Курс математики

В. М. ТИХОМИРОВ
Московский государственный университет
им. М. В. Ломоносова
e-mail: vmtikh@googlemail.com

УДК 510

Ключевые слова: альтернатива Фредгольма, метод Ньютона, принцип Лагранжа, теорема Гильберта—Шмидта, теорема Пуанкаре.

Аннотация

Обсуждаются проблемы современного математического образования. В статье собраны некоторые фундаментальные вопросы математики, рассматриваемой в её единстве.

Abstract

V. M. Tikhomirov, Course of mathematics, Fundamentalnaya i prikladnaya matematika, vol. 19 (2014), no. 4, pp. 153—196.

This paper is devoted to problems of modern mathematical education. Fundamental problems of mathematics are considered in their unity.

Цель этой статьи — выделить то основное, что составляет теоретическую базу математики, преподающейся у нас в школе, вузе и университете, и потому изложение ведётся одновременно на трёх уровнях: специальной математической школы (A), технического или экономического вуза (B) и естественно-научного университета (C).

Вопрос о том, как разумно вести преподавание математики, обсуждается во всём мире. Эта статья выражает точку зрения автора на стиль и цели математического образования.

В соответствии с традициями образования в России преподавание группировалось по различным «предметам». В этой статье нет разбиения на «предметы»; математика рассматривается как единое целое с её приложениями.

Эта статья предназначена для обсуждения содержания математического образования с учителем школы, преподавателем вуза и с моими коллегами по университетскому преподаванию. Любому из этих моих собеседников мне хотелось показать, что математика, лежащая в основе преподавания любого уровня (школьного, институтского, университетского), проста и по сути своей едина, что основные её понятия формулируются совершенно одинаковым образом, а путь от школьной математики к вершинам университетского образования поразительно краток.

Фундаментальная и прикладная математика, 2014, том 19, № 4, с. 153—196.

I. Понятия

1. Вещественные числа и векторные пространства

(A) *Вещественная прямая* $\mathbb{R}$ *и декартова плоскость* $\mathbb{R}^2$.

Вещественная прямая — важнейшее математическое понятие. Она обозначается символом $\mathbb{R}$, в ней выделяются *натуральный ряд* $\mathbb{N}$, *совокупность целых чисел* $\mathbb{Z}$ и *дробей* $\mathbb{Q}$. Числовая прямая описывается двояко: набором свойств (аксиом) и «арифметической моделью». Аксиоматически $\mathbb{R}$ — *полное упорядоченное поле*. Это означает, что вещественные числа можно складывать и умножать, их можно сравнивать, и эти операции удовлетворяют, как учат в школе, переместительному, сочетательному и распределительному законам, а в институте и университете их называют коммутативным, ассоциативным и дистрибутивным законами. Согласно этим законам сумма не зависит от порядка слагаемых, а произведение — от порядка сомножителей; сумма не зависит от группировки её слагаемых, а произведение — от группировки его сомножителей; произведение суммы двух чисел на третье равно сумме произведений каждого из этих двух чисел на третье число. Упорядочение чисел также имеет естественные законы порядка. Всем этим законам удовлетворяют и дроби. Но не все величины можно выразить дробями, например, диагональ квадрата со стороной длины единица, дробью не выражается. Прямая представляет собой *пополнение* совокупности дробей. Существует несколько равносильных аксиом полноты (Больцано, Вейерштрасса, Дедекинда, Кантора, Коши), с ними со всеми приходится сталкиваться, но сейчас, для завершения разговора, приведём лишь аксиому Кантора, согласно которой *последовательность вложенных друг в друга числовых отрезков, длины которых стремятся к нулю, имеет единственное принадлежащее им всем число*. При этом каждое вещественное число можно однозначно представить в виде десятичной дроби $n, q_1 q_2 \ldots$, где n — целое число, а q_i — одно из чисел 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 и отброшены последовательности, оканчивающиеся на одни девятки. Такая реализация вещественных чисел и есть *арифметическая модель* $\mathbb{R}$.

Декартова плоскость $\mathbb{R}^2$ состоит из векторов-столбцов

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Вещественные числа x_1 и x_2 называются *координатами вектора* x. Векторы можно покоординатно складывать и умножать на вещественные числа.

(B) *Конечномерное пространство* $\mathbb{R}^n$. *Векторное пространство*. Пространство $\mathbb{R}^n$ есть n-я степень $\mathbb{R}$. Оно представляет собой совокупность векторов-столбцов

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(иногда располагаемых нами в строку как $(x_1, \dots, x_n)^{\mathrm{T}}$, где T — знак транспонирования). Числа $x_1, \dots, x_n$ называются *координатами вектора* x. Векторы можно покоординатно складывать и умножать на вещественные числа.

Векторным пространством $X = \{x\}$ называется совокупность элементов $\{x\}$, которые можно складывать, образуя для векторов $x, x' \in X$ их сумму $x + x'$, и умножать на вещественные числа, образуя для вектора $x \in X$ и числа $\alpha \in \mathbb{R}$ вектор αx. При этом удовлетворяются естественные законы, как у чисел.

(C) *Нормированное пространство* $(X, \|\cdot\|_X)$. Нормированное пространство — это векторное пространство, оснащённое *нормой*, т. е. функцией $\|\cdot\|_X \colon X \to \mathbb{R}$, обладающей следующими свойствами:

а) если $\|x\|_X \geqslant 0$ для всех $x \in X$ и $\|x\|_X = 0$, то $x = 0$,

б) $\|\alpha x\|_X = |\alpha|\, \|x\|_X$ для всех $x \in X$, $\alpha \in \mathbb{R}$,

в) $\|x + x'\|_X \leqslant \|x\|_X + \|x'\|_X$ для всех $x, x' \in X$ (неравенство треугольника).

В $\mathbb{R}$, $\mathbb{R}^n$ и $(X, \|\cdot\|_X)$ можно ввести *расстояние* $d_X(x, x')$ между x, x' из X формулами

$$d_{\mathbb{R}}(x, x') = |x - x'|, \quad d_{\mathbb{R}^n}(x, x') = |x - x'| = \left(\sum_{i=1}^{n} (x_i - x_i')\right)^{1/2}, \\ d_X(x, x') = \|x - x'\|_X.$$

2. Порядок и общая топология

В этих разделах математики изучаются свойства упорядоченности, непрерывности и компактности.

На первых порах нам понадобится только одно «порядковое» определение.

Число a, обозначаемое $\sup A$ ($\inf A$) называется *верхней* (*нижней*) гранью числового множества A, если для любого $\varepsilon > 0$ пересечение $A \cap [a - \varepsilon, a]$ ($A \cup [a, a + \varepsilon]$) непусто. Из аксиомы Кантора легко выводится следующая лемма.

Лемма Вейерштрасса о верхних и нижних гранях. *Ограниченное сверху* (*снизу*) *числовое множество имеет верхнюю* (*нижнюю*) *грань.*

Пусть X — это $\mathbb{R}$ для (A), $\mathbb{R}^n$ для (B) и нормированное пространство $(X, \|\cdot\|_X)$ для (C). Приведём основные общетопологические определения (они формулируются одинаково для (A), (B) и (C)).

