{T}f(x)\,dx_j\equiv 0$. Это отличие не влияет на аппроксимативные свойства, см. обсуждение в [98, §3.2].

Следствие 4.1. *Для любых $d\in\mathbb{N}$, $d\geqslant 2$, $r>0$, $\mathbf{r}:=(r,\dots,r)$, $2\leqslant p\leqslant\infty$, $1\leqslant q\leqslant 2$, $n\in\mathbb{N}$ справедливы неравенства*

$$c(r,d)n^{-\frac{rd}{d-1}}\leqslant\Theta_n(W_p^\mathbf{r},L_q(\mathbb{T}^d))\leqslant C(r,d)n^{-\frac{rd}{d-1}}.$$

Доказательство. Верхняя оценка была получена Темляковым [75, Теорема 4.1]:

$$\Theta_n(W_2^{\mathbf{r}}, L_2(\mathbb{T}^d)) \lesssim n^{-\frac{rd}{d-1}}.$$

Нижнюю оценку достаточно установить для $p = \infty$, $q = 1$. Как и в [88], используем включение $n^{-rd}\mathcal{T}(n, n, \dots, n)_\infty \subset c(r, d)W_\infty^{\mathbf{r}}$, справедливое в силу аналога неравенства Бернштейна для дробной гладкости, см. [139, §3.3]. Неравенство (4.6) сразу даёт нужную нам оценку:

$$\Theta_n(W_\infty^{\mathbf{r}}, L_1(\mathbb{T}^d)) \gtrsim n^{-\frac{rd}{d-1}}.$$

□

Обобщение на случай групп переменных. Пусть переменные разбиты на группы (блоки):

$$[d] = \Delta_1 \sqcup \Delta_2 \sqcup \ldots \sqcup \Delta_b, \quad |\Delta_j| \geqslant 1. \tag{4.8}$$

Для $x \in \mathbb{R}^d$ через $x[j]$ обозначим вектор из координат j-й группы, то есть $x[j] := (x_i)_{i\in\Delta_j}$.

Рассмотрим следующее обобщение полилинейной аппроксимации. Вместо функций $u^1(x_1)\cdots u^d(x_d)$ будем использовать произведение функций, каждая из которых зависит от всех переменных соответствующей группы: $u^1(x[1])\cdots u^b(x[b])$. Следующее утверждение является обобщением теоремы 4.1.

Утверждение 4.7. *Пусть $d \in \mathbb{N}$, $d \geqslant 2$ и переменные разбиты на группы (4.8). Для любого $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$ существует тригонометрический полином $t \in \mathcal{T}_{\mathbf{m}} \cap B_\infty$ такой, что для всякого*

$$n \leqslant c_1(d)\frac{\prod_{i=1}^d m_i}{\sum_{j=1}^b \prod_{i\in\Delta_j} m_i}$$

и любых функций $u^{j,k} \in L_1(\mathbb{T}^{\Delta_j})$, $j = 1, \dots, b$, $k = 1, \dots, n$, справедливо неравенство

$$\left\| t(x) - \sum_{k=1}^n u^{1,k}(x[1])\cdots u^{b,k}(x[b]) \right\|_{L_1(\mathbb{T}^d)} \geqslant c_2(d).$$

Доказательство. Работают те же рассуждения, что и в теореме 4.1. Для дискретизации используем в точности тот же оператор $\mathcal{I}\colon L(\mathbb{T}^d) \to \mathbb{R}^{2m_1'\times\cdots\times 2m_d'}$

$(m'_j \asymp m_j)$, что и в теореме. Вместо аппроксимации тензорами ранга n возникает аппроксимация тензорами следующего вида:

$$T_I \approx \sum_{k=1}^{n} u^{1,k}_{I[1]} \cdots u^{b,k}_{I[b]},$$

где $I = (i_1, \ldots, i_d)$, $I[j] = (i_s \colon s \in \Delta_j)$. Здесь также применим метод Шапиро–Уоррена, только количество параметров больше:

$$K = n \sum_{j=1}^{b} \prod_{i \in \Delta_j} 2m'_i.$$

□

Нетрудно видеть, что оценка на n точна по порядку.

Глава 5

Избранные задачи о поперечниках в $\mathbb{R}^N$

5.1 Поперечники шаров в смешанных нормах

Рассмотрим пространство $\mathbb{R}^N$; предположим, что все координаты разбиты на b блоков $\Delta_1, \dots, \Delta_b$ размера s каждый; тем самым, $N = sb$. Для вектора $x = (x_i) \in \mathbb{R}^N$ обозначаем через $x[j] \in \mathbb{R}^s$ вектор, образованный координатами из j-го блока, т.е.

$$x[j] := (x_i)_{i \in \Delta_j}, \quad \Delta_j := \{s(j-1)+1, \dots, sj\}.$$

Напомним понятие смешанной нормы. Пусть заданы некоторые нормы Φ и Ψ на $\mathbb{R}^s$ и $\mathbb{R}^b$, соответственно. Тогда смешанная норма вектора $x \in \mathbb{R}^N$ определена как

$$\|x\| := \Psi(z), \quad z := (\Phi(x[1]), \dots, \Phi(x[b])). \tag{5.1}$$

Напомним, что норма называется симметричной, если её единичный шар — безусловное множество, инвариантное относительно перестановок координат. Из результатов главы 1 вытекает, что шар смеси симметричных норм — жёсткий в ℓ_q^N при $q \leqslant 2$.

Утверждение 5.1. *Пусть $s, b \in \mathbb{N}$, $N = sb$; Φ и Ψ — некоторые симметричные нормы на $\mathbb{R}^s$ и $\mathbb{R}^b$, соответственно. Пусть $B \subset \mathbb{R}^N$ — единичный шар смешанной нормы* (5.1)*. Тогда для любых $1 \leqslant q \leqslant 2$, $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$, имеем*

$$d_n(B, \ell_q^N) \geqslant c(q,\varepsilon) \sup_{x \in B} \|x\|_q = c(q,\varepsilon) \sup_{\Phi(y) \leqslant 1} \|y\|_q \sup_{\Psi(z) \leqslant 1} \|z\|_q.$$

Доказательство. Запишем координаты $\mathbb{R}^N$ в виде пар (i,j), где $i \in [s]$ — координата внутри блока, а $j \in [b]$ — номер блока. На множестве координат действует группа $S_s \times S_b$, переставляющая координаты внутри блоков и переставляющая сами блоки:

$$(\sigma,\tau)\colon (i,j) \mapsto (\sigma(i),\tau(j)).$$

Очевидно, что это действие транзитивно. Применима теорема 1.3, которая и даёт жёсткость. □

Важный частный случай смешанных норм — смесь $\ell_{q_1}^s$ и $\ell_{q_2}^b$:

$$\|x\|_{q_1,q_2} := (\sum_{j=1}^{b}(\sum_{s(j-1)<i\leqslant sj}|x_i|^{q_1})^{q_2/q_1})^{1/q_2}$$

(с очевидными изменениями для $q_1 = \infty$ или $q_2 = \infty$). Полученное пространство обозначается $\ell_{q_1,q_2}^{s,b}$, а $B_{q_1,q_2}^{s,b}$ — его единичный шар. В этой главе мы изучим поперечники

$$d_n(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}).$$

Параметров много; запомнить, какой к чему относится, можно так: индекс 1 у внутренних норм, они применяются *первыми*; индекс 2 у внешних норм. как и в других главах, p задаёт приближаемое тело, а q задаёт метрику. Наконец, s — это размер блока, ***s**ize*; b — это количество блоков, ***b**locks*.

Мы будем применять двойственность поперечников по Колмогорову и Гельфанду, (1.22); в случае смешанных норм она даёт равенство

$$d_n(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}) = d^n(B_{q_1',q_2'}^{s,b}, \ell_{p_1',p_2'}^{s,b}).$$

Мы уже видели, что $d_0(B_p^N, \ell_q^N) = N^{(1/q-1/p)_+}$. Для смешанных норм имеем

$$d_0(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}) = s^{(\frac{1}{q_1}-\frac{1}{p_1})_+} b^{(\frac{1}{q_2}-\frac{1}{p_2})_+}.$$

Частным случаем утверждения 5.1 является

Следствие 5.1. *Пусть* $1 \leqslant p_1, p_2 \leqslant \infty$; $1 \leqslant q \leqslant 2$; $N = sb$. *Тогда для* $n \leqslant N(1-\varepsilon)$, $\varepsilon \in (0,1)$, *справедливо неравенство*

$$c(q,\varepsilon) s^{(\frac{1}{q}-\frac{1}{p_1})_+} b^{(\frac{1}{q}-\frac{1}{p_2})_+} \leqslant d_n(B_{p_1,p_2}^{s,b}, \ell_q^N) \leqslant s^{(\frac{1}{q}-\frac{1}{p_1})_+} b^{(\frac{1}{q}-\frac{1}{p_2})_+}.$$

Произведение октаэдров в $\ell_{2,1}$.

Теорема 5.1. *Для всех $s, b \in \mathbb{N}$, $\varepsilon \in (0,1)$, $N := sb$ имеет место неравенство*

$$d_n(B_{1,\infty}^{s,b}, \ell_{2,1}^{s,b}) \geqslant c(\varepsilon)b, \quad \text{при } n \leqslant N(1-\varepsilon).$$

Идея доказательства состоит в переходе к поперечнику по Гельфанду, построению векторов для разных блоков и взятии их линейной комбинации. Для оценки норм построенного вектора нам потребуется один результат о "балансировке".

Пусть γ_d — гауссова мера в $\mathbb{R}^d$:

$$\gamma_d(A) = (2\pi)^{-d/2} \int_A \exp(-|x|^2/2)\, dx, \quad A \subset \mathbb{R}^d.$$

Положим $\Psi(t) := \gamma_1(-t, t)$. Нам понадобится простое неравенство

$$\Psi(t) \geqslant 1 - e^{-t^2/2} \geqslant \exp(-2\exp(-t^2/2)), \quad t \geqslant 1, \tag{5.2}$$

которое вытекает из оценки (1.27) и численного неравенства $1 - u \geqslant e^{-2u}$, $0 \leqslant u \leqslant 3/4$.

Теорема 5.A (S-неравенство)**.** *Пусть $K \subset \mathbb{R}^d$ — выпуклое замкнутое центрально–симметричное тело, $P = \{x : |\langle w, x\rangle| \leqslant 1\}$, $w \in \mathbb{R}^d$, — полоса той же гауссовской меры, $\gamma_d(K) = \gamma_d(P)$. Тогда выполнены неравенство $\gamma_d(tK) \geqslant \gamma_d(tP)$ при $t \geqslant 1$ и неравенство $\gamma_d(tK) \leqslant \gamma_d(tP)$ при $t \in (0,1)$.*

Доказательство получено Р. Латала и К. Олешкевичем [114].

Теорема 5.B (Gaussian Correlation Conjecture)**.** *Для любых выпуклых центрально–симметричных тел $K_1, K_2 \subset \mathbb{R}^d$ выполнено неравенство $\gamma_d(K_1 \cap K_2) \geqslant \gamma_d(K_1)\gamma_d(K_2)$.*

Теорема 5.B была доказана в работе Т. Ройена [129]; для наших целей достаточно частного случая этого результата, когда K_1 — эллипсоид (доказано в [108]).

Кроме того, нам понадобится следующий результат о "балансировке" векторов, представляющий собой частный случай [104, Lemma 3.1].

Лемма 5.A. *Для любого выпуклого центрально-симметричного тела $V \subset \mathbb{R}^d$ такого, что $\gamma_d(V) \geqslant 2^{-d/7}$, и любых векторов $u_1, \ldots, u_d \in B_2^d$ найдётся вектор $\tau \in \{0, 1, -1\}^d$, $\|\tau\|_1 \geqslant d/2$ такой, что $\sum_{i=1}^d \tau_i u_i \in 4V$.*

Сформулируем утверждение, которое непосредственно будет использоваться в доказательстве теоремы.

Лемма 5.1. *Пусть $\mathcal{X}_1, \dots, \mathcal{X}_m$ — конечные наборы векторов в $\mathbb{R}^d$, причём $\sum_{x\in\mathcal{X}_j} |x|^2 \leqslant 1$ для всех j. Тогда найдётся вектор $\tau \in \{0, 1, -1\}^d$, $\|\tau\|_1 \geqslant cd$, такой, что*

$$\max_{j=1,\dots,m} \left(\sum_{x\in\mathcal{X}_j} \langle \tau, x\rangle^2 \right)^{1/2} \leqslant C \log^{1/2} \left(\frac{m}{d} + 2 \right).$$

Лемма обобщает Теорему 3 из работы Глускина [28], в которой рассматривались одноэлементные множества $\mathcal{X}_j = \{x_j\}$.