Определение 1 (предел). Элемент $\hat{x}$ называется *пределом последовательности* $\{x_k\}_{k \in \mathbb{N}}$ ($x_k \in X$), если для любого числа $\varepsilon > 0$ найдётся такое число N, что для любого $k > N$ $d_{\mathbb{R}}(x_k, \hat{x}) < \varepsilon$ для (A), $d_{\mathbb{R}^n}(x_k, \hat{x}) < \varepsilon$ для (B), $d_X(x_k, \hat{x}) < \varepsilon$ для (C). При этом пишут $x_n \to \hat{x}$ или $\hat{x} = \lim\limits_{n \to \infty} x_n$. Последовательность, имеющая предел, называется *сходящейся*.

Определение 2 (полнота). Последовательность $\{x_k\}_{k \in \mathbb{N}}$ называется *фундаментальной* в пространстве X, если для любого числа $\varepsilon > 0$ найдётся такое число N, что если k_1 и k_2 превосходят N, то $d_{\mathbb{R}}(x_{k_1}, x_{k_2}) < \varepsilon$ для (A), $d_{\mathbb{R}^n}(x_{k_1}, x_{k_2}) < \varepsilon$ для (B), $d_X(x_{k_1}, x_{k_2}) < \varepsilon$ для (C). Пространство, в котором каждая фундаментальная последовательность является сходящейся, называется *полным*. Полное нормированное пространство называется *банаховым*.

Из аксиомы Кантора выводится полнота вещественной прямой $\mathbb{R}$, откуда следует, что пространства $\mathbb{R}^n$ являются банаховыми. Банаховы пространства — основной класс пространств в этой работе.

Множество $U_X(\hat{x}, \varepsilon) = \{x \in X \mid d_X(x, \hat{x}) < \varepsilon\}$ ($B_X(\hat{x}, \varepsilon) = \{x \in X \mid d_X(x, \hat{x}) \leqslant \varepsilon\}$) называется *открытым* (*замкнутым*) *шаром с центром в* $\hat{x}$ *радиуса* ε. Подмножество $V \subset X$ называется *открытым*, если для любой точки $x \in V$ имеется число $\varepsilon = \varepsilon(x)$, такое что $U_X(x, \varepsilon) \subset V$. Если открытое множество V содержит точку $\hat{x}$, то это множество называют *окрестностью точки* $\hat{x}$. Совокупность окрестностей точки $\hat{x}$ будем обозначать $\mathcal{O}(\hat{x}, X)$. Множество называется *замкнутым*, если его дополнение открыто. Множество $C \subset X$ называется *всюду плотным в пространстве* X, если для любой точки $x \in X$ и любого $\varepsilon > 0$ пересечение $U_X(x, \varepsilon) \cap A$ непусто. Пространство (X, d_X), в котором есть счётное всюду плотное множество, называется *сепарабельным*.

Определение 3 (непрерывность). Пусть X — это $\mathbb{R}$ для (A), $\mathbb{R}^n$ для (B) и нормированное пространство $(X, \|\cdot\|_X)$ для (C), Y — это $\mathbb{R}$ для (A), $\mathbb{R}^m$ для (B) и нормированное пространство $(Y, \|\cdot\|_Y)$ для (C), $V \in \mathcal{O}(\hat{x}, X)$ и $F \colon V \to Y$ — отображение из V в Y. Отображение F называется *непрерывным в точке* $\hat{x} \in X$, если для любого числа $\varepsilon > 0$ найдётся такое число $\delta > 0$, что $d_Y\bigl(F(x), F(\hat{x})\bigr) < \varepsilon$ для любого x, такого что $d_X(x, \hat{x}) < \delta$. Этому определению эквивалентно определение *непрерывности по Гейне* (или *секвенциальной непрерывности*), согласно которому отображение F является *непрерывным в точке* $\hat{x}$, если для всякой последовательности $\{x_k\}_{k \in \mathbb{N}}$, сходящейся к $\hat{x}$, последовательность $\{F(x_k)\}_{k \in \mathbb{N}}$ сходится к $F(\hat{x})$. Если отображение F непрерывно в каждой точке подмножества $T \subset X$, говорят, что F *непрерывно на* T и пишут, что $F \in C(T, Y)$.

Определение 4 (компактность). Пусть X — это $\mathbb{R}$ для (A), $\mathbb{R}^n$ для (B) и нормированное пространство $(X, \|\cdot\|_X)$ для (C). Множество $C \subset X$ называют *компактом*, если из любой последовательности элементов этого множества можно выделить сходящуюся подпоследовательность.

3. Линейность

Класс линейных отображений из одного векторного пространства в другое играет фундаментальную роль во всей математике.

(A) Пусть $a \in \mathbb{R}$. Функция $x \mapsto ax$ (или, как пишут в школе, $y = ax$) называется *линейной функцией* (одного переменного). Пусть a_1 и a_2 — два вещественных числа и

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (x \in \mathbb{R}^2).$$

Функция $x \mapsto a_1x_1 + a_2x_2$ (или, как пишут, $y = a_1x_1 + a_2x_2$) называется *линейной функцией двух переменных*.

Пусть a_{11}, a_{12}, a_{21}, a_{22} — четыре вещественных числа. Отображение, задаваемое формулой

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \in \mathbb{R}^2,$$

называется *линейным отображением из* $\mathbb{R}^2$ *в* $\mathbb{R}^2$. Такому отображению будем ставить в соответствие таблицу

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

из двух строк и двух столбцов. Её называют *матрицей размера* 2×2. Она определяет линейное отображение из $\mathbb{R}^2$ в $\mathbb{R}^2$.

(B) Пусть $a_1, \ldots, a_n$ — n вещественных чисел и

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (x \in \mathbb{R}^n).$$

Функция $x \mapsto a_1x_1 + \ldots + a_nx_n$ (или, как пишут, $y = a_1x_1 + \ldots + a_nx_n$) называется *линейной функцией* n *переменных*. Будем ставить в соответствие линейной функции $y = a_1x_1 + \ldots + a_nx_n$ вектор-строку $a = (a_1, \ldots, a_n)$. Такой вектор определяет линейный функционал $x \mapsto \sum\limits_{i=1}^{n} a_ix_i$ на $\mathbb{R}^n$. Совокупность линейных функционалов над $\mathbb{R}^n$ обозначим $(\mathbb{R}^n)^*$, а действие линейной функции $a \in (\mathbb{R}^n)^*$ на вектор $x \in \mathbb{R}^n$ будем записывать как $\langle a, x \rangle = a_1x_1 + \ldots + a_nx_n$.

Пусть a_{ij}, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, — mn вещественных чисел. Отображение, задаваемое формулой

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_2 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m,$$

называется *линейным отображением из* $\mathbb{R}^n$ *в* $\mathbb{R}^m$. Такому отображению будем ставить в соответствие таблицу

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix},$$

из m строк и n столбцов. Её называют *матрицей размера* $m \times n$. Она определяет линейное отображение из $\mathbb{R}^n$ в $\mathbb{R}^m$.

(C) Пусть X и Y — два векторных пространства.

Определение 5 (линейное отображение). Отображение $A\colon X \to Y$ называется *линейным*, если для любых x, x' из X и любых α, α' из $\mathbb{R}$ выполнено равенство $A(\alpha x + \alpha' x') = \alpha Ax + \alpha' Ax'$. Линейные отображения называют ещё *операторами*. Совокупность линейных непрерывных отображений из банахова пространства $(X, \|\cdot\|_X)$ в $(Y, \|\cdot\|_Y)$ обозначают $\mathcal{L}(X, Y)$. Совокупность *линейных функционалов* (т. е. $\mathcal{L}(X, \mathbb{R})$) называют *сопряжённым пространством* с X и обозначают X^*.

4. Математический анализ. Дифференцируемость

Математический анализ — фундаментальнейшая ветвь в математике (в частности, потому, что математический анализ явился языком естествознания). Он состоит из двух разделов, именуемых дифференциальным и интегральным исчислениями.

Производная и интеграл — основные понятия математического анализа, они были введены в математику Ньютоном и Лейбницем. Приведём слова Ньютона, где он объясняет суть математического анализа. «Для пояснения искусства анализа, — пишет Ньютон, — нужно привести некоторые примеры задач». Вот эти примеры.

1. Пусть длина пути известна. Нужно узнать скорость в данный момент времени.
2. Пусть известна скорость движения. Надо узнать длину пройденного пути.

(Мы чуть модернизировали текст Ньютона.)

(A) Поясним сказанное (на школьном уровне). Пусть к моменту t экипажем, движущимся по прямолинейной дороге, был преодолён путь $s(t)$, и пусть эта функция, ставящая в соответствие моменту времени t длину пройденного пути $s(t)$, известна. Продумывание первого вопроса, вопроса о скорости, естественно приводит к идее, что скорость $v(\tau)$ экипажа в некоторый момент времени τ (или, как принято говорить, *мгновенная скорость*) примерно равна *средней скорости*

$$\frac{s(\tau + \Delta t) - s(\tau)}{\Delta t}$$

на малом участке времени от τ до $\tau + \Delta t$, а сама скорость равна *пределу этого отношения при Δt, стремящемся к нулю*. Математики пишут при этом

$$v(\tau) = \lim_{\Delta t \to 0} \frac{s(\tau + \Delta t - s(\tau)}{\Delta t}$$

и говорят, что скорость — это *производная пути по времени*. Ньютон обозначил этот предел точкой над s: $v(\tau) = \dot{s}(\tau)$. Итак, *производная по Ньютону — это скорость*.

Если исходной позицией для Ньютона было естествознание, то Лейбниц во многом исходил из геометрии. При изложении концепции Лейбница представляется естественным использовать более привычные обозначения, когда аргумент обозначается буквой x, а функция — буквой f. Тогда производная функции f в точке $\hat{x}$ (её стали обозначать $f'(\hat{x})$) есть предел

$$\lim_{x \to \hat{x}} \frac{f(x) - f(\hat{x})}{x - \hat{x}},$$

а с геометрической точки зрения — это тангенс угла наклона касательной к графику функции f в точке $\hat{x}$. Лейбниц осознал важнейшее свойство производной: она задаёт *наилучшую локальную линейную аппроксимацию нелинейной функции*. Это свойство допускает конечномерное и бесконечномерное толкование понятия производной, которое составляет основу институтского и университетского уровней дифференциального исчисления. Приведём определение производной на этих уровнях.

Пусть X — это $\mathbb{R}$ для (A), $\mathbb{R}^n$ для (B) и нормированное пространство $(X, \|\cdot\|_X)$ для (C), Y — это $\mathbb{R}$ для (A), $\mathbb{R}^m$ для (B) и нормированное пространство $(Y, \|\cdot\|_Y)$ для (C) и V — окрестность точки $\hat{x}$ в X.

Определение 6 (дифференцируемость). Отображение $F\colon V \to Y$ называется *дифференцируемым в точке* $\hat{x}$, если существует линейный непрерывный оператор $A\colon X \to Y$, такой что для любого числа $\varepsilon > 0$ найдётся такое число $\delta > 0$, что из неравенства $d_X(x, \hat{x}) < \delta$ следует, что $d_Y(F(x) - Ax, F(\hat{x}) - A\hat{x}) < \varepsilon d_X(x, \hat{x})$.

Оператор A этим соотношением определяется однозначно, его называют *производной отображения F в точке $\hat{x}$* и обозначают $F'(\hat{x})$. Если F дифференцируемо в точке $\hat{x}$, мы пишем $F \in D^1(\hat{x}, X)$.

Линейное отображение $x \mapsto F'(\hat{x})x$ называют *дифференциалом F в точке $\hat{x}$*.

(C) Дадим ещё определение строгой дифференцируемости, которое, как это будет видно в дальнейшем, оказывается достаточно мотивированным. Отображение F называется *строго дифференцируемым в точке* $\hat{x}$, если оно дифференцируемо в $\hat{x}$ и для любого числа $\varepsilon > 0$ найдётся такое число $\delta > 0$, что из неравенств $\|x_i - \hat{x}\|_X < \delta$, $i = 1, 2$, следует, что $\|F(x_1) - F(x_2) - F'(\hat{x})(x_1 - x_2)\|_Y < \varepsilon\|x_1 - x_2\|_X$. Если F строго дифференцируемо в точке $\hat{x}$, мы пишем $F \in SD^1(\hat{x}, X)$.

5. Отделимость

В двадцатом веке выделился особый раздел в математике, где изучают выпуклые объекты в векторных пространствах: выпуклые множества, выпуклые функции и выпуклые экстремальные задачи. Этот раздел, получивший название выпуклого анализа, весьма важен, и его разумно ввести в математическое образование любого уровня. Основным понятием выпуклого анализа является понятие отделимости.

Пусть X — векторное пространство, у которого имеется нетривиальное сопряжённое пространство X^* (в том смысле, что для любого $x \in X$ найдётся $x^* \in X^*$, для которого не равна нулю величина $\langle x^*, x\rangle$ — действие функционала x^* на элемент $x \in X$, и для любого $x^* \in X^*$ найдётся элемент $x \in X$ с тем же свойством). Разумеется, таковы $\mathbb{R}^n$ и $(\mathbb{R}^n)^*$, банахово пространство также обладает нетривиальным сопряжённым. Широкий класс векторных пространств с нетривиальными сопряжёнными называется классом локально выпуклых топологических векторных пространств. Он и есть истинный «ареал» выпуклого анализа (также как класс банаховых пространств — истинный ареал гладкого анализа), но нам всюду будет хватать банаховых пространств и при рассмотрениях выпуклых объектов.