Доказательство леммы. Положим

$$h_j(v) := (\sum_{x\in\mathcal{X}_j} \langle v, x\rangle^2)^{1/2}, \quad \mathcal{E}_j := \{v \in \mathbb{R}^d \colon h_j(v) \leqslant 1\}.$$

Для стандартного гауссовского вектора ξ в $\mathbb{R}^d$ имеем

$$\mathsf{E} h_j(\xi)^2 = \sum_{x\in\mathcal{X}_j} |x|^2 \leqslant 1,$$

откуда получаем оценку

$$\gamma_d(2\mathcal{E}_j) = 1 - \mathsf{P}(h_j(\xi) > 2) \geqslant 1 - \frac{\mathsf{E} h_j(\xi)^2}{4} \geqslant 3/4 > \Psi(1) = 0.68\dots. \tag{5.3}$$

В силу S–неравенства, используя (5.3) и (5.2), получаем

$$\gamma_d(2t\mathcal{E}_j) \geqslant \Psi(t) \geqslant \exp(-2\exp(-t^2/2)), \quad t \geqslant 1.$$

Пусть $V := \bigcap_{j=1}^m (2t\mathcal{E}_j)$; используя Gaussian Correlation Conjecture, получаем

$$\gamma_d(V) \geqslant \prod_{j=1}^m \gamma_d(2t\mathcal{E}_j) \geqslant \exp(-2m\exp(-t^2/2)).$$

При $t \asymp \log^{1/2}(m/d + 2)$ мы получим $\gamma_d(V) \geqslant 2^{-d/7}$. Следовательно, к телу V применима лемма 5.A, которая для стандартного базиса $e_1, \dots, e_d$ даёт вектор $\sum \tau_j e_j = \tau \in 4V$. Таким образом, $h_j(\tau) \leqslant 8t$ для всех j, что и требовалось. $\square$

Доказательство теоремы. Считаем без ограничения общности, что s и b достаточно велики (иначе поперечник для смешанных норм оценивается через поперечник для обычных норм).

В доказательстве мы обозначаем i-ю координату вектора через $x(i)$; нижний индекс зарезервирован для номера вектора. Как и раньше, $x[m] \in \mathbb{R}^s$ это вектор с координатами из m-го блока Δ_m.

В силу двойственности (1.22) и равенства $d_n(B_1^N, \ell_2^N) = d^n(B_2^N, \ell_\infty^N) = \sqrt{1-n/N}$, для любого подпространства $Z \subset \mathbb{R}^N$ имеем:

$$\dim Z \geqslant \varepsilon N/2 \quad \Rightarrow \quad \exists x \in Z \ \ |x| = 1,\ \|x\|_\infty \geqslant \delta := (\varepsilon/2)^{1/2}. \tag{5.4}$$

В силу той же двойственности,

$$d_n(B_{1,\infty}^{s,b}, \ell_{2,1}^{s,b}) = d^n(B_{2,\infty}^{s,b}, \ell_{\infty,1}^{s,b}).$$

Для оценки поперечника по Гельфанду снизу нужно в любом пространстве $Q \subset \mathbb{R}^N$ размерности $\geqslant \varepsilon N$ найти вектор x, для которого $\|x\|_{\infty,1} \gtrsim b$, $\|x\|_{2,\infty} \lesssim 1$. Эти условия будут выполнены, если

- ℓ_2-норма каждого блока ограничена: $|x[m]| \lesssim 1$;
- найдется $\asymp b$ блоков m таких, что $|x(i_m)| \gtrsim 1$ для некоторого $i_m \in \Delta_m$.

Искомый вектор будет построен как линейная комбинация векторов $x_1, \dots, x_l \in Q$, где вектор $x_j \in Z_j$ и подпространство $Z_j \subset Z_{j-1} \subset Q$ строятся на j-м шаге описываемой ниже конструкции. На первом шаге полагаем $Z_1 := Q$ и выбираем x_1 произвольным образом, применив (5.4). Имеем $x_1 \in Q$, $|x_1| = 1$ и $\|x_1\|_\infty \geqslant \delta$. Следовательно, $|x_1(i_1^*)| \geqslant \delta$ для некоторой координаты i_1^*; эту координату будем называть "большой"; если таких координат несколько, выберем какую-то одну. Большую координату, возникшую на j-м шаге, обозначим через i_j^*.

Построенные на первых j шагах числа $(x_k(i))_{\substack{1\leqslant k\leqslant j\\ 1\leqslant i\leqslant N}}$ удобно представлять себе в виде матрицы, где k будет номером строки, а i — номером столбца.

После j-го шага мы составляем множество "зануляемых" координат $\Lambda_j \subset \{1, \dots, N\}$. В него входят:

(i) координаты i, для которых набралась большая сумма квадратов: $x_1(i)^2 + \dots + x_j(i)^2 \geqslant 1/s$;

(ii) координаты из всех блоков m, в которых встречалась "большая" координата: $i_k^* \in \Delta_m$, $1 \leqslant k \leqslant j$;

(iii) координаты из всех блоков m, в которых набралась большая сумма квадратов: $|x_1[m]|^2 + \ldots + |x_j[m]|^2 \geqslant 1$.

На $(j+1)$-м шаге мы применяем (5.4) к подпространству $Z_{j+1} := \{x \in Q\colon x(i) = 0\ \forall i \in \Lambda_j\}$ и получаем вектор x_{j+1}.

Утверждение (5.4) применимо, если $|\Lambda_j| \leqslant \varepsilon N/2$. Проверим, что это условие выполняется при $j < (\varepsilon/6)b$. Действительно, на каждом шаге появляется один блок с большой координатой, следовательно, так занулим не более $js < \varepsilon N/6$ координат. Далее, так как сумма (по всем $k \leqslant j$ и всем $i \leqslant N$) квадратов всех построенных чисел $x_k(i)$ не превосходит j, то найдётся не более $j/(1/s) = js < \varepsilon N/6$ столбцов с суммой квадратов $\geqslant 1/s$. Аналогично оценивается количество зануляемых координат из блоков с большой суммой квадратов.

Итак, мы сделаем $l \asymp b$ шагов и построим векторы $x_1, \ldots, x_l$. Все числа $x_j(i)$ можно разделить на три группы:

- "большие": $x_j(i_j^*)$;
- "средние": те $x_j(i)$, для которых $\sum_{k \leqslant j} x_k(i)^2 \geqslant 1/s$;
- "малые": все остальные.

Отметим, что по построению, если $x_j(i)$ большое или среднее, то на всех следующих шагах i-я координата зануляется: $x_{j+1}(i) = \ldots = x_l(i) = 0$. Представим x_j в виде суммы $v_j + w_j$, где в v_j поместим большие и средние числа, а в w_j — малые.

Мы будем строить искомый вектор x в виде

$$x = \sum_{j=1}^{l} \tau_j x_j = \sum_{j=1}^{l} \tau_j v_j + \sum_{j=1}^{l} \tau_j w_j = v + w,$$

где $\tau = (\tau_j)_{j=1}^l \in \{0, 1, -1\}^l$ будет взят из леммы 5.1. Очевидно, $x \in Q$. Оценим сверху $|x[m]| \leqslant |v[m]| + |w[m]|$. Поскольку для каждого i в столбце $(v_j(i))_{j=1}^l$ лишь одно ненулевое число, то

$$|v[m]|^2 = \sum_{i \in \Delta_m} v(i)^2 = \sum_{i \in \Delta_m} (\sum_{j=1}^{l} \tau_j v_j(i))^2 \leqslant \sum_{i \in \Delta_m} \sum_{j=1}^{l} v_j(i)^2 = \sum_{j=1}^{l} |v_j[m]|^2 \lesssim 1.$$

Последнее неравенство выполнено, так как, по построению в силу (iii), справедливо даже $\sum_{j=1}^{l} |x_j[m]|^2 \leqslant 2$, а $|v_j[m]| \leqslant |x_j[m]|$.

Для завершения доказательства остаётся найти вектор τ, $\|\tau\|_1 \asymp b$ такой, что 1) ℓ_2-норма w в каждом блоке ограничена константой 2) $|w(i)| \leqslant \delta/2$ для всех $i \in \{i_1^*, \dots, i_l^*\}$. Условие 1) обеспечит $|x[m]| \lesssim 1$, а условие 2) даст для всех j с $|\tau_j| = 1$ оценку

$$|x(i_j^*)| \geqslant |v(i_j^*)| - |w(i_j^*)| \geqslant |v(i_j^*)| - \delta/2 = |v_j(i_j^*)| - \delta/2 \geqslant \delta/2.$$

Отсюда $\|x\|_{\infty,1} \geqslant (\delta/2)\|\tau\|_1$.

Рассмотрим столбцы $f_i := (w_j(i))_{j=1}^{l} \in \mathbb{R}^l$; имеем $|f_i| \leqslant 1/\sqrt{s}$ в силу того, что "средние" числа не попадают в w. Воспользуемся леммой 5.1 для следующих наборов векторов: составим для каждого блока $m = 1, \dots, b$ наборы $\{f_i \colon i \in \Delta_m\}$ и добавим к ним наборы по одному вектору $\{f_i/|f_i|\}$ для $i \in \{i_1^*, \dots, i_l^*\}$. Лемма выдаст нам вектор τ, для которого

$$|w[m]|^2 = \sum_{i \in \Delta_m} \langle \tau, f_i \rangle^2 \lesssim \log(2 + (b+l)/l) \lesssim 1,$$

а также $|w(i)| = |\langle \tau, f_i \rangle| \lesssim |f_i| \leqslant 1/\sqrt{s} < \delta/2$ для $i \in \{i_1^*, \dots, i_l^*\}$ при больших s. Что и требовалось. □

Следствие для функциональных классов. Поперечник произведения октаэдров возник у Галеева в связи с задачей о линейных поперечниках классов Гёльдера–Никольского. Дадим необходимые определения.

Линейным n-поперечником множества K в нормированном пространстве X называется величина $\lambda_n(K, X) := \inf_{L_n} \sup_{x \in K} \|x - L_n x\|$, где инфимум берётся по линейным операторам $L_n \colon X \to X$ ранга не более n. Очевидно, что $\lambda_n(K, X) \geqslant d_n(K, X)$.

Определим класс Гёльдера–Никольского $H_p^{\mathbf{r}}(\mathbb{T}^d)$, где $d \in \mathbb{N}$, $\mathbf{r} = (r_1, \dots, r_d)$, $r_j > 0$, $1 \leqslant p \leqslant \infty$. Этот класс состоит из функций $f \in L_p(\mathbb{T}^d)$ таких, что $\int_{\mathbb{T}} f(x)\, dx_j \equiv 0$, $j = 1, \dots, d$, и для всех $\mathbf{h} = (h_1, \dots, h_d)$:

$$\|\Delta_{\mathbf{h}}^l f\|_p \leqslant \prod_{j=1}^{d} |h_j|^{r_j}, \quad l := \lfloor \max r_j \rfloor + 1,$$

где $\Delta_{\mathbf{h}}^l$ — смешанная l-я разность с шагом h_j по переменной x_j.

Подробнее о классах смешанной гладкости см. [73, 98], а также [15, 17].

Следствие 5.2. *Пусть* $\mathbf{r} = (r, r, \dots, r)$, $r > 0$. *Справедливы порядковые равенства (константа в* $\asymp$ *может зависеть от параметров* r, d, p, q*):*

$$\lambda_n(H_p^{\mathbf{r}}(\mathbb{T}^d), L_q) \asymp \begin{cases} \left(\frac{\log^{d-1} n}{n}\right)^{r-1/2+1/q} \log^{(d-1)/q} n, & \frac{1}{p} + \frac{1}{q} < 1,\ p \leqslant 2,\ r > 1 - \frac{1}{q}, \\ \left(\frac{\log^{d-1} n}{n}\right)^{r-1/p+1/q} \log^{(d-1)/q} n, & 2 \leqslant p < q,\ r > \frac{1}{p} - \frac{1}{q}. \end{cases}$$

Доказательство. Верхние оценки поперечника были получены в работе [17], см. пп.5), 6) основной теоремы. Нижняя оценка сводилась там к поперечнику в смешанной норме:

$$\lambda_n(H_p^{\mathbf{r}}, L_q) \gtrsim b^{1/q-1} 2^{m(-r+1/2-1/q)} \lambda_n(B_{1,\infty}^{2^m,b}, \ell_{2,1}^{2^m,b}), \quad b \asymp m^{d-1},\ b2^m \asymp n.$$

Галеев далее оценивал $\lambda_n \geqslant d_n$, использовал неравенство (0.16) и, как следствие, в оценке возникал лишний повторный логарифм n. Если применить теорему 5.1, получим точную по порядку нижнюю оценку. □

Критерий жёсткости для смешанных норм.

Теорема 5.2. *Пусть* $1 \leqslant p_1, p_2, q_1, q_2 \leqslant \infty$; $s, b \in \mathbb{N}$ *и* $N := sb$.