Определение 7 (отделимость). Пусть X — векторное пространство, X^* — сопряжённое с ним пространство, A и B — два подмножества X. Говорят, что *множества A и B отделимы*, если существует элемент x^* из сопряжённого пространства X^* (отличный от нуля), для которого

$$\inf_{x\in A}\langle x^*, x\rangle \geqslant \sup_{\xi\in B}\langle x^*, \xi\rangle.$$

Если выполнено неравенство

$$\inf_{x\in A}\langle x^*, x\rangle > \sup_{\xi\in B}\langle x^*, \xi\rangle,$$

то множества A и B называют *строго отделимыми*.

6. Два примера банаховых пространств

1. Пространство непрерывных функций на отрезке $[a,b]$. Совокупность непрерывных функций на отрезке $[a,b]$ обозначают $C([a,b])$. Это векторное пространство, которое становится нормированным относительно нормы

$$\|x(\cdot)\|_{C([a,b])} = \max_{t\in[a,b]}|x(t)|.$$

Это *универсальное пространство* для всех сепарабельных банаховых пространств: любое сепарабельное банахово пространство изометрично некоторому подпространству пространства $C([a,b])$.

2. Пространство l_2 состоит из последовательностей $x = (x_1, x_2, \ldots)$, для которых $\sum\limits_{k\in\mathbb{N}} x_k^2 < \infty$, со скалярным произведением

$$\langle x, y\rangle = \sum_{k\in\mathbb{N}} x_k y_k.$$

Конечномерное пространство, состоящее из векторов $x = (x_1, \ldots, x_n)$, со скалярным произведением

$$\langle x, y\rangle = \sum_{k\in\mathbb{N}}^{n} x_k y_k$$

обозначают l_2^n или $\mathbb{E}^n$. Пространство l_2 — *уникальное пространство*: любое его бесконечномерное подпространство ему изометрично.

Векторное пространство $(X, \langle\cdot,\cdot\rangle)$, оснащённое скалярным произведением $\langle\cdot,\cdot\rangle\colon X\times X\to\mathbb{R}$ со свойствами

$$\langle x, x\rangle \geqslant 0 \ \text{ и } \ \langle x, x\rangle = 0 \iff x = 0,$$
$$\langle x, y\rangle = \langle y, x\rangle, \quad \langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$$

для всех $x, y \in X$, $\alpha, \beta \in \mathbb{R}$ в конечномерном случае называется *евклидовым пространством*, а в бесконечномерном случае при полноте пространства относительно нормы $\|x\|_X = \sqrt{\langle x, x\rangle}$ называется *гильбертовым пространством*. Любое бесконечномерное сепарабельное пространство изометрично l_2.

Линейный оператор, действующий из гильбертова пространства в себя называется *компактным*, если он замкнутые шары отображает в компактные множества, и *предкомпактным*, если замыкание образа компактно.

7. Интегрируемость

Рассмотрим теперь вторую задачу Ньютона. Пусть снова экипаж движется по прямолинейной дороге, нам известна мгновенная скорость экипажа $v(t)$ в любой момент времени t на отрезке времени $[t_0, t_1]$, и мы хотели бы найти длину пути, преодолённого экипажем за этот период времени. Естественно допустить, что на малом отрезке времени $\Delta = [\tau, \tau + \Delta\tau]$ скорость меняется мало и путь примерно равен длине отрезка времени $|\Delta\tau|$, умноженной на значение скорости в какой-то момент времени в этом промежутке. Таким образом, весь путь примерно равен сумме $v(\theta_1)|\Delta_1| + \ldots + v(\theta_N)|\Delta_N|$, где $\{\Delta_i\}_{i=1}^N$ — разбиение отрезка $[t_0, t_1]$ на N отрезочков и θ_i — какая-то точка отрезка Δ_i. Это приводит к определению интеграла по Риману, остаётся только перейти к пределу при стремлении $\max|\Delta_i|$ к нулю. Дадим более чёткие определения.

Пусть f — ограниченная функция на конечном отрезке $[a, b]$, D (от англ. «division») — разбиение отрезка точками x_j: $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_N = b$,

$$M(D) = \max\{|x_j - x_{j-1}|,\ 1 \leqslant j \leqslant N\} —$$

«диаметр» разбиения D, S (от англ. «selection») — выбор по точке ξ_j в каждом отрезочке $\Delta_j = [x_i, x_{j-1}]$, $1 \leqslant j \leqslant N$, $|\Delta_j| = x_{j+1} - x_j$. Выражение

$$R(f, D, S) = \sum_{j=1}^{n} f(\xi_j)|\Delta_j|$$

называется *римановой суммой функции f по разбиению D при выборе S*. Число I (если таковое существует) называется *определённым интегралом* (*Римана*) *от функции f на $[a, b]$* (обозначение: $\int\limits_a^b f(x)\,dx$), если для любого $\delta > 0$ найдётся такое число $\varepsilon > 0$, что для любого D с $M(D) < \varepsilon$ при любом выборе S выполнено неравенство

$$|R(f, D, S) - I| < \varepsilon.$$

Сумма

$$\overline{R}(f, D) = \sum_{j=1}^{n} \sup_{x \in \Delta_j} f(x)|\Delta_j|$$

называется *верхней суммой Дарбу*, а сумма

$$\underline{R}(f,D)=\sum_{j=1}^{n}\inf_{x\in\Delta_j} f(x)|\Delta_j| -$$

нижней суммой Дарбу функции f *по разбиению* D *при выборе* S. Очевидны неравенства

$$\underline{R}(f,D)\leqslant R(f,D,S)\leqslant \overline{R}(f,D)$$

для всех S.

Лемма I.1. *Пусть*

$$D=\bigcup_{i=1}^{N}\Delta_i \quad \text{и} \quad D'=\bigcup_{j=1}^{N'}\Delta'_j -$$

два разбиения $[a,b]$. *Тогда* $\underline{R}(f,D)\leqslant \overline{R}(f,D')$.

Доказательство. Обозначим через Δ_{ij} общую часть отрезков Δ_i и Δ'_j,

$$m_{ij}=\min_{x\in\Delta_{ij}} f(x),\quad M_{ij}=\max_{x\in\Delta_{ij}} f(x).$$

через m'_{ij} и M'_{ij} — аналогичные величины для разбиения D' (при этом мы пользуемся теоремой Вейерштрасса о достижении непрерывной функцией верхней и нижней граней). Имеем:

$$\underline{R}(f,D)\overset{\text{def}}{=}\sum_{i=1}^{N} m_i|\Delta_i| \overset{\text{def}}{=}\sum_{i=1}^{N}\sum_{j=1}^{N'} m_i|\Delta_{ij}| \overset{m_i\leqslant m_{ij}}{\leqslant} \sum_{i=1}^{N}\sum_{j=1}^{N'} m_{ij}|\Delta_{ij}| \overset{m_{ij}\leqslant M_{ij}}{\leqslant}$$

$$\leqslant \sum_{i=1}^{N}\sum_{j=1}^{N'} M_{ij}|\Delta_{ij}| \overset{M_{ij}\leqslant M'_i}{\leqslant} \sum_{j=1}^{N'} M'_j|\Delta'_j| \overset{\text{def}}{=} \overline{R}(f,D'). \quad \square$$

Теорема о существовании определённого интеграла. *Определённый интеграл по отрезку от непрерывной функции существует.*

Доказательство. Верхние и нижние суммы Дарбу в силу теоремы Вейерштрасса (о ней и теореме Кантора говорится в школе и будет рассказано в этой работе чуть далее) ограничены снизу и сверху соответственно:

$$\overline{R}(f,D)\geqslant \Big(\min_{x\in[a,b]} f(x)\Big)(b-a),\quad \underline{R}(f,D)\leqslant \Big(\max_{x\in[a,b]} f(x)\Big)(b-a).$$

Значит, в силу полноты $\mathbb{R}$ они имеют (соответственно) нижнюю $\underline{I}=\inf\limits_{D}\underline{R}(f,D)$ и верхнюю $\overline{I}=\sup\limits_{D}\overline{R}(f,d)$ грани. По лемме $\underline{I}\leqslant\overline{I}$. Эти грани совпадают в силу теоремы Кантора (так как для любого $\delta>0$ найдётся $\varepsilon>0$, такое что $M(D)<\delta$, следовательно, $\overline{R}(f,D)-\underline{R}(f,D)<\varepsilon$). Из определения вытекает, что это общее число $I=\underline{I}=\overline{I}$ и есть искомый определённый интеграл. $\square$

Таким образом, геометрический смысл определённого интеграла

$$\int\limits_a^b f(x)\,dx$$

от положительной функции f — это площадь под графиком функции f.

II. Теория

Предварительные результаты

Предложение II.1. *Из любой последовательности $\{x_k\}_{k\in\mathbb{N}}$ точек отрезка в $\mathbb{R}$* (для (A))*, точек ограниченного замкнутого множества в $\mathbb{R}^n$* (для (B)) *можно извлечь сходящуюся подпоследовательность. Иначе говоря, отрезок вещественной прямой* (для (A)) *и ограниченные замкнутые подмножества в $\mathbb{R}^n$* (для (B)) *являются компактами.*

Доказательство в случае (A) осуществляется методом деления отрезка пополам, а в случае (B), после помещения множества в n-мерный куб, методом деления куба на 2^n равных частей. Выбрав любой элемент последовательности в качестве первого, производим деление, выбираем ту часть, где имеется бесконечное число элементов последовательности, и там выбираем один из элементов с номером, большим, чем номер первого элемента, в качестве второго и затем повторяем всю процедуру. Использование аксиомы Кантора о полноте приводит к доказательству требуемого утверждения. □

Теорема Вейерштрасса о достижении максимумов и минимумов. *Непрерывная функция на отрезке* (для (A))*, на ограниченном замкнутом подмножестве в $\mathbb{R}^n$* (для (B))*, на компакте* (для (C)) *достигает максимума и минимума.*

Доказательство можно вести одновременно на школьном, институтском и университетском уровне. Пусть f — непрерывная вещественная функция на отрезке (на ограниченном замкнутом подмножестве $\mathbb{R}^n$, на компакте) C. Если предположить, что $\sup\limits_{x\in C} f(x) = +\infty$, то по определению существует последовательность $\{x_k\}_{k\in\mathbb{N}}$ элементов из C, такая что $f(x_k) \to +\infty$. Из предложения II.1 следует, что найдётся точка $\hat{x}$, являющаяся пределом подпоследовательности выбранной последовательности. Но это противоречит определению непрерывности функции f в точке $\hat{x}$ (ибо f должна быть ограничена в окрестности точки $\hat{x}$). Случай $\inf\limits_{x\in X} f(x) = -\infty$ рассматривается аналогично.

Докажем существование максимума. Пусть $M = \sup\limits_{x\in X} f(x)$. Если допустить, что $f(x) \neq M$ для всех x, то по определению верхней грани существует последовательность $\{\xi_k \in X\}_{k\in\mathbb{N}}$, такая что $f(\xi_k) \to M$. Из определения компакта следует, что существует точка $\hat{\xi}$, являющаяся пределом подпоследовательности выбранной последовательности. Из определения секвенциальной непрерывности функции f в точке $\hat{\xi}$ следует, что $f(\hat{\xi}) = M$. Случай минимума разбирается аналогично. $\square$

Теорема о равномерной непрерывности. *Непрерывная функция на отрезке* (*для* (A)), *ограниченном замкнутом подмножестве в* $\mathbb{R}^n$ (*для* (B)), *на компакте в нормированном пространстве* (*для* (C)) *равномерно непрерывна.*

У нас принято называть эту теорему теоремой Кантора, Бурбаки называет этот результат теоремой Гейне.

Доказательство одинаково на всех уровнях. Проведём его сразу на компакте. Пусть для некоторого $\varepsilon > 0$ существуют две последовательности точек компакта $\{x_j\}_{j\in\mathbb{N}}$ и $\{\xi_j\}_{j\in\mathbb{N}}$, для которых $|x_j - \xi_j| \to 0$ при $j \to \infty$ в случае $\mathbb{R}$ (или $\mathbb{R}^n$) или $\|x_j - \xi_j\|_X \to 0$ при $j \to \infty$ в нормированном случае, в то время как

$$\text{расстояние между точками } f(x_j) \text{ и } f(\xi_j) \text{ больше } \varepsilon. \tag{$*$}$$

Из того, что X — отрезок (замкнутое ограниченное множество в $\mathbb{R}^n$ или компакт), следует, что в X существует точка $\hat{x}$, такая что $x_{k_j} \to \hat{x}$ при $j \to \infty$. Но тогда и $\xi_{k_j} \to \hat{x}$ при $j \to \infty$. При этом соотношение $(*)$ противоречит определению непрерывности f в точке $\hat{x}$. $\square$

Первая теорема отделимости. *Пусть* X — *пространство,* $X = \mathbb{R}^2$ (*для* (A)), $X = \mathbb{R}^n$ (*для* (B)), $(X, \|\cdot\|_X)$ — *банахово пространство* (*или локально выпуклое топологическое векторное пространство*) (*для* (A)), A *и* B — *выпуклые подмножества в* X, *причём внутренность* A *непуста и не пересекается с* B. *Тогда* A *и* B *отделимы.*

Доказательство. Рассмотрим разность

$$C = A - B = \{x = a - b \mid a \in A,\ b \in B\}.$$

Этим рассмотрением теорема редуцируется к случаю отделимости множества C с непустой внутренностью от нуля, не принадлежащего этой внутренности. Приведём эскиз доказательства редуцированного варианта теоремы, предполагая, что нуль принадлежит замыканию C (случай, когда нуль не принадлежит замыканию, доказывается проще).