A) Предположим, что набор (p_1, p_2, q_1, q_2) *удовлетворяет трём условиям: (i)* $q_1 \leqslant \max\{p_1, 2\}$; *(ii)* $q_2 \leqslant \max\{p_2, 2\}$; *(iii) он не попадает в "исключительный случай"*

$$q_1 < \min\{p_1, q_2\} \text{ и } p_2 < q_2 \leqslant 2. \tag{5.5}$$

Тогда для любого $\varepsilon > 0$ *справедливо неравенство*

$$d_n(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}) \geqslant c(q_2, \varepsilon) d_0(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}), \quad \text{при } n \leqslant N(1-\varepsilon).$$

B) Для любого набора (p_1, p_2, q_1, q_2), *не удовлетворяющего A), существует* $\gamma > 0$ *такое, что для достаточно больших* $t := \min(s, b)$ *имеем*

$$d_n(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}) \leqslant t^{-\gamma} d_0(B_{p_1,p_2}^{s,b}, \ell_{q_1,q_2}^{s,b}), \quad \text{если } n \geqslant Nt^{-\gamma}.$$

Проиллюстрируем "исключительный случай" (5.5) примером: $d_n(B_{\infty,1}^{s,b}, \ell_{1,2}^{s,b})$. Будем смотреть на вектор x как на матрицу размера $s \times b$; блок $x[j]$ — это j-й столбец x. Ясно, что крайние точки $B_{\infty,1}^{s,b}$ — это матрицы с носителем в одном из столбцов (и элементами ± 1). Пусть $s = b$, через x^T обозначим транспонированную матрицу x. Для любой крайней точки x мы имеем

$$\|x - (x - x^T)\|_{1,2} = \|x^T\|_{1,2} \leqslant s^{1/2},$$

поскольку в x^T в каждом столбце не больше одного ненулевого элемента. Следовательно, x аппроксимируется кососимметричной матрицей $x - x^T$ (а такие матрицы лежат в подпространстве размерности $s(s-1)/2$) с погрешностью не более $s^{1/2}$. Отсюда

$$d_{s(s-1)/2}(B^{s,s}_{\infty,1}, \ell^{s,s}_{1,2}) \leqslant s^{1/2},$$

в то время как $d_0(B^{s,s}_{\infty,1}, \ell^{s,s}_{1,2}) = s$.

Мы видим, что смесь двух жёстких множеств может не быть жёсткой в смешанной норме.

Оставшаяся часть параграфа посвящена доказательству теоремы 5.2.

Доказательство. A) Нижние оценки мы выведем из следствия 5.1 и теоремы 5.1.

Каждое из условий (i), (ii) распадается на два непересекающихся случая: $p_i \geqslant q_i$ или $p_i < q_i \leqslant 2$; мы рассмотрим их по отдельности.

Случай a) $p_1 \geqslant q_1$, $p_2 \geqslant q_2$. Имеем $d_0 = s^{(1/q_1 - 1/p_1)} b^{(1/q_2 - 1/p_2)}$. В силу включения $B^{s,b}_{p_1,p_2} \supset s^{-1/p_1} b^{-1/p_2} B^N_\infty$, достаточно доказать жёсткость для куба, $p_1 = p_2 = \infty$. Если $q_1 \leqslant q_2$, то из (1.12) мы получим

$$d_n(B^N_\infty, \ell^{s,b}_{q_1,q_2}) \geqslant b^{1/q_2 - 1/q_1} d_n(B^N_\infty, \ell^{s,b}_{q_1,q_1}) \geqslant b^{1/q_2 - 1/q_1} \varepsilon N^{1/q_1} = \varepsilon s^{1/q_1} b^{1/q_2}.$$

Если $q_1 > q_2$, рассуждения аналогичны.

Случай b) $p_1 < q_1 \leqslant 2$, $p_2 < q_2 \leqslant 2$; тогда $d_0 = 1$. Имеем

$$d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \geqslant d_n(B^N_1, \ell^N_2) = (1 - n/N)^{1/2} \geqslant \varepsilon^{1/2}.$$

Отметим, что это рассуждение работает и для $p_1 = q_1 \leqslant 2$ или $p_2 = q_2 \leqslant 2$.

Случай c) $p_1 < q_1 \leqslant 2$, $p_2 \geqslant q_2$. Имеем $d_0 = b^{1/q_2 - 1/p_2}$. Заменим пару (p_1, q_1) на $(1, 2)$ (это не изменит d_0 и не увеличит d_n). Далее, используя включение $B^{s,b}_{1,p_2} \supset b^{-1/p_2} B^{s,b}_{1,\infty}$ и оценку $\|x\|_{2,1} \leqslant b^{1-1/q_2} \|x\|_{2,q_2}$, сводим дело к теореме 5.1:

$$d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \geqslant d_n(B^{s,b}_{1,p_2}, \ell^{s,b}_{2,q_2}) \geqslant b^{-1/p_2} d_n(B^{s,b}_{1,\infty}, \ell^{s,b}_{2,q_2}) \geqslant$$
$$\geqslant b^{-1/p_2} b^{-1+1/q_2} d_n(B^{s,b}_{1,\infty}, \ell^{s,b}_{2,1}) \geqslant c(\varepsilon) b^{1/q_2 - 1/p_2}.$$

Случай d) $p_1 \geqslant q_1$, $p_2 < q_2 \leqslant 2$; $d_0 = s^{1/q_1 - 1/p_1}$. Рассмотрим несколько подслучаев.

Случай d1) $q_1 \geqslant q_2$. Используем следствие 5.1:

$$d_n(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \geqslant s^{1/q_1 - 1/q_2} d_n(B^{s,b}_{p_1,p_2}, \ell^N_{q_2}) \geqslant c(q_2, \varepsilon) s^{1/q_1 - 1/p_1}.$$

Случай d2) $p_1 = q_1 \leqslant 2$. Здесь работают те же рассуждения, что и в b).

Если параметры не попадают в подслучаи d1), d2), это означает, что $q_1 < q_2 \leqslant 2$, $q_1 < p_1$, $p_2 < q_2 \leqslant 2$. То есть мы попали в исключительный случай.

B) Верхние оценки. Докажем, что жёсткость не имеет место для всех четвёрок (p_1, p_2, q_1, q_2), не попавших в A). Необходимым условием для жёсткости шара в смешанной норме является жёсткость $B^s_{p_1}$ в $\ell^s_{q_1}$ и $B^b_{p_2}$ в $\ell^b_{q_2}$. Как мы обсуждали во введении (см. (0.17) и (0.18)), шар B^N_p является жёстким в ℓ^N_q в точности при $q \leqslant \max\{p, 2\}$. Следовательно, условия (i) и (ii) необходимы для жёсткости, и тут нам достаточно рассматривать только исключительный случай: $p_1 > q_1$, $p_2 < q_2 \leqslant 2$, $q_1 < q_2$. Отметим, что $d_0 = s^{1/q_1 - 1/p_1}$.

Наша конструкция представляет собой обобщение примера $d_n(B^{s,s}_{\infty,1}, \ell^{s,s}_{2,1})$. Вместо транспонирования мы строим более сложный оператор, который "распределяет" координаты вектора по разным столбцам так, чтобы минимизировать смешанную норму.

Потребуются некоторые приготовления. Мы отождествляем векторы из $\mathbb{R}^N$, $N = sb$, с матрицами размера $s \times b$ и будем использовать далее оба термина (матрица и вектор); j-й блок, $x[j]$, отождествляем с j-м столбцом матрицы.

Пусть $\Gamma = (\Gamma_1, \dots, \Gamma_m)$ — некоторое разбиение множества $[s] \times [b]$:

$$[s] \times [b] = \Gamma_1 \sqcup \Gamma_2 \sqcup \dots \sqcup \Gamma_m.$$

Назовём разбиение Γ (m, r, l)-разбиением, если выполняются три условия:

(I) $|\Gamma_k| \leqslant r$, $k = 1, \dots, m$;

(II) любой столбец $\mathcal{C}^j := [s] \times \{j\}$ содержит не более одной точки из каждого из множеств: $|\mathcal{C}^j \cap \Gamma_k| \leqslant 1$, $j = 1, \dots, b$, $k = 1, \dots, m$;

(III) любая пара столбцов может пересекать одновременно не более l множеств:

$$|\{k \colon \mathcal{C}^{j_1} \cap \Gamma_k \neq \varnothing,\ \mathcal{C}^{j_2} \cap \Gamma_k \neq \varnothing\}| \leqslant l,$$

для $1 \leqslant j_1 < j_2 \leqslant b$.

Отметим, что если существует (m, r, l)-разбиение множества $[s'] \times [b']$, то для любых $s \leqslant s'$, $b \leqslant b'$ существует (m, r, l)-разбиение $[s] \times [b]$. Действительно, разбиение большего множества индуцирует разбиение меньшего.

Лемма 5.2. *Предположим, $A_1, \ldots, A_m \subset [b]$ — множества мощности не больше r, удовлетворяющие условиям: 1) каждая точка $j \in [b]$ принадлежит как минимум s множествам A_k; 2) каждая пара $\{j_1, j_2\} \subset [b]$, $j_1 \neq j_2$, содержится не более чем в l множествах A_k. Тогда существует (m, r, l)-разбиение $[s] \times [b]$.*

Доказательство. Построим $\Gamma_1, \ldots, \Gamma_m$; начинаем с пустых множеств. Для каждого $j \in [b]$ возьмём какие-то s множеств $A_{k_1}, \ldots, A_{k_s}$, содержащих j. Добавим точку $(1, j)$ в Γ_{k_1}, точку $(2, j)$ в Γ_{k_2} и так далее (s, j) в Γ_{k_s}. Свойство (II) не будет нарушено. Если мы проделаем это для всех j, то заполним весь прямоугольник $[s] \times [b]$. Заметим, что $\mathcal{C}^j \cap \Gamma_k \neq \varnothing$, только если $j \in A_k$; это влечёт (I) и (III), что и требовалось. (При этом некоторые из Γ_k могут остаться пустыми.) □

Для построения регулярных систем множеств удобно использовать комбинаторные дизайны. Необходимые определения и базовые факты о дизайнах можно найти в книге [135, Гл.1, 5].

Напомним, что (b, r, l)-дизайном называется последовательность множеств $A_1, \ldots, A_m \subset [b]$ таких, что $|A_k| = r$ для всех k и что любая пара $\{j_1, j_2\} \subset [b]$, $j_1 \neq j_2$, содержится ровно в l множествах A_k. Хорошо известно, что каждая точка $j \in [b]$ лежит ровно в $l(b-1)/(r-1)$ множествах дизайна; кроме того, $m = lb(b-1)/(r^2 - r)$. Дизайн с параметрами (b, r, l) даёт (b, r, lh)-дизайн ($h \in \mathbb{N}$): достаточно повторить каждое множество h раз. Если r — степень простого числа и $d \in \mathbb{N}$, $d \geqslant 2$, то существует $(r^d, r, 1)$-дизайн; он состоит из аффинных прямых в пространстве $\mathbb{F}_r^d$.

Лемма 5.3. *Для любых $s, b, d \geqslant 2$, $s \geqslant b$, существует (m, r, l)-разбиение множества $[s] \times [b]$ с параметрами*

$$b^{1/d} \leqslant r \leqslant 2b^{1/d}, \quad m \leqslant C_1(d)sb/r, \quad l \leqslant C_2(d)sr/b.$$

Доказательство. Предположим сначала, что $b = 2^{ud}$, $u \in \mathbb{N}$. Пусть $r = 2^u$, тогда $b = r^d$ и существует $(b, r, 1)$-дизайн. Каждая точка $[b]$ лежит в $(b-1)/(r-1)$ множествах дизайна. Пусть l — минимальное целое число, удовлетворяющее неравенству $l(b-1)/(r-1) \geqslant s$. Возьмём (b, r, l)-дизайн, полученный повторением каждого множества l раз. Заметим, что он удовлетворяет условиям

леммы 5.2 и даёт (m, r, l)-разбиение с параметрами

$$r = b^{1/d}, \quad l = \left\lceil \frac{s(r-1)}{b-1} \right\rceil, \quad m = \frac{lb(b-1)}{r(r-1)}.$$

Требуемые неравенства выполнены.

Общий случай следует из рассмотренного, поскольку всегда $2^{-d}b' < b \leqslant b'$, $b' = 2^{ud}$, при подходящем u. □

Каждое (m, r, l)-разбиение $[s] \times [b] = \Gamma_1 \sqcup \ldots \sqcup \Gamma_m$ определяет m-мерное пространство $Q_\Gamma \subset \mathbb{R}^{s\times b}$ векторов, постоянных на множествах Γ_k. Рассмотрим линейный оператор $D_\Gamma \colon \mathbb{R}^{s\times b} \to Q_\Gamma$:

$$D_\Gamma(e_{i,j}) = \sum_{(u,v)\in\Gamma_k} e_{u,v}, \quad \text{где } \Gamma_k \ni (i,j).$$

Лемма 5.4. *Пусть Γ — (m, r, l)-разбиение $[s] \times [b]$; тогда для любой матрицы $x \in \mathbb{R}^{s\times b}$ с носителем в одном столбце и любых $p, q_1, q_2 \in [1, +\infty]$ имеем*

$$\|x - D_\Gamma x\|_{q_1,q_2} \leqslant l^{(1/q_1-1/p)_+} b^{(1/q_2-1/p)_+} (r-1)^{1/p} \|x\|_p.$$

Доказательство. Обозначим $z := x - D_\Gamma x$; ключевое свойство z состоит в том, что в каждом столбце у него не более l ненулевых элементов (это следует из свойства (III) разбиения). Следовательно, $\|z[j]\|_{q_1} \leqslant l^{(1/q_1-1/p)_+} \|z[j]\|_p$ и

$$\|z\|_{q_1,q_2} \leqslant b^{(1/q_2-1/p)_+} \|z\|_{q_1,p} \leqslant b^{(1/q_2-1/p)_+} l^{(1/q_1-1/p)_+} \|z\|_p.$$

Каждая ненулевая координата x встречается среди координат z не более чем $(r-1)$ раз; следовательно, $\|z\|_p^p \leqslant (r-1)\|x\|_p^p$, что завершает доказательство. □

Вернёмся к доказательству теоремы. Сначала рассмотрим случай $s \geqslant b$. Возьмём число d (уточним этот выбор позже) и применим леммы 5.3 и 5.4. Для любой матрицы x с носителем в одном столбце мы имеем

$$\begin{aligned}\|x - D_\Gamma x\|_{q_1,q_2} &\leqslant l^{(1/q_1-1/p_1)_+} b^{(1/q_2-1/p_1)_+} r^{1/p_1} \|x\|_{p_1} \leqslant \\ &\leqslant C(d)(sr/b)^{1/q_1-1/p_1} b^{(1/q_2-1/p_1)_+} r^{1/p_1} \|x\|_{p_1} = C(d) s^{1/q_1-1/p_1} b^{-\alpha} r^{1/q_1} \|x\|_{p_1},\end{aligned}$$

где $\alpha := (\frac{1}{q_1} - \frac{1}{p_1}) - (\frac{1}{q_2} - \frac{1}{p_1})_+$. Заметим, что поскольку $q_1 < q_2$ и $q_1 < p_1$, то $\alpha > 0$, поэтому

$$\|x - D_\Gamma x\|_{q_1,q_2} \leqslant C(d) s^{1/q_1-1/p_1} b^{-\alpha/2} \|x\|_{p_1}, \tag{5.6}$$

если число $d = d(p_1, p_2, q_1, q_2)$ выбрано достаточно большим.