В случае (A) пусть $C \subset \mathbb{R}^2$ — плоское выпуклое множество, содержащее внутреннюю точку, такое что нуль не является внутренней точкой этого множества. Тогда и замыкание множества C обладает тем же свойством: нуль не станет внутренней точкой замыкания, так что сразу можно считать, что C замкнуто. Пусть c — внутренняя точка C. Значит, в C содержится круг

$$U_{\mathbb{R}^2}(c, \varepsilon) = \{x \in c \mid |x - c| < \varepsilon\}.$$

Проведём через нуль и c прямую l. Эта прямая распадается на два открытых луча: луч l_+, содержащий c, и противоположный луч l_-, не содержащий точек из C (иначе точка нуль была бы внутренней точкой C). Любую точку ξ

луча l_- можно отделить от C следующим образом: из теоремы Вейерштрасса следует, что в замкнутом множестве C найдётся точка η, ближайшая к ξ, и тогда перпендикуляр к отрезку $[\xi, \eta]$, проведённый через точку η, отделит C от ξ. Далее надо взять последовательность точек луча l_-, стремящуюся к нулю, и из отделяющих прямых выбрать предельную.

В случае (B) применяем индукцию. Пусть теорема доказана для m-мерных выпуклых тел. Докажем её для тел $m+1$-мерных. Прямую, проходящую через нуль и точку c, включим в m-мерное пространство L_m. По предположению индукции существует опорная L_{m-1}-мерная плоскость $\hat{L}_{m-1}$, опорная к сечению $C \cap L_m$. Профакторизуем пространство $\mathbb{R}^{m+1}$, содержащее C, по $\hat{L}_{m-1}$. Фактором является плоское множество, имеющее внутреннюю точку. По доказанному существует опорная прямая $\hat{l}$ к этому плоскому множеству. Тогда сумма $\hat{L}_{m-1}+\hat{l}$ есть искомая опорная m-мерная плоскость.

В случае (C) надо применить лемму Цорна (это аксиома для упорядоченных систем, подробное описание опускаем, отсылая к литературе). Пусть дано семейство вложенных подпространств, проходящих через точку нуль и не пересекающих внутренность выпуклого тела. Объединение этих подпространств будет мажорантой этой совокупности. По лемме Цорна имеется максимальное подпространство, обладающее перечисленными свойствами. Замыкание этого подпространства также не будет пересекаться с внутренностью. Если линейная оболочка этого подпространства и точки c не совпадает со всем X, рассмотрим оболочку этого пространства и элемента, не принадлежащего линейной оболочке его с внутренней точкой c. Профакторизуем пространство X по максимальному элементу, получим плоскую проблему, разрешив которую, придём к противоречию. $\square$

Следствие (вторая теорема отделимости). *Точка, не принадлежащая выпуклому замкнутому множеству, строго отделима от него.*

Доказательство. Надо взять выпуклую окрестность точки, не пересекающуюся со множеством, и отделить их по первой теореме отделимости. $\square$

Основная часть

Решение линейных уравнений входило в начальное математическое образование во все времена. Теория линейных уравнений составляет содержание начальных курсов математики в институтах и университетах. Решение систем линейных уравнений — база огромной доли приложений современной математики. На линейной аппроксимации основано дифференциальное исчисление. В этом заключается мотивировка этой темы в математическом образовании. Далее проходится путь от истоков до начала XX века.

1. Теория линейных уравнений

Теорема II.1 (теорема Фредгольма). *Пусть X — пространство, $X = \mathbb{R}^2$ (для* (A)*), $X = \mathbb{R}^n$ (для* (B)*), $(X, \|\cdot\|_X)$ — банахово пространство (для* (C)*),*

$A\colon X \to X$ — линейный оператор (в банаховом бесконечномерном случае — линейный непрерывный оператор, являющийся суммой единичного и компактного операторов)[1]. *Тогда имеет место альтернатива Фредгольма: либо уравнение $Ax = y$ разрешимо для любого $y \in X$, либо однородное уравнение $Ax = 0$ имеет ненулевое решение.*

В бесконечномерном случае, чтобы избежать некоторых затруднений, мы доказываем теорему для $X = l_2$ и $A = I + B$, где I — единичный оператор, а матрица оператора B квадратично суммируема: $\sum\limits_{i,j \in \mathbb{N}} b_{ij}^2 < \infty$. Этот результат был доказан профессором Стокгольмского университета И. Фредгольмом в 1903 г.

Доказательство складывается из двух частей.

1. Докажем, что если $\operatorname{Im} A = X$, то $\operatorname{Ker} A = 0$.

Начнём с (A). Допустим, что $\operatorname{Im} A = \mathbb{R}^2$, но найдётся вектор $y^1 \neq 0$, такой что $Ay^1 = 0$. Тогда найдём такой вектор y^2, что $Ay^2 = y^1$, и такой вектор y^3, что $Ay^3 = y^2$. Три вектора в $\mathbb{R}^2$ линейно зависимы, значит, найдутся числа a_1, a_2, a_3, не равные одновременно нулю и такие, что $a_1y^1 + a_2y^2 + a_3y^3 = 0$ (этот факт легко сводится к простому утверждению, что два однородных линейных уравнения с тремя неизвестными имеют ненулевое решение). Подействовав на равенство оператором A^2, получим, что $a_3y^1 = 0$, откуда следует, что $a_3 = 0$. Аналогично доказываем, что $a_2 = a_1 = 0$, и приходим к противоречию с тем, что не все a_i равны нулю.

Доказательство в случае (B) совершенно аналогично: строим вектор $y^1 \neq 0$, для которого $Ay^1 = 0$, и далее $n{+}1$ вектор $Ay^{i+1} = y^i$, $i = 1, \ldots, n$. Они линейно зависимы, следовательно, для не равных одновременно нулю чисел $a_1, \ldots, a_{n+1}$ выполнится соотношение $\sum\limits_{i=1}^{n+1} a_iy^i = 0$. Подействовав на это соотношение последовательно операторами $A^n, \ldots, A$, приходим к противоречию с тем, что a_i не все нули.

Для доказательства теоремы на университетском уровне (C) потребуются «выходы в бесконечномерное пространство».

Лемма II.1 (о последовательности единичных векторов). *Не существует бесконечной последовательности единичных векторов $\{f^i\}_{i \in \mathbb{N}}$, такой что расстояния между Bf^i и Bf^j при $i \neq j$ превосходят некоторое положительное число.*

Это первый выход в бесконечномерное пространство; доказательство проведём в разделе «Дополнение».

Используя лемму II.1, переходим к доказательству первого утверждения теоремы. Допустим, что $\operatorname{Im} A = l_2$, но найдётся вектор $e^1 \neq 0$, такой что $Ae^1 = 0$ и

[1]Причина условия вскроется при доказательстве.