Для $y \in \mathbb{R}^b$ и $k \leqslant b$ рассмотрим k–членное приближение

$$\sigma_k(y)_q := \min\{\|y - z\|_q \colon \#\operatorname{supp} z \leqslant k\}.$$

Будем использовать неравенство (см., например, [139], Lemma 7.6.6):

$$\sigma_{k-1}(y)_q \leqslant k^{-(1/p-1/q)}\|y\|_p, \quad 1 \leqslant p < q \leqslant \infty. \tag{5.7}$$

Возьмём теперь произвольный $x \in B^{s,b}_{p_1,p_2}$. Тогда вектор $y = (\|x[1]\|_{p_1}, \ldots, \|x[b]\|_{p_1})$ имеет норму $\|y\|_{p_2} \leqslant 1$. Следовательно,

$$\sigma_{k-1}(y)_{q_2} \leqslant k^{-(1/p_2-1/q_2)}\|y\|_{p_2} \leqslant k^{-(1/p_2-1/q_2)} =: \delta,$$

где число k будет выбрано позже. Итак, есть множество индексов Λ, $|\Lambda| < k$ такое, что $\|y^\Lambda - y\|_{q_2} \leqslant \delta$, где $y^\Lambda \in \mathbb{R}^b$ совпадает с y на Λ и зануляется вне Λ. Следовательно, $\|x - x^\Lambda\|_{p_1,q_2} \leqslant \delta$, где $x^\Lambda = \sum_{j\in\Lambda} x^j$, а $x^j \in \mathbb{R}^N$ — это матрица, у которой j-й столбец равен $x[j]$, а на остальных местах нули. В соответствии с неравенством (5.6),

$$\|x^\Lambda - D_\Gamma x^\Lambda\|_{q_1,q_2} \leqslant \sum_{j\in\Lambda} \|x^j - D_\Gamma x^j\|_{q_1,q_2} \leqslant C(d)s^{1/q_1-1/p_1}b^{-\alpha/2}\sum_{j\in\Lambda}\|x^j\|_{p_1} \leqslant$$
$$\leqslant C(d)ks^{1/q_1-1/p_1}b^{-\alpha/2}.$$

Следовательно, вектор x приближен вектором $D_\Gamma x^\Lambda \in Q_\Gamma$ с погрешностью

$$\|x - D_\Gamma x^\Lambda\|_{q_1,q_2} \leqslant \|x - x^\Lambda\|_{q_1,q_2} + \|x^\Lambda - D_\Gamma x^\Lambda\|_{q_1,q_2} \leqslant \delta s^{1/q_1-1/p_1} + C(d)ks^{1/q_1-1/p_1}b^{-\alpha/2}.$$

Откуда ($\dim Q_\Gamma \leqslant m$):

$$d_m(B^{s,b}_{p_1,p_2}, \ell^{s,b}_{q_1,q_2}) \leqslant \delta s^{1/q_1-1/p_1} + C(d)ks^{1/q_1-1/p_1}b^{-\alpha/2}.$$

Отметим, что $m \leqslant C(d)sb/r \leqslant C(d)Nb^{-1/d}$. Выбор $k \asymp b^{\alpha/4}$ обеспечивает требуемую верхнюю оценку поперечника.

Наконец, нужно рассмотреть случай $s < b$. Сгруппируем все блоки в $g \asymp b/s$ групп размера примерно s:

$$x = x^{(1)} + \ldots + x^{(g)}, \quad x^{(j)} \in \ell^{s,b_j}_{q_1,q_2}, \quad s/2 \leqslant b_j \leqslant s.$$

В каждой группе возьмём подпространство размерности $\asymp s^{2-\gamma}$, дающее хорошую аппроксимацию:

$$\|x^{(j)} - y^{(j)}\|_{q_1,q_2} \lesssim s^{-\gamma}\|x^{(j)}\|_{p_1,p_2}.$$

Тогда сумма этих подпространств имеет размерность $\asymp s^{-\gamma}N$ и обеспечивает нужную нам аппроксимацию во всём пространстве (здесь используем, что $p_2 < q_2$):

$$\|x-y\|_{q_1,q_2} = (\sum_{j=1}^{g} \|x^{(j)} - y^{(j)}\|_{q_1,q_2}^{q_2})^{1/q_2} \lesssim s^{-\gamma}(\sum_{j=1}^{g} \|x^{(j)}\|_{p_1,p_2}^{q_2})^{1/q_2} \leqslant$$
$$\leqslant s^{-\gamma}(\sum_{j=1}^{g} \|x^{(j)}\|_{p_1,p_2}^{p_2})^{1/p_2} = s^{-\gamma}\|x\|_{p_1,p_2}.$$

Теорема доказана. □

5.2 Свойства матриц, приближающих единичную

С матрицей $A \in \mathbb{R}^{N\times N}$ свяжем функцию распределения

$$F_A^*(\gamma) := N^{-2}|\{(i,j)\colon |A_{i,j}| > \gamma\}|.$$

Теорема 5.3. *Пусть A — матрица размера $N \times N$, ранга n такая, что*

$$\max_{1\leqslant i,j\leqslant N} |A_{i,j} - \delta_{i,j}| \leqslant 1/3. \tag{5.8}$$

Тогда справедливо неравенство

$$F_A^*(\gamma) \geqslant c\frac{\log N}{n\log(2+\frac{n}{\log N})}, \quad \text{где } \gamma := c\max\{n^{-3/2}\log N,\, n^{-1}\}.$$

Замечание. Вместо 1/3 можно взять любую константу, меньшую 1/2.

Следствие 5.3. *В условиях теоремы, если $n \leqslant K\log N$ для некоторого K, то не менее $c(K)N^2$ элементов матрицы A по модулю больше $c(K)n^{-1/2}$.*

Это следствие даёт положительный ответ на вопрос Кашина. Вероятностный метод показывает, что порядок $n^{-1/2}$ нельзя увеличить (см. (5.10)).

Следствие 5.4. *В условиях теоремы, доля ненулевых элементов матрицы A не меньше $c\log(N)/(n\log(2+n/\log N))$.*

При этом нетрудно построить пример матрицы (блочно-диагональной), приближающей единичную, для которой доля ненулевых элементов $\asymp \log(N)/n$.

Общие оценки для матриц можно применять в частных случаях. Так, из следствия 5.3 получаем следующее

Следствие 5.5. *Если $N \geqslant \exp(\delta n)$, то для любых единичных векторов $x^1, \dots, x^N \in S^{n-1}$ таких, что $\max_{i\neq j} |\langle x^i, x^j\rangle| \leqslant 1/3$, по крайней мере $c(\delta)N^2$ пар удовлетворяют неравенству $|\langle x^i, x^j\rangle| > c(\delta)n^{-1/2}$.*

Действительно, достаточно применить следствие 5.3 к матрице Грама системы $\{x^k\}$.

Мы будем использовать следующие утверждения.

Лемма 5.B ([103])**.** *Пусть $K \subset \mathbb{R}^n$ — выпуклое центрально-симметричное тело, $D \supset K$ — его эллипсоид минимального объёма, $|\cdot|_D$ — евклидова норма, соответствующая D. Для любого $\varepsilon \in (0,1)$ найдутся векторы $x^1, \dots, x^k \in K \cap \partial D$, $k \geqslant n(1-\varepsilon)$ такие, что для любого набора коэффициентов $t_1, \dots, t_k$ справедливо неравенство*

$$|\sum_{\nu=1}^{k} t_\nu x^\nu|_D \geqslant c_0\varepsilon n^{-1/2}\sum_{\nu=1}^{k}|t_\nu|. \tag{5.9}$$

Эта лемма была доказана в работе [137] со множителем $\varepsilon^{3/2}n^{-1/2}$; в [103] результат улучшен до $\varepsilon n^{-1/2}$. Лемма использовалась для оценки расстояния Банаха–Мазура от произвольного n-мерного пространства до ℓ_∞^n. Так, в [103] было доказано, что $d(X_n, \ell_\infty^n) \leqslant Cn^{5/6}$.

Лемма 5.5. *Пусть $K \subset \mathbb{R}^n$ — выпуклое центрально-симметричное тело. Существуют $x^1, \dots, x^n \in \operatorname{extr} K$ такие, что любой элемент $x \in K$ представляется в виде $x = \sum_{\nu=1}^n t_\nu x^\nu$ с некоторыми $t_1, \dots, t_n \in [-1, 1]$.*

Искомый набор $x^1, \dots, x^n$ — это просто базис Ауэрбаха в нормированном пространстве с единичным шаром K. Базис строится максимизацией объёма [116, Vol.I, §1.c], и можно все векторы выбирать из $\operatorname{extr} K$.

Перейдём к доказательству теоремы.

Доказательство. Элементы $|A_{i,j}| > \gamma$ будем называть "большими", а долю больших элементов будем называть "плотностью".

Матрица A имеет плотность $F_A^*(\gamma)$. Пусть максимальная плотность столбцов A равна $\varkappa$. Можно считать, что $\varkappa = O(F_A^*(\gamma))$ и оценивать снизу $\varkappa$. Действительно, столбцов с плотностью $> 2F_A^*(\gamma)$ не больше половины; столбцы и строки с этими индексами можно вычеркнуть, перейдя к главной подматрице половинного размера.

Пусть $v^1, \dots, v^N$ — столбцы матрицы A; $V := \operatorname{span}\{v^j\}_{j=1}^N$ — n-мерное подпространство $\mathbb{R}^N$; $K := \operatorname{conv}\{\pm v^j\}_{j=1}^N$; $D \subset V$ — эллипсоид минимального объёма, содержащий K; $\|\cdot\|_K$ — норма в V с единичным шаром K; $\langle\cdot,\cdot\rangle_D$ и $|\cdot|_D$ — скалярное произведение и норма в V, соответствующие эллипсоиду D.

Зафиксируем $\varepsilon \in (0,1)$; точное значение выберем позже. В соответствии с леммой 5.B, найдутся векторы $x^1, \dots, x^k \in K \cap \partial D$, $k \geqslant n(1-\varepsilon)$ такие, что выполняется неравенство (5.9). Поскольку $K \cap \partial D \subset \operatorname{extr} K \subset \{\pm v^j\}_{j=1}^N$, то, без ограничения общности, $x^1 = v^{j_1}, \dots, x^k = v^{j_k}$ для некоторых $j_1, \dots, j_k$. (Мы воспользуемся этим чуть позже.)

Дополним векторы $\{x^\nu\}_1^k$ до базиса пространства V векторами $y^1, \dots, y^{n-k}$, ортогональными $\{x^\nu\}_1^k$ в смысле скалярного произведения $\langle\cdot,\cdot\rangle_D$. Разложим все столбцы матрицы по этому базису:

$$v^j = \sum_{\nu=1}^{k} t_\nu^j x^\nu + \sum_{l=1}^{n-k} s_l^j y^l, \quad j = 1, \dots, N.$$

При этом, поскольку $v^j \in K$, то мы имеем

$$1 \geqslant \|v^j\|_K \geqslant |v^j|_D \geqslant |\sum_{\nu=1}^{k} t_\nu^j x^\nu|_D \geqslant c_0 \varepsilon n^{-1/2} \sum_{\nu=1}^{k} |t_\nu^j|.$$

Таким образом,

$$A_{i,j} = (v^j)_i = \sum_{\nu=1}^{k} t_\nu^j x_i^\nu + \sum_{l=1}^{n-k} s_l^j y_i^l = \langle x_i, t^j\rangle + \langle y_i, s^j\rangle, \quad \|t^j\|_1 \leqslant (c_0\varepsilon)^{-1} n^{1/2},$$

где $x_i, t^j \in \mathbb{R}^k$; $y_i, s^j \in \mathbb{R}^{n-k}$.

Векторы $x_i = (x_i^1, \dots, x_i^k) = (v_i^{j_1}, \dots, v_i^{j_k})$ — это строки матрицы, составленной из столбцов A с индексами $j_1, \dots, j_k$. Эта матрица имеет плотность $\leqslant \varkappa$. Следовательно, по крайней мере половина строк имеет плотность $\leqslant 2\varkappa$; обозначим через I множество индексов этих строк. Итак, плотность x_i не больше $2\varkappa$ при $i \in I$.

Представим x_i в виде $x_i = w_i + z_i$, где в w_i мы оставили большие координаты, а в z_i — все остальные. Имеем

$$|\langle x_i, t^j\rangle - \langle w_i, t^j\rangle| \leqslant \|t^j\|_1 \cdot \gamma \leqslant (c_0\varepsilon)^{-1} n^{1/2}\gamma \leqslant 1/15$$

при условии $\gamma \leqslant c_0\varepsilon/(15n^{1/2})$. Рассмотрим следующую матрицу:

$$B_{i,j} = \langle w_i, t^j\rangle + \langle y_i, s^j\rangle, \quad i, j \in I.$$

Будем использовать три свойства: количество ненулевых элементов в каждом векторе w_i не превосходит $m := \lfloor 2\varkappa k\rfloor$; размерность векторов y_i, s^j не превосходит $n\varepsilon$; матрица B хорошо приближает единичную матрицу: $|B_{i,j} - \delta_{i,j}| \leqslant 2/5$.

Рассмотрим строки матрицы B. С одной стороны, они попарно удалены на расстояние не менее $1/5$ в ℓ_∞^I. С другой стороны, можно построить $(1/11)$–сеть в ℓ_∞^I для множества строк следующим образом. Для каждого $i \in I$ множество $\Lambda = \operatorname{supp} w_i$ имеет мощность не более m и в силу (1.35) количество вариантов для Λ не превосходит $(en/m)^m$. (Не исключён пока случай $m = 0$; в этом случае, формально, положим $(en/m)^m = 1$.)

Для фиксированного множества Λ рассмотрим подматрицу B_Λ матрицы B, состоящую из всех строк i, для которых $\operatorname{supp} w_i = \Lambda$. Для этой подматрицы $B_{i,j} = \langle w_i|_\Lambda, t^j\big|_\Lambda\rangle + \langle y_i, s^j\rangle$, откуда $\operatorname{rank} B_\Lambda \leqslant m + n\varepsilon$. Все строки B_Λ лежат в шаре фиксированного радиуса в пространстве размерности не более $m + n\varepsilon$; размер минимальной сети для такого шара не превосходит $\exp(C(m + n\varepsilon))$.

Поскольку размер сети не меньше количества строк B, имеем:

$$(en/m)^m \cdot \exp(C(m + n\varepsilon)) \geqslant |I| \geqslant N/2.$$

Логарифмируя, получим: $m\ln(en/m) + Cm + Cn\varepsilon \geqslant \ln(N/2)$. Положим

$$\varepsilon := c_1 \frac{\ln(N/2)}{n},$$

где $c_1 \leqslant 1/(2C)$, тогда $m\ln(en/m) + Cm = m\ln(C_1 n/m) \geqslant \ln(N/2)/2$. Отсюда, в частности, следует, что $m \geqslant 1$.

Свойство аппроксимации (5.8) гарантирует, что $n \geqslant c\log N$ (см. обсуждение во введении). Поэтому за счёт малого c_1 можно считать, что $\varepsilon < 1/2$; тогда $k \geqslant n(1-\varepsilon) \geqslant n/2$; $2\varkappa n \geqslant m = \lfloor 2\varkappa k\rfloor \geqslant \varkappa k \geqslant \varkappa n/2$, следовательно:

$$\varkappa \ln(2C_1/\varkappa) \geqslant \frac{\ln(N/2)}{4n}.$$

Отсюда вытекает доказываемая оценка для $\varkappa$ при $\gamma \asymp n^{-1/2}\varepsilon \asymp n^{-3/2}\log N$.

Для $\gamma \asymp n^{-1}$ используем лемму 5.5; рассуждения аналогичны (и даже проще: $\varepsilon = 0$ и $k = n$). Теорема доказана. □

Наметим ещё один подход к задаче оценки количества больших элементов матрицы. Он даёт оценки чуть хуже оптимальных, но применим для широкого диапазона γ.

Нам потребуется вариант теоремы Турана. Рассматриваются обычные графы: неориентированные, без петель и кратных рёбер.

Теорема 5.C ([82], Ch.41)**.** *Если граф G, имеющий m вершин, не содержит полного подграфа на r вершинах ($r \geqslant 2$), то количество рёбер G не превосходит $(1 - 1/(r-1))m^2/2$.*

Утверждение 5.2. *Для достаточно больших N существует матрица A размера $N \times N$ и ранга не выше $C\log N$, для которой выполнено условие* (5.8)*, а также справедливо неравенство*

$$F_A^*(\gamma) \leqslant \exp(-c_1\gamma^2 \log N), \quad \text{при } C\log^{-1/2} N \leqslant \gamma \leqslant c_2. \tag{5.10}$$

При фиксированном $K > 0$ и достаточно большом N для любой $N \times N$ матрицы A ранга не выше $K\log N$ из условия (5.8) *следует неравенство*

$$F_A^*(\gamma) \geqslant \exp(-c(K)\gamma^2 \log(1/\gamma)\log N), \quad C(K)\log^{-1/2} N \leqslant \gamma \leqslant c(K).$$

Доказательство. Положим $n := \lceil 100 \ln N \rceil$. Для построения матрицы используем вероятностный метод: $A_{i,j} := \frac{1}{n}\langle x^i, x^j\rangle$, $x^1, \ldots, x^N \sim \mathrm{Unif}\{-1,1\}^n$. Ясно, что $A_{i,i} = 1$, а при $i \neq j$ величина $\langle x^i, x^j\rangle$ распределена как сумма n независимых ± 1. При фиксированном γ из неравенства Хёфдинга (1.23) получаем оценку

$$p_\gamma := \mathsf{P}(|A_{i,j}| > \gamma) \leqslant \hat{p}_\gamma := 2\exp(-n\gamma^2/2).$$

При больших N имеем $N^2\hat{p}_{1/3} < 1/3$, поэтому с вероятностью $\geqslant 2/3$ условие (5.8) выполняется.

Запишем функцию распределения в следующем виде:

$$F_A^*(\gamma) = N^{-2}\sum_{i,j=1}^{N} \mathbf{1}\{|A_{i,j}| > \gamma\} = N^{-1} + N^{-2}\sum_{i\neq j} \mathbf{1}\{|A_{i,j}| > \gamma\} =: N^{-1} + F(\gamma).$$

При фиксированном γ величина $F(\gamma)$ является случайной величиной, причём $\mathsf{E}F(\gamma) = (1-1/N)p_\gamma < \hat{p}_\gamma$. В силу того же неравенства Хёфдинга:

$$\mathsf{P}(F(\gamma) \geqslant 2\hat{p}_\gamma) \leqslant \mathsf{P}(|F(\gamma) - \mathsf{E}F(\gamma)| \geqslant \hat{p}_\gamma) \leqslant 2\exp(-2N^2\hat{p}_\gamma^2).$$

При $\gamma \leqslant c_2$ имеем $\hat{p}_\gamma \gtrsim N^{-1/2}$, откуда $\mathsf{P}(F(\gamma) \geqslant 2\hat{p}_\gamma) \leqslant 2\exp(-cN)$.

Достаточно доказать (5.10) для $\gamma = 2^{-\nu}$, $\nu \in \mathbb{N}$, $\nu_{\min} \leqslant \nu \leqslant \nu_{\max}$, где $\nu_{\min} \asymp 1$, $\nu_{\max} \asymp \log\log N$. Поскольку $\nu_{\max} \cdot 2\exp(-cN) \leqslant 1/3$ при больших N, то с вероятностью $\geqslant 2/3$ выполняются неравенства $F(\gamma) \leqslant 2\hat{p}_\gamma$, $\gamma = 2^{-\nu}$, $\nu_{\min} \leqslant \nu \leqslant \nu_{\max}$. Следовательно, с положительной вероятностью матрица A удовлетворяет условию (5.8) и неравенствам

$$F_A^*(\gamma) \leqslant N^{-1} + 4\exp(-n\gamma^2/2), \quad C\log^{-1/2} N \leqslant \gamma \leqslant c_2.$$

Отсюда (при подходящих константах) вытекает требуемая оценка.

Более интересна оценка снизу. Пусть $n = \operatorname{rank} A \leqslant K\log N$ и выполнено условие (5.8). Построим симметричную матрицу B с единицами на диагонали следующим образом. Сначала разделим строки матрицы на числа $A_{i,i} \in [2/3, 4/3]$, получим матрицу $\tilde{A}$, потом симметризуем: $B := (\tilde{A} + \tilde{A}^T)/2$. При этом $\operatorname{rank} B \leqslant 2\operatorname{rank} A$ и

$$F_A^*(2\gamma/3) \geqslant F_{\tilde{A}}^*(\gamma) \geqslant \frac{1}{2}F_B^*(\gamma).$$

Следовательно, достаточно оценивать функцию распределения матрицы B.

Предположим, число M таково, что выполняется неравенство

$$d_n(B_1^M, \ell_\infty^M) > \gamma. \tag{5.11}$$

Рассмотрим граф G на N вершинах, в котором i, j соединены, если $|B_{i,j}| \leqslant \gamma$. Тогда G не содержит полного подграфа на M вершинах, иначе мы могли бы перейти к соответствующей $M \times M$ подматрице и она бы, в силу условия $B_{i,i} = 1$, приближала единичную матрицу с ошибкой $\leqslant \gamma$; это противоречит (5.11). По теореме 5.C, количество рёбер в G не превосходит $(1-(M-1)^{-1})N^2/2$. Значит, количество пар $\{i,j\}$, для которых $|B_{i,j}| > \gamma$, не меньше

$$\binom{N}{2} - \left(1 - \frac{1}{M-1}\right)\frac{N^2}{2} = \frac{N^2}{2}\left(\frac{1}{M-1} - \frac{1}{N}\right). \tag{5.12}$$

Из оценки Глускина (1.13) видно, что в (5.11) можно взять M, для которого $\log(M)/n \asymp \gamma^2\log(1/\gamma)$. При $\gamma \leqslant c(K)$ имеем $M < N/2$, и оценка (5.12)

даёт $\asymp N^2/M$ больших элементов матрицы; подставляя значение M, получаем требуемую оценку. $\square$

Отметим, что в доказательстве нижней оценки из условия (5.8) мы использовали только хорошее приближение единичной матрицы на диагонали. Можно ещё ослабить требование к матрице A, например, до следующего условия: $|A_{i,i}| \geqslant 1/2$, $i = 1, \dots, N$.

5.3 "Свидетели" поперечника

Теорема 5.4. *При всех $1 \leqslant p, q \leqslant \infty$, за исключением случая $1 < p < 2$, $q = \infty$, для любого $N \in \mathbb{N}$ существует множество $W \subset B_p^N$ мощности $|W| \leqslant CN \log N$ такое, что*

$$d_n(W, \ell_q^N) \geqslant c(p,q) d_n(B_p^N, \ell_q^N) \quad \text{для всех } 1 \leqslant n \leqslant cN.$$

Доказательство. При $p = 1$ доказывать нечего, берём $W := \{e_1, \dots, e_N\}$ для любого q. К октаэдру сводится и случай $p < q \leqslant 2$.

Случай $p \geqslant q$, $d_n \asymp N^{1/q-1/p}$ сводится к поперечнику куба. Действительно, если $W \subset B_\infty^N$ и $d_n(W, \ell_q^N) \gtrsim N^{1/q}$, то $N^{-1/p} W \subset B_p^N$ и $d_n(N^{-1/p}W, \ell_q^N) \gtrsim N^{1/q-1/p}$. Известно (см. §4.2 и утверждение 4.2), что существует матрица A размера $N \times N$ с элементами ± 1 такая, что для любой матрицы B ранга $\leqslant cN$ имеем $A_{i,j} B_{i,j} \leqslant 0$ для $\asymp N^2$ элементов. Ясно, что в качестве W можно взять строки матрицы A.

Остаётся рассмотреть случай $q > \max\{p, 2\}$.

Пусть $q < \infty$. Порядок поперечника даётся неравенствами (1.16). Рассмотрим сначала случай $p < 2 < q < \infty$. Нужно будет, в соответствии с оценкой поперечника, доказать два неравенства: $d_n(W, \ell_q^N) \gtrsim \min\{1, N^{1/q} n^{-1/2}\}$ и $d_n(W, \ell_q^N) \gtrsim N^{1/q-1/p}$. Возьмём в качестве W объединение вершин октаэдра — это обеспечивает первое неравенство, в силу (1.15), — и нормированного множества "свидетелей" поперечника $d_n(B_\infty^N, \ell_q^N)$ — это даёт второе неравенство.

При $2 \leqslant p < q < \infty$ нижние оценки получаются вписыванием в шар B_p^N множества $k^{-1/p} V_k^N$ с подходящим k. Напомним, что $V_k^N = \operatorname{conv}\{\pm e_{i_1} \pm \dots \pm e_{i_k}\}$. Мы покажем, что V_k^N можно заменить на его небольшое подмножество.