при этом можно считать, что $\|e^1\|_X = 1$. Обозначим через $L_k = \operatorname{Ker} A^k$. Доказано, что L_1 нетривиально. Решение уравнения $Ae^2 = e^1$ и далее $Ae^{k+1} = e^k$ показывает, что все L_k нетривиальны. При этом из леммы II.1 вытекает, что все L_k конечномерны (иначе мы построили бы в L_k ортонормированную систему из любого числа векторов g^k, таких что $\|Bg^i - Bg^j\|_X = \|g^i - g^j\|_X = \sqrt{2}$). Рассмотрим последовательность Be^n. Пусть $m > n$. Легко понять, что $z_n = e^n - Ae^n + Ae^m \in \in \operatorname{Ker} A^{m-1}$. Значит, $\|Be^m - Be^n\|_X = \|z_n - e_m\|_X \geqslant 1/2$. Пришли к противоречию с компактностью оператора B. Следовательно, $\operatorname{Ker} A = 0$.

2. Докажем, что если $AX \neq X$, то $\operatorname{Ker} A \neq 0$ (рис. 1).

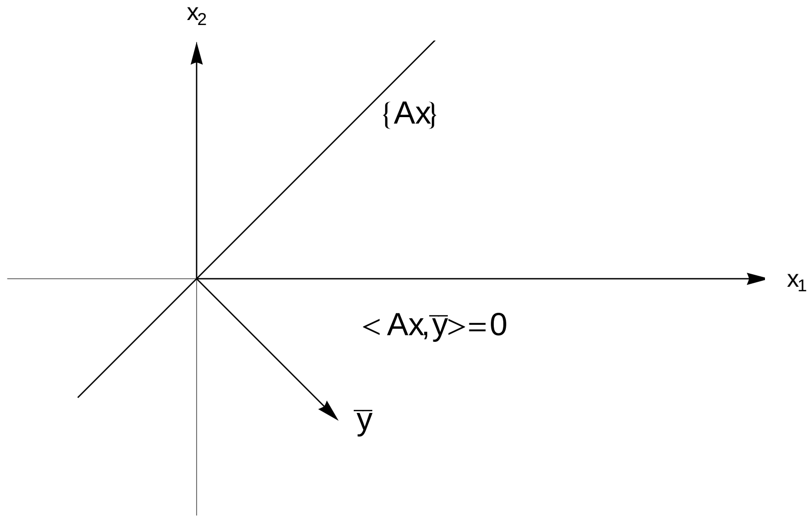

Рис. 1

Снова начнём с (А). Если образ A не есть $\mathbb{R}^2$, то этот образ либо нуль, либо прямая, проходящая через нуль. В первом случае всё доказано, ибо тогда любой вектор принадлежит ядру. Во втором случае к прямой, являющейся образом A, найдётся ортогональный вектор $\bar{y}$:

$$0 = \langle \bar{y}, Ax \rangle = \langle A^{\mathrm{T}} \bar{y}, x \rangle$$

(см. рис. 1). Значит, $\langle A^{\mathrm{T}} \bar{y}, x \rangle = 0$ для всех $x \in \mathbb{R}^2$, следовательно, $A^{\mathrm{T}} \bar{y} = 0$, т. е. $\bar{y}$ — решение однородного уравнения для транспонированной матрицы. По утверждению, доказанному чуть выше, оператор A^{T} не сюръективен. Применив к нему только что проведённое рассуждение, придём к тому, что однородное уравнение $A^{\mathrm{TT}} = A = 0$ имеет ненулевое решение.

Доказательство в случае (В) аналогично. Если $\operatorname{Im} A \neq \mathbb{R}^n$, то этот образ — собственное подпространство $\mathbb{R}^n$. Собственное подпространство задаётся системой линейных однородных уравнений. Пусть $\sum\limits_{k=1}^{n} \bar{y}_k x_k = 0$ — одно из них, т. е.

$$0 = \langle \bar{y}, Ax \rangle = \langle A^{\mathrm{T}} \bar{y}, x \rangle.$$

Значит, $\langle A^{\mathrm{T}} \bar{y}, x \rangle = 0$ для всех $x \in \mathbb{R}^n$, следовательно, $A^{\mathrm{T}} \bar{y} = 0$, т. е. $\bar{y}$ — решение однородного уравнения для транспонированной матрицы. В силу первого утверждения альтернативы Фредгольма оператор A^{T} не сюръективен. Применив к нему только что проведённое рассуждение, придём к тому, что однородное уравнение $A^{\mathrm{TT}} = A = 0$ имеет ненулевое решение. Альтернатива Фредгольма доказана.

В бесконечномерном случае (C) доказательство аналогично, но нужны два выхода в бесконечномерность.

Лемма II.2 (о замкнутости). *Образ оператора $A = I + B$, где B квадратично суммируем, замкнут.*

Лемма II.3 (о перпендикуляре). *Пусть X — гильбертово пространство, L — собственное замкнутое подпространство X. Тогда существует элемент $y \neq 0$, перпендикулярный L.*

Это технические леммы, которые мы также докажем в разделе «Дополнение». Ещё потребуется лемма о нетривиальности аннулятора, согласно которой для замкнутого подпространства гильбертова пространства существует ортогональный вектор. Эта лемма является простым следствием второй теоремы отделимости (и в более общем виде доказывается далее).

Далее для доказательства второго пункта теоремы надо повторить рассуждения для (A) или (B). □

2. Разрешимость нелинейных уравнений

Нелинейные уравнения возникают почти во всех математических моделях, описывающих явления и процессы: в естествознании, инженерии, теории управления, экономике и т. п., и нахождение приближённых решений таких уравнений — одна из основных задач математики и её приложений. Но начать придётся с разрешимости линейных уравнений.

Лемма о правой обратимости в линейном случае. *Пусть пространства X и Y — это либо $\mathbb{R}$ и $\mathbb{R}$ (для (A)), либо $\mathbb{R}^n$ и $\mathbb{R}^m$ (для (B)), либо банаховы пространства (для (C)) и A — линейный (непрерывный в банаховом случае) сюръективный оператор из X в Y. Тогда существуют отображение $R\colon Y \to X$ (правое обратное) и число $\gamma > 0$, такие что $AR(y) = y$ и $\|Ry\|_X \leqslant \gamma \|y\|_X$ для всех $y \in Y$.*

Доказательство. В случае (A) «оператор A» — это просто число $A \neq 0$ и $R = A^{-1}$. В случае (B) надо для системы базисных векторов $\{e_j\}_{j=1}^m$ в $\mathbb{R}^m$ найти (по условию существующие) векторы $\{f_k\}_{k=1}^m$, $f_k \in \mathbb{R}^n$, такие что $Af_k = e_k$. Для каждого вектора

$$y = \sum_{k=1}^{m} y_k e_k \in \mathbb{R}^m$$

(здесь y_k — числа) положим

$$R(y) = \sum_{k=1}^{m} y_k f_k.$$

Тогда равенство $AR(y) = y$ следует из линейности оператора A, а второе свойство следует из очевидных оценок. В случае (C) по теореме Банаха об обратном операторе (см. учебник по функциональному анализу, например, А. Н. Колмогорова и С. В. Фомина; в «Дополнении» мы выведем лемму в случае (C) непосредственно) найдётся такое число $\delta > 0$, что $\Lambda U_X(0,1) \supset U_Y(0,\delta)$ ($U_Z(z_0, r) = \{z \in Z \mid \|z - z_0\|_Z < r\}$ — открытый шар в пространстве Z с центром в z_0 радиуса r). Это значит, что для любого $y \in U_Y(0,\delta)$ найдётся такой элемент $x(y) \in U_X(0,1)$, что $\Lambda x(y) = y$, и остаётся положить

$$R(y) = \frac{2\|y\|_Y}{\delta} x\left(\frac{\delta y}{2\|y\|_Y}\right)$$

и проверить выполнимость утверждений леммы. □

Доказательство следующей основной в этом разделе теоремы одинаково в одномерном, конечномерном и бесконечномерном случаях, и потому мы формулируем её, как и лемму, сразу для всех этих трёх случаев.