Лемма 5.6. *Пусть $q \in (2,\infty)$, N и k — это степени двойки, $k \leqslant N$. Существует множество $W \subset \operatorname{extr} V_k^N$ из N точек такое, что $d_n(W, \ell_q^N) \geqslant (1/2)k^{1/q}$ при $n \leqslant (1/16)N^{2/q}k^{1-2/q}$.*

Доказательство. Разделим N координат на блоки размера k. Возьмём матрицу Уолша–Адамара размера $k \times k$. В каждом блоке разместим k векторов, равных строкам этой матрицы (и с нулевыми координатами вне блока), получим множество W из N векторов. Пусть $\xi \sim \operatorname{Unif} W$. Ясно, что $\mathsf{E}\xi_i\xi_j = 0$ и $\mathsf{E}\xi_i^2 = k/N$, поэтому к вектору ξ применимо утверждение 1.10, которое и даёт нужную оценку поперечника (аналогично следствию 1.4). □

Продолжим доказательство. Можно считать, что N — это степень двойки. В качестве W возьмём объединение множеств $k^{-1/p}W(k)$ по всем k степеням двойки, где $W(k)$ получается из предыдущей леммы. При этом считаем, что $W(1) = \{e_1, \dots, e_N\}$. Докажем неравенство $d_n(W, \ell_q^N) \gtrsim \min\{1, N^{1/q}n^{-1/q}\}^\omega$, $\omega = (1/p - 1/q)/(1/2 - 1/q)$, соответствующее порядку исходного поперечника (см. (1.16)). Если $n \leqslant N^{2/q}$, достаточно применить оценку $d_n(W, \ell_q^N) \geqslant d_n(B_1^N, \ell_q^N) \gtrsim 1$. Если $n > N^{2/q}$, возьмём в качестве k минимальную степень двойки, для которой $n \leqslant (1/16)N^{2/q}k^{1-2/q}$, тогда

$$d_n(W, \ell_q^N) \geqslant k^{-1/p} d_n(W(k), \ell_q^N) \geqslant \frac{1}{2}k^{1/q-1/p} \asymp (N^{1/q}n^{-1/2})^\omega.$$

По построению $|W| \asymp N \log N$.

Остаётся рассмотреть случай $q = \infty$, $2 \leqslant p < \infty$, см. (1.17). В работе Гарнаева и Глускина оценка снизу сводится к поперечнику множества V_k^N, и мы снова покажем, что можно заменить V_k^N на небольшое его подмножество.

Напомним, что грассманианом $G(N,n)$ называется множество (многообразие) всех n-мерных линейных подпространств $\mathbb{R}^N$. На грассманиане вводится стандартная метрика:

$$\rho_G(Q,S) := \inf_{U \in \mathrm{O}(N)\colon UQ=S} \|\mathrm{Id} - U\|, \quad Q, S \in G(N,n),$$

где $\|\cdot\|$ — операторная норма ($\|T\| = \max |Tx|/|x|$), а $\mathrm{O}(N)$ — группа ортогональных операторов.

Отметим, что для любых $x \in \mathbb{R}^N$ и $Q, S \in G(N,n)$ справедливо неравенство:

$$|\rho(x,Q)_\infty - \rho(x,S)_\infty| \leqslant 2N^{1/2}\rho_G(Q,S)\|x\|_\infty. \tag{5.13}$$

Действительно, пусть $\rho(x,Q)_\infty = \|x-z\|_\infty$, $z \in Q$. Ясно, что $\|z\|_\infty \leqslant \|x\|_\infty + \|x-z\|_\infty \leqslant 2\|x\|_\infty$. По определению метрики ρ_G найдётся оператор U такой, что $S = UQ$, $\|\mathrm{Id} - U\| = \rho_G(Q,S)$. Для $z' := Uz$ имеем: $z' \in S$ и

$$\|z-z'\|_\infty \leqslant |z-z'| = |(\mathrm{Id}-U)z| \leqslant \rho_G(Q,S)|z| \leqslant \\ \leqslant N^{1/2}\rho_G(Q,S)\|z\|_\infty \leqslant 2N^{1/2}\rho_G(Q,S)\|x\|_\infty.$$

Поскольку $\rho(x,S) \leqslant \rho(x,Q) + \|z-z'\|_\infty$, мы получаем нужное неравенство.

Нам потребуется оценка ε-энтропии грассманиана.

Теорема 5.D ([136], Lemma 7.3)**.** *Для любых $1 \leqslant n < N$ и $\varepsilon \in (0, \sqrt{2}]$ справедливо неравенство*

$$(c_1/\varepsilon)^{n(N-n)} \leqslant N_\varepsilon(G(N,n),\rho_G) \leqslant (c_2/\varepsilon)^{n(N-n)}.$$

Лемма 5.7. *Пусть $N,k,n \in \mathbb{N}$, $k \leqslant N$ и выполнено неравенство $n \leqslant (1/5)k\log_2(2N/k)$. Тогда существует множество $W \subset \operatorname{extr} V_k^N$ такое, что $|W| \leqslant CN\log N$ и $d_n(W,\ell_\infty^N) \geqslant 1/6$.*

Доказательство. Множество $V := \operatorname{extr} V_k^N$ содержит $2^k\binom{N}{k} \geqslant (2N/k)^k$ точек с попарными расстояниями в ℓ_∞^N не меньше 1. Пусть $Q_n \in G(N,n)$. Обозначим через $M(Q_n) := \{v \in V\colon \rho(v,Q_n)_\infty \leqslant 1/3\}$ множество "хорошо приближаемых" точек. Покажем, что $|M(Q_n)| \leqslant 8^n$. Действительно, каждой точке $v \in M(Q_n)$ можно сопоставить $v' \in Q_n$, $\|v-v'\|_\infty \leqslant 1/3$. Полученные точки лежат в шаре радиуса $4/3$ в n-мерном подпространстве, и попарные расстояния между ними не меньше $1/3$. Стандартное неравенство (1.29) даёт $|M(Q_n)| \leqslant (1 + (8/3)/(1/3))^n = 9^n$.

Неравенство в условии леммы означает, что $(2N/k)^k \geqslant 32^n$, откуда $|M(Q_n)|/|V| < 2^{-n}$. Возьмём случайную выборку точек $\xi^1,\dots,\xi^m$ во множестве V. Имеем $\mathsf{P}(\xi^1 \in M(Q_n)) \leqslant 2^{-n}$, откуда, в силу независимости выбора,

$$\mathsf{P}(\xi^1,\dots,\xi^n \in M(Q_n)) \leqslant 2^{-mn}.$$

Другими словами, если взять

$$W := \{\xi^1(\omega),\dots,\xi^m(\omega)\},$$

то $\mathsf{P}(\rho(W,Q_n)_\infty \leqslant 1/3) \leqslant 2^{-mn}$. Пусть $\mathcal{N}$ — минимальная ε-сеть в $G(N,n)$ для $\varepsilon = 1/(12N^{1/2})$. Если выполнено неравенство $|\mathcal{N}|\cdot 2^{mn} < 1$, то с положительной

вероятностью для множества W имеем

$$\min_{Q\in\mathcal{N}} \rho(W,Q)_\infty > 1/3. \tag{5.14}$$

Выведем из (5.14), что $d_n(W,\ell_\infty^N) \geqslant 1/6$. Возьмём произвольно $S \in G(N,n)$; найдётся близкое подпространство $Q \in \mathcal{N}$, $\rho(S,Q) \leqslant \varepsilon$, и плохо приближаемая точка $v \in W$, $\rho(v,Q) > 1/3$. В силу (5.13) и нашего выбора ε, $\rho(v,S) \geqslant \rho(v,Q) - 2\varepsilon N^{1/2} \geqslant 1/3 - 1/6 = 1/6$.

Применим теорему 5.D:

$$|\mathcal{N}|2^{-mn} \leqslant (c_2/\varepsilon)^{n(N-n)}2^{-mn} \leqslant (c_3 N)^{Nn}2^{-mn} < 1,$$

если взять $m \asymp N\log N$. Таким образом, мы доказали существование множества $W \subset V$, $|W| \leqslant m \asymp N\log N$, для которого $d_n(W,\ell_\infty^N) \geqslant 1/6$. □

Продолжим доказательство теоремы. Будем считать, что N — степень двойки. Пусть $n = N2^{-2r}$, $r \in \mathbb{N}$. Положим $N' := N2^{-r}$. Из формулы (1.17) видно, что $d_n(B_2^N,\ell_\infty^N) \asymp d_n(B_2^{N'},\ell_\infty^{N'})$. Воспользуемся доказанной леммой для $N' = N2^{-r}$, $n = N2^{-2r}$ и минимально возможного $k = k(r)$; получим множество $W(r)$ из $\asymp N'\log N' \asymp N\log(N)2^{-r}$ элементов такое, что $d_n(W(r),\ell_\infty^N) \geqslant 1/6$. При этом $W(r) \subset k(r)^{1/p}B_p^N$ и $k(r) \asymp n/\log(2N/n)$. Возьмём в качестве множества свидетелей объединение $k(r)^{1/p}W(r)$ по всем возможным r. □

Заключение

В диссертации развит метод получения нижних оценок через усреднённые поперечники. Найдены порядки поперечников в ряде важных и интересных случаев. Исследованы некоторые новые задачи.

Исследования по теме диссертации могут быть продолжены в различных направлениях. Приведём ряд вопросов для дальнейшего изучения.

Глава 1, §1.3. Теорема 1.1 позволяет доказать жёсткость систем независимых функций в разных симметричных пространствах. Было бы интересно описать все такие пространства, хотя бы в избранных шкалах (Лоренца, Орлича). Теорема 1.2 накладывает ограничительные условия (I), (U), (D) для L_1-жёсткости; можно ли от них отойти? Будет ли хаос Радемахера $\{r_i(t)r_j(t)\}_{1\leqslant i<j\leqslant N}$ жёстким в L_1? Жёсткость позволила бы улучшить оценку $n \gtrsim \log N$ размерности, необходимой для хорошей аппроксимации в L_1 первых N функций системы Уолша.

§1.4. Мы увидели, что системы функций с богатой структурой иногда можно хорошо приблизить. Вероятно, что тригонометрическая система не является жёсткой, хотя бы в L_1. Формально, верно ли, что

$$d_n(\{e^{ikx}\}_{k=-N}^{N}, L_1^{\mathbb{C}}(\mathbb{T})) = o(1) \quad \text{при подходящем } n = o(N),\ N \to \infty?$$

Верно ли, что для фиксированного вектора $u \in [-1,1]^N$ имеем

$$d_n(\{u, Tu, \dots, T^{N-1}u\}, \ell_1^N) = o(N), \quad \text{при } n = o(N),$$

где T — оператор циклического сдвига? З. Двирь и А. Лиу [99] доказали "нежёсткость" такого циркулянта в смысле Вэлианта. Возможно, аналоги утверждений 1.15 и 1.16 о приближении в L_0 пространствами очень малой размерности справедливы и для произвольной группы характеров $\widehat{G}$.

Глава 2, §2.1. Классы Соболева $W_p^r[a,b]$ можно определить и для нецелой гладкости $r > 0$ с помощью производной по Вейлю, см. §4.3. Б.С. Кашин показал, что в случае малой гладкости порядок поперечника имеет другой вид, см. [41]; некоторые результаты получены Куланиным в [51] и его диссертации, однако остаются ещё нерешённые случаи. Интересно было бы рассмотреть аналог класса W_1^1 для функций нескольких переменных. Связанная с W_1^1 задача о приближении верхнетреугольной матрицы $\Delta^{(N)}$ (стр. 85) в чебышёвской

норме остаётся нерешённой (подробнее см. [163]). Параграф 2.2. Помимо пространств Бесова, имеет смысл изучить критический случай в других шкалах пространств, например, Лизоркина–Трибеля.

В главе 4 с использованием метода Шапиро–Уоррена доказано существование плохо приближаемых тензоров и функций. Было бы интересно получить нижние оценки малоранговой аппроксимации для конкретных тензоров. Например, оценить минимальный ранг тензора, хорошо аппроксимирующего d-мерный тензор равенства $\mathrm{Id}_{i_1,\ldots,i_d} := \mathbf{1}\{i_1 = \ldots = i_d\}$, $i_j \in [N]$, при больших N и d. Подобные оценки, конечно, требуют развития других методов.

Глава 5, §5.3. Хотелось бы получить оценки с минимально возможным по порядку количеством свидетелей, $|W| = O(N)$. Также интересно получить какой-нибудь общий результат о свидетелях, не использующий явный вид порядков поперечников.

Литература

[1] Р.Р. Акопян, В.В. Арестов, В.Г. Тимофеев, “Задача Стечкина о приближении оператора дифференцирования в равномерной норме на полупрямой”, *Матем. заметки*, **115**:6 (2024), 807–824.

[2] С.В. Асташкин, Ф.А. Сукочев, “Независимые функции и геометрия банаховых пространств”, *УМН*, **65**:6(396) (2010), 3–86.

[3] В.Ф. Бабенко, “Приближения в среднем при наличии ограничений на производные приближающих функций”, *Вопросы анализа и приближения. Сб. научн. тр. Киев: ИМ АН УССР* (1989), 9–18.

[4] В.Ф. Бабенко, “О наилучших равномерных приближениях сплайнами при наличии ограничений на их производные”, *Матем. заметки*, **50**:6 (1991), 24–30.

[5] В.Ф. Бабенко, “О наилучших L_1-приближениях сплайнами при наличии ограничений на их производные”, *Матем. заметки*, **51**:5 (1992), 12–19.

[6] М.В. Балашов, “Метод проекции градиента на матричных многообразиях”, *Ж. вычисл. матем. и матем. физ.*, **60**:9 (2020), 1453–1461.

[7] Э.С. Белинский, “Приближение периодических функций “плавающей” системой экспонент и тригонометрические поперечники”, *Исследования по теории функций многих вещественных переменных*, Ярославль (1984), 10–24.

[8] О.В. Бесов, “Продолжение некоторых классов дифференцируемых функций за пределы области”, *Тр. МИАН СССР*, **77** (1965), 35–44.

[9] А.П. Буслаев, “Поперечники случайных векторов и их применения в теории приближений”, *Матем. заметки*, **43**:4 (1988), 474–497.

[10] В.М. Бухштабер, “Поперечники Колмогорова, многообразия Грассмана и развертка временных рядов”, *Матем. сб.*, **216**:3 (2025), 49–68.

[11] Н.Н. Вахания, В.И. Тариеладзе, С.А. Чобанян, *Вероятностные распределения в банаховых пространствах*, 1985.

[12] А.А. Васильева, “Оценки колмогоровских поперечников пересечения двух шаров в смешанной норм”, *Матем. сб.*, **215**:1 (2024), 82–98.

[13] С.М. Воронин, Н.Т. Темиргалиев, “О некоторых приложениях меры Банаха”, *Изв. АН КазССР, сер.физ.-мат.*, **5** (1984), 8–11.

[14] Э.М. Галеев, “Поперечники по Колмогорову пересечения классов периодических функций и конечномерных множеств”, *Матем. заметки*, **29**:5 (1981), 749–760.

[15] Э.М. Галеев, “Поперечники по Колмогорову классов периодических функций одной и нескольких переменных”, *Изв. АН СССР. Сер. матем.*, **54**:2 (1990), 418–430.

[16] Э.М. Галеев, “Поперечники по Колмогорову некоторых конечномерных множеств в смешанной норме”, *Матем. заметки*, **58**:1 (1995), 144–148.

[17] Э.М. Галеев, “Линейные поперечники классов Гёльдера–Никольского периодических функций многих переменных”, *Матем. заметки*, **59**:2 (1996), 189–199.

[18] Э.М. Галеев, “Поперечники классов Бесова $B^r_{p,\theta}(\mathbb{T}^d)$”, *Матем. заметки*, **69**:5 (2001), 656–665.

[19] Э.М. Галеев, “Поперечники функциональных классов и конечномерных множеств”, *Владикавк. матем. журн.*, **13**:2 (2011), 3–14.

[20] В.Ф. Гапошкин, “Лакунарные ряды и независимые функции”, *УМН*, **21**:6(132) (1966), 3–82.

[21] А.Л. Гаркави, “О совместном приближении периодической функции и ее производных тригонометрическими полиномами”, *Изв. АН СССР. Сер. матем.*, **24**:1 (1960), 103–128.

[22] А.Ю. Гарнаев, Е.Д. Глускин, “О поперечниках евклидова шара”, *Докл. АН СССР*, **277**:5 (1984), 1048–1052.

[23] Е.Д. Глускин, “Об одной задаче о поперечниках”, *Докл. АН СССР*, **219**:3 (1974), 527–530.

[24] Е.Д. Глускин, “О некоторых конечномерных задачах теории поперечников”, *Вестн. ЛГУ*, **13** (1981), 5–10.

[25] Е.Д. Глускин, “Нормы случайных матриц и поперечники конечномерных множеств”, *Матем. сб.*, **120(162)**:2 (1983), 180–189.

[26] Е.Д. Глускин, “Октаэдр плохо приближается случайными подпространствами”, *Функц. анализ и его прил.*, **20**:1 (1986), 14–20.

[27] Е.Д. Глускин, “Пересечения куба с октаэдром плохо аппроксимируются подпространствами малой размерности”, *Приближение функций специальными классами операторов*, Вологда, 1987, 35–41.

[28] Е.Д. Глускин, “Экстремальные свойства ортогональных параллелепипедов и их приложения к геометрии банаховых пространств”, *Матем. сб.*, **136**:1 (1988), 85–96.

[29] М.Г. Григорян, В.Г. Кротов, “Теорема исправления Лузина и коэффициенты разложений Фурье по системе Фабера–Шаудера”, *Матем. заметки*, **93**:2 (2013), 172–178.

[30] Динь Зунг, “Приближение функций многих переменных на торе тригонометрическими полиномами”, *Матем. сб.*, **131(173)**:2(10) (1986), 251–271.

[31] Н.Л. Замарашкин, С.В. Морозов, Е.Е. Тыртышников, “Об алгоритме наилучшего приближения матрицами малого ранга в норме Чебышёва”, *Ж. вычисл. матем. и матем. физ.*, **62**:5 (2022), 723–741.

[32] В.И. Иванов, В.А. Юдин, "О тригонометрической системе в L_p, $0 < p < 1$", *Матем. заметки*, **28**:6 (1980), 859–868.

[33] Р.С. Исмагилов, "Об n-мерных поперечниках компактов в гильбертовом пространстве", *Функ. ан. и прил.*, **2**:2 (1968), 32–39.

[34] Р.С. Исмагилов, "Поперечники множеств в линейных нормированных пространствах и приближение функций тригонометрическими многочленами", *УМН*, **29**:3(177) (1974), 161–178.

[35] А.Д. Изаак, "Поперечники по Колмогорову в конечномерных пространствах со смешанной нормой", *Матем. заметки*, **55**:1 (1994), 43–52.

[36] А.Д. Изаак, "Поперечники классов Гёльдера–Никольского и конечномерных множеств в пространствах со смешанной нормой", *Матем. заметки*, **59**:3 (1996), 459–461.

[37] Б.С. Кашин, "О колмогоровских поперечниках октаэдров" *Докл. АН СССР*, **214**:5 (1974), 1024–1026.

[38] Б.С. Кашин, "О поперечниках октаэдров", *УМН*, **30**:4(184) (1975), 251–252.

[39] Б.С. Кашин, "Поперечники некоторых конечномерных множеств и классов гладких функций", *Изв. АН СССР. Сер. матем.*, **41**:2 (1977), 334–351.

[40] Б.С. Кашин, "О некоторых свойствах матриц ограниченных операторов из пространства ℓ_2^n в ℓ_2^m", *Изв. АН Арм.ССР. Матем.*, **15**:5 (1980), 379–394.

[41] Б.С. Кашин, "О поперечниках классов Соболева малой гладкости", *Вестн. МГУ. Сер. 1. Матем., мех.*, **5** (1981), 50–54.

[42] Б.С. Кашин, "Аналог теоремы Меньшова "об исправлении" для дискретных ортонормированных систем", *Матем. заметки*, **46**:6 (1989), 67–74.

[43] Б.С. Кашин, А.А. Саакян, *Ортогональные ряды*. Москва, 1999.

[44] Б.С. Кашин, В.Н. Темляков, “Замечание о задаче сжатого измерения”, *Матем. заметки*, **82**:6 (2007), 829–837.

[45] А.Н. Колмогоров, “О наилучшем приближении функций заданного класса”, *А.Н. Колмогоров. Избранные труды. Математика и механика*. Наука, М., 1985. Перевод; оригинал: “Ueber die beste Annäherung von Funktionen einer gegebenen Funktionen-klasse”, *Ann.Math.*, **37** (1936), 107–110.

[46] А.Н. Колмогоров, А.А. Петров, Ю.М. Смирнов, “Одна формула Гаусса по теории метода наименьших квадратов”, *Изв. АН СССР, сер. мат.*, **11**:6 (1947), 561–566.

[47] В.Н. Коновалов, “Оценки поперечников типа Колмогорова для классов дифференцируемых периодических функций”, *Матем. заметки*, **35**:3 (1984), 369–380.

[48] В.Н. Коновалов, “Приближение классов Соболева их конечномерными сечениями”, *Матем. заметки*, **72**:3 (2002), 370–382.

[49] Н.П. Корнейчук, *Экстремальные задачи теории приближения*, М.: Наука, 1976.

[50] Е.Д. Куланин, “О поперечниках класса функций ограниченной вариации в пространстве $L^q(0,1)$, $2 < q < \infty$”, *УМН*, **38**:5(233) (1983), 191–192.

[51] Е.Д. Куланин, “Оценки поперечников классов Соболева малой гладкости”, *Вестн. Моск. ун-та. Сер. 1. Матем., мех.*, **2** (1983), 24–30.

[52] В.Е. Майоров, “Интегральные и колмогоровские поперечники классов дифференцируемых функций”, *Докл. АН СССР*, **318**:5 (1991), 1082–1085.

[53] В.Е. Майоров, “Поперечники пространств, наделенных мерой Гаусса”, *Докл. РАН*, **323**:2 (1992), 233–237.

[54] В.Е. Майоров, “Колмогоровские (n,δ)-поперечники пространств гладких функций”, *Матем. сб.*, **184**:7 (1993), 49–70 .

[55] Ю.И. Маковоз, “Об одном приеме оценки снизу поперечников множеств в банаховых пространствах”, *Матем. сб.*, **87(129)**:1 (1972), 136–142.

[56] А.И. Мальцев, “Замечание к работе А. Н. Колмогорова, А. А. Петрова и Ю. М. Смирнова “Одна формула Гаусса из теории наименьших квадратов””, *Изв. АН СССР. Сер. матем.*, **11**:6 (1947), 567–568.

[57] В.Н. Малоземов, А.Б. Певный, *Полиномиальные сплайны*, Изд. Лен. Унив., Ленинград, 1986.

[58] В.Д. Мильман, “Эссе о замечательной теореме Кашина 1977 г. о разложении”, *Труды МИАН*, **319** (2022), 213–222.

[59] С.М. Никольский, “О некоторых методах приближения тригонометрическими суммами”, *Изв. АН СССР. Сер. матем.*, **4**:6 (1940), 509–520.

[60] И.Я. Новиков, В.Ю. Простасов, М.А. Скопина, *Теория всплесков*, М.: Физматлит, 2006.

[61] Ал.А. Привалов, “О росте степеней полиномиальных базисов и приближении тригонометрических проекторов”, *Матем. заметки*, **42**:2 (1987), 207–214.

[62] А.А. Разборов, “О квантовой коммуникационной сложности симметрических предикатов”, *Изв. РАН. Сер. матем.*, **67**:1 (2003), 159–176.

[63] А.С. Романюк, “Приближение классов Бесова периодических функций многих переменных в пространстве L_q”, *Укр. матем. журн.*, **43**:10 (1991), 1398–1408.

[64] А.С. Романюк, “О наилучших тригонометрических приближениях и колмогоровских поперечниках классов Бесова функций многих переменных”, *Укр. матем. журн.*, **45**:5 (1993), 663–675.

[65] А.С. Романюк, “Колмогоровские и тригонометрические поперечники классов Бесова $B^r_{p,\theta}$ периодических функций многих переменных”, *Матем. сб.*, **197**:1 (2006), 71–96.

[66] М.И. Стесин, “Александровские поперечники конечномерных множеств и классов гладких функций”, *Докл. АН СССР*, **220**:6 (1975), 1278–1281.

[67] С.Б. Стечкин, “О наилучшем приближении заданных классов функций любыми полиномами”, *УМН*, **9**:1(59) (1954), 133–134.

[68] М.З. Соломяк, В.М. Тихомиров, “О геометрических характеристиках вложения классов W_p^α в C”, *Изв. вузов. Матем.*, **10** (1967), 76–82.

[69] Ю.Н. Субботин, С.А. Теляковский, “Точные значения относительных поперечников классов дифференцируемых функций”, *Матем. заметки*, **65**:6 (1999), 871–879.

[70] Ю.Н. Субботин, С.А. Теляковский, “Сплайны и относительные поперечники классов дифференцируемых функций”, *Теория приближений. Асимптотические разложения*, Сборник статей, Тр. ИММ УрО РАН, **7**:1 (2001), 208–216.

[71] Ю.Н. Субботин, С.А. Теляковский, “Относительные поперечники классов дифференцируемых функций в метрике L^2”, *УМН*, **56**:4(340) (2001), 159–160.

[72] А.В. Сульдин, “ Мера Винера и ее приложения к приближенным методам. I”, *Изв. вузов. Матем.*, **6** (1959), 145–158.

[73] В.Н. Темляков, “Приближение функций с ограниченной смешанной производной”, *Тр. МИАН СССР*, **178** (1986), 3–113.

[74] В.Н. Темляков, “Оценки наилучших билинейных приближений функций двух переменных и некоторые их приложения”, *Матем. сб.*, **134(176)**:1(9) (1987), 93–107.

[75] В.Н. Темляков, “Оценки наилучших билинейных приближений периодических функций”, *Тр. МИАН СССР*, **181** (1988), 250–267.

[76] В.Н. Темляков, “Оценки наилучших билинейных приближений функций и аппроксимативных чисел интегральных операторов”, *Матем. заметки*, **51**:5 (1992), 125–134.

[77] В.М. Тихомиров, "Поперечники множеств в функциональных пространствах и теория наилучших приближений", *УМН*, **15**:3(93) (1960), 81–120.

[78] В.М. Тихомиров, "Наилучшие методы приближения и интерполирования дифференцируемых функций в пространстве $C[-1,1]$", *Матем. сб.*, **80(122)**:2(10) (1969), 290–304.

[79] В.М. Тихомиров, *Некоторые вопросы теории приближений*, МГУ, М., 1976.

[80] В.М. Тихомиров, *Теория приближений.* Итоги науки и техн. Сер. Современные проблемы математики. Фундаментальные направления. **14**, ВИНИТИ, М., 1987, 103–260.

[81] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Dept. Commerce, NBS, Tenth Printing, 1972.

[82] M. Ainger, G.M. Ziegler, *Proofs from THE BOOK*, Springer, 6-th ed., 2018.

[83] J. Alman, R. Williams, "Probabilistic rank and matrix rigidity", *STOC 2017: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing* (2017), 641–652.

[84] N. Alon, "Perturbed Identity Matrices Have High Rank: Proof and Applications", *Combinatorics, Probability and Computing*, **18**:1-2 (2009), 3–15.

[85] N. Alon, P. Frankl, V. Rödl, "Geometrical realization of set systems and probabilistic communication complexity", *Proc. 26th Ann. Symposium on Foundations of Computer Science* (1985), 277-280.

[86] N. Alon, T. Lee, A. Shraibman, S. Vempala, "The approximate rank of a matrix and its algorithmic applications", *Proc. STOC* 2013, 675–684.

[87] S. Arstein–Avidan, A. Giannopoulos, V.D. Milman, *Asymptotic Geometric Analysis*, Part I. Mathematical surveys and monographs, 202. AMS, Providence, 2015.

[88] D. Bazarkhanov, V. Temlyakov, "Nonlinear tensor product approximation of functions", *J. Complexity*, **31**:6 (2015), 867–884.

[89] T. Bendokat, R. Zimmermann, P.-A. Absil, "A Grassmann manifold handbook: basic geometry and computational aspects", *Adv. Comput. Math.*, **50**:6 (2024).

[90] I. Berkes, "On the uniform theory of lacunary series". *In: Number Theory — Diophantine Problems, Uniform Distribution and Applications.* Springer, Cham. 2017.

[91] S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities*, Oxford University Press, 2013.

[92] J. Bourgain, "Bounded orthogonal systems and the $\Lambda(p)$-set problem", *Acta Math.* **162** (1989), 227-245.

[93] S. Brazitikos, A. Giannopoulos, P. Valettas, B-H. Vritsiou, *Geometry of isotropic convex bodies*, Mathem. surveys and monographs 196, AMS, Providence, Rhode Island, 2014.

[94] J. Creutzig, "Relations between Classical, Average, and Probabilistic Kolmogorov Widths", *J. Complexity*, **18** (2002), 287–303.

[95] R. DeVore, G. Lorentz, *Constructive Approximation*, Springer-Verlag Berlin Heidelberg 1993.

[96] S. Dirksen, T. Ullrich, "Gelfand numbers related to structured sparsity and Besov space embeddings with small mixed smoothness", *J. Complexity*, **48** (2018), 69–102.

[97] R.M. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2002.

[98] D. Dung, V. Temlyakov, T. Ullrich, *Hyperbolic Cross Approximation.* Birkhäuser, Cham. 2018.

[99] Z. Dvir, A. Liu, "Fourier and Circulant Matrices Are Not Rigid", *Theory Comput.*, **16** (2020), 1–48.

[100] A. Edelman, T. Arias, S. Smith, "The Geometry of Algorithms with Orthogonality Constraints", *SIAM Journal of Matrix Analysis and Applications*, **20**:2 (1998), 303–353.

[101] S. Foucart, H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser, 2013.

[102] G. Garrigós, A. Seeger, T. Ullrich, "Basis Properties of the Haar System in Limiting Besov Spaces", *GAFA*, **45** (2021), 361–424.

[103] A. Giannopoulos, "A Note on the Banach-Mazur Distance to the Cube". *GAFA*, **77** (1995), 67–73.

[104] A. Giannopoulos, "On some vector balancing problems", *Stud. Math.*, **122**:3 (1997), 225–234.

[105] B. Green, "On the width of transitive sets: Bounds on matrix coefficients of finite groups", *Duke Math. J.*, **169** (2020), 551–578.

[106] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer-Verlag, Heidelberg, 2012.

[107] D. Haroske, L.Skrzypczak, "Entropy and Approximation Numbers of Embeddings of Function Spaces with Muckenhoupt Weights, I", *Rev. Mat. Complut.*, **21**:1 (2008), 135–177.

[108] G. Hargé, "A particular case of correlation inequality for the Gaussian measure", *Ann.Prob.*, **27** (1999), 1939–1951.

[109] A. Hinrichs, J. Prochno, J. Vybíral, "Gelfand numbers of embeddings of Schatten classes", *Math. Ann.*, **380** (2021), 1563–1593.

[110] C.R. Hobby, J.R. Rice, "A Moment Problem in L_1 Approximation", *Proc. AMS*, **16**:4 (1965), 665–670.

[111] K. Hollig, "Approximationszahlen von Sobolev-Einbettungen", *Math. Ann.*, **242**:3 (1979), 273–281.

[112] S. Karlin, "Some variational problems on certain Sobolev spaces and perfect splines", *Bull. AMS*, **79**:1 (1973), 124–128.

[113] A.R. Klivans, A.A. Sherstov, "Lower Bounds for Agnostic Learning via Approximate Rank", *comput. complex.*, **19** (2010), 581–604.

[114] R. Latala, K. Oleszkiewicz, "Gaussian measures of dilatations of convex symmetric sets", *Ann. Probab.*, **27** (1999), 1922–1938.

[115] T. Lee, A. Shraibman, "Lower Bounds in Communication Complexity", *Foundations and Trends in Theoretical Computer Science*, **3**:4 (2009), 263–399.

[116] J. Lindenstrauss, L. Tzafriri, *Classical Banach Spaces I and II*, Springer. 1996.

[117] S.V. Lokam, "Complexity Lower Bounds using Linear Algebra", *Foundations and Trends in Theoretical Computer Science*, **4**:1-2 (2008), 1–155.

[118] G.G. Lorentz, "Lower Bounds for the Degree of Approximation", *Trans. AMS*, **97**:1 (1960), 25–34.

[119] G.G. Lorentz, M. Golitschek, Y. Makovoz, *Constructive Approximation: Advanced Problems*, Springer, 1996.

[120] S. Mendelson, R. Vershynin, "Entropy and the combinatorial dimension", *Invent. math.* **152** (2003), 37–55.

[121] T. Oikhberg, M.I. Ostrovskii, "Dependence of Kolmogorov Widths on the Ambient Space", *Zh. Mat. Fiz. Anal. Geom.*, **9**:1 (2013), 25–50.

[122] I. Oseledets, E. Tyrtyshnikov, "TT-cross approximation for multidimensional arrays", *Linear Algebra and its Applications*, **432**:1 (2010), 70–88.

[123] N.V. Parfinovych, "Exact order of relative widths of classes W_1^r in the space L_1", *Urk.Math.J.*, **57**:10 (2005), 1652–1662.

[124] R. Paturi, J. Simon, "Probabilistic communication complexity", *Journal of Computer and System Sciences*, vol. **33**, no. 1, pp. 106–123.

[125] A. Pietsch, "s-Numbers of operators in Banach space", *Studia Math.*, **51** (1974), 201–223.

[126] A. Pinkus, "A simple proof of the Hobby-Rice theorem", *Proc. AMS*, **60** (1976), 82–84.

[127] A. Pinkus, *n-Widths in Approximation Theory*, Springer Berlin, 1985.

[128] J. Prochno, M. Strzelecki, "Approximation, Gelfand, and Kolmogorov numbers of Schatten class embeddings", *J. Appr. Theory*, **277** (2022), article 105736.

[129] T. Royen, "A simple proof of the gaussian correlation conjecture extended to some multivariate gamma distributions", *Far East Journ. Theoret. Statist.*, **48**:2 (2014), 139–145.

[130] W. Rudin, "L^2-approximation by partial sums of orthogonal developments", *Duke Math J.*, **19**:1 (1952), 1–4.

[131] A. Sah, M. Sawhney, Y. Zhao, "The cylindrical width of transitive sets", *Isr. J. Math.*, **253** (2023), 647–672.

[132] I.J. Schoenberg, *Cardinal Spline Interpolation*, SIAM, Philadelphia, 1973.

[133] R. Schneider, A. Uschmajew, "Approximation rates for the hierarchical tensor format in periodic Sobolev spaces", *J. Complexity*, **30**:2 (2014), 56–71.

[134] H.S. Shapiro, "Some negative theorems of approximation theory", *Michigan Math. J.*, **11**(3), 211–217.

[135] D.R. Stinson, *Combinatorial Designs: Constructions and Analysis.* Springer, 2004.

[136] S.J. Szarek, "The finite dimensional basis problem with an appendix on nets of Grassmann manifolds", *Acta Math.* **151** (1983), 153–179.

[137] S.J. Szarek, M. Talagrand, "An 'isomorphic' version of the Sauer-Shelah lemma and the Banach-Mazur distance to the cube", *GAFA*, **1376** (1989).

[138] T. Tao, V.H. Vu, *Additive combinatorics*. Cambridge University Press, 2010.

[139] V. Temlyakov, *Multivariate approximation*, Cambridge University Press, 2018.

[140] V.M. Tikhomirov, "Some remarks on relative diameters", *Banach Center Publications*, **22**:1 (1989), 471–474.

[141] J.F. Traub, G.W. Wasilkowski, H. Wozniakowski, *Information-Based Complexity*, Academic Press, 1988.

[142] H. Triebel, *Function Spaces and Wavelets on Domains*, EMS Tracts in Mathematics, 2008.

[143] H. Triebel, *Theory of Function Spaces II*, Birkhäuser Verlag, 1992.

[144] H. Triebel, *Theory of Function Spaces III*, Birkhäuser Verlag, 2006.

[145] R.M. Trigub, E.S. Belinskii, *Fourier Analysis and Approximation of Functions*, Springer, 2004.

[146] J. Vaaler, "A geometric inequality with applications to linear forms", *Pacific J. Math.* **83**(2) (1979), 543–553.

[147] L.G. Valiant, "Graph-theoretic arguments in low-level complexity", *Mathematical Foundations of Computer Science (MFCS)* (1977), 162–176.

[148] A.A. Vasil'eva, "Kolmogorov and linear widths of the weighted Besov classes with singularity at the origin", *J. Approx. Theory*, **167** (2013), 1–41.

[149] A.A. Vasil'eva, "Kolmogorov widths of intersections of finite-dimensional balls", *J.Complexity*, **72** (2022), 101649.

[150] A.A. Vasil'eva, "Kolmogorov widths of an intersection of a family of balls in a mixed norm", *J. Appr. Theory*, **301** (2024), article 106046.

[151] R. Vershynin, *High-Dimensional Probability. An Introduction with Applications in Data Science*. Cambridge University Press, 2018.

[152] H.E. Warren, “Lower bounds for approximation by nonlinear manifolds”, *Trans. AMS*, **133**:1 (1968), 167–178.

[153] P. Wojtaszczyk, *A Mathematical Introduction to Wavelets*, Cambridge University Press, 1997.

[154] V.A. Zorich, *Mathematical Analysis I*, Springer, 2004.

Работы автора по теме диссертации

[155] Y. Malykhin, “Matrix and tensor rigidity and L_p-approximation”, *J. Complexity*, **72** (2022), 101651, 13 pp.

[156] Ю.В. Малыхин, “Поперечники и жесткость”, *Матем. сб.*, **215**:4 (2024), 117–148.

[157] Ю.В. Малыхин, К.С. Рютин, “Поперечники и жесткость безусловных множеств и случайных векторов”, *Изв. РАН. Сер. матем.*, **89**:2 (2025), 45–59.

[158] Ю.В. Малыхин, К.С. Рютин, “Произведение октаэдров плохо приближается в метрике $\ell_{2,1}$”, *Матем. заметки*, **101**:1 (2017), 85–90.

[159] Ю.В. Малыхин, “Колмогоровские поперечники классов Бесова $B^1_{1,\theta}$ и произведений октаэдров”, *Труды МИАН*, **312** (2021), 224–235.

[160] Ю.В. Малыхин, “Асимптотические свойства чебышёвских сплайнов с фиксированным числом узлов”, *Фундамент. и прикл. матем.*, **19**:5 (2014), 143–166.

[161] Ю.В. Малыхин, “Полное описание относительных поперечников классов Соболева в равномерной метрике”, *Труды ИММ УрО РАН*, **28**:4 (2022), 137–142.

[162] Ю.В. Малыхин, “Относительные поперечники классов Соболева в равномерной и интегральной метриках”, *Труды МИАН*, **293** (2016), 217–223.

[163] Б.С. Кашин, Ю.В. Малыхин, К.С. Рютин, “Поперечник по Колмогорову и аппроксимативный ранг”, *Труды МИАН*, **303** (2018), 155–168.

[164] Ю.В. Малыхин, “О структуре матриц малого ранга, приближающих единичную матрицу”, *Матем. заметки*, **117**:5 (2025), 795–798.

[165] Ю.В. Малыхин, “Колмогоровские поперечники класса W_1^1”, *Матем. заметки*, **117**:6 (2025), 922–927.