Теорема II.2 (о правом обратном отображении). *Пусть либо* $X = \mathbb{R}$ с $\|x\|_X = |x|$, $Y = \mathbb{R}$ с $\|y\|_Y = |y|$ *(для* (A)), *либо* $X = \mathbb{R}^n$ с $\|x\|_X = \Big(\sum\limits_{i=1}^{n} x_i^2\Big)^{1/2}$, $Y = \mathbb{R}^m$ с $\|y\|_X = \Big(\sum\limits_{i=1}^{m} y_i^2\Big)^{1/2}$ *(для* (B)), *либо* $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ — *банаховы пространства (для* (C)), V — *окрестность точки* $\hat{x} \in X$, $F\colon V \to Y$ — *нелинейное отображение. Для решения уравнения* $F(x) = y$, $y \in Y$, *применяем модифицированный метод Ньютона*:

$$x_k = x_{k-1} + R\big(y - F(x_{k-1})\big), \quad k \in \mathbb{N}, \quad \hat{x} = x_0, \tag{II.1}$$

где R — *правый обратный для линейного сюръективного оператора* $A\colon Y \to X$. *Пусть при этом для чисел* ϑ, $0 < \vartheta < 1$, *и* $\delta > 0$ *выполнено неравенство*

$$\|F(x) - F(x') - A(x - x')\|_Y \leqslant \frac{\vartheta}{\gamma}\|x - x'\|_X \tag{II.2}$$

для любых x, x', *для которых* $\|x - \hat{x}\|_X < \delta$, $\|x' - \hat{x}\|_X < \delta$. *Тогда для любого вектора* $y \in Y$, *отстоящего по норме в* Y *от* $F(\hat{x})$ *не больше чем на* $\big(\delta(1-\vartheta)\big)/\gamma$, *последовательность* (II.1) *будет стремиться к вектору* $\varphi(y)$, *отстоящему от* $\hat{x}$ *по норме в* X *не больше чем на* δ, *такому что* $F\big(\varphi(y)\big) = y$ *и при этом* $\|\varphi(y) - \hat{x}\|_X \leqslant K\|y - F(\hat{x})\|_Y$, где $K = \gamma/(1 - \vartheta)$.

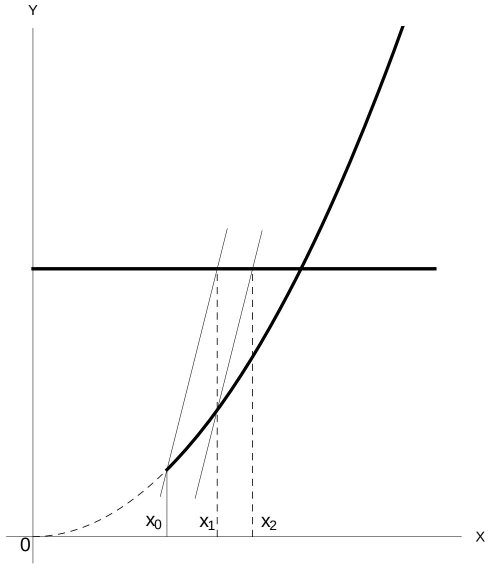

Рис. 2

Этот результат был доказан профессором Чикагского университета Л. Грейвсом в 1950 г.

Доказательство иллюстрирует рис. 2. Докажем, что

а) элементы x_k, которые строятся согласно (II.1), для всех $k \geqslant 0$ лежат в $U_X(0,\delta)$;

б) последовательность $\{x_k\}_{k\geqslant 0}$ фундаментальна.

Утверждение а) докажем по индукции. Начальный элемент x_0 принадлежит $U_X(0,\delta)$ по определению. Пусть $x_s \in U_X(0,\delta)$, $1 \leqslant s \leqslant k$. Докажем, что $x_{k+1} \in U_X(0,\delta)$. Используя последовательно (II.2) и равенство

$$A(x_s - x_{s-1}) - y + F(x_{s-1}) = 0,$$

следующее из (II.2), после чего применяя (II.1) и затем повторяя процедуру, получаем

$$\begin{aligned}\|x_{k+1} - x_k\|_X = \|R(y - F(x_{k-1}))\|_X \leqslant \gamma\|y - F(x_k)\|_Y =\\ = \gamma\|y - F(x_k) - y + F(x_{k-1}) + A(x_k - x_{k-1})\|_Y \leqslant\\ \leqslant \theta\|x_k - x_{k-1}\|_X \leqslant \ldots \leqslant \theta^n\|x_1 - x_0\|_X.\end{aligned} \tag{II.3}$$

Теперь, используя неравенство для модулей, формулу для суммы геометрической прогрессии, (II.3) при $k = 1$ и выбор y, получаем, что

$$\|x_{k+1} - \hat{x}\|_X \leqslant \|x_{k+1} - x_k\|_X + \ldots + \|x_1 - \hat{x}\|_X \leqslant$$
$$\leqslant (\theta^k + \theta^{k-1} + \ldots + 1)\|x_1 - \hat{x}\|_X < \frac{1}{(1-\theta)}\|y - F(\hat{x})\|_Y \leqslant \delta,$$

т. е. элементы x_k определены для всех k.

Докажем утверждение б). Для любых $n, m \in \mathbb{N}$ имеем

$$\|x_{k+l} - x_k\|_X \leqslant \|x_{k+l} - x_{k+l-1}| + \ldots + |x_{k+1} - x_k\|_X \leqslant$$
$$\leqslant \left(\sum_{i=k}^{k+l} \theta^i\right)\|x_1 - \hat{x}\|_X < \frac{\theta^k}{1-\theta}\|y - F(\hat{x})\|_Y \leqslant \delta\theta^k, \quad \text{(II.4)}$$

откуда вытекает, что $\{x_k\}_{n\in\mathbb{N}}$ — фундаментальная последовательность. Значит, она сходится. Обозначим её предел $\varphi(y)$. Переход к пределу в неравенстве (II.2) (который существует из-за непрерывности F, следующей из строгой дифференцируемости) приводит к равенству $F\big(\varphi(y)\big) = y$, а переход к пределу в (II.4) с учётом того, что $\|x_1 - \hat{x}\|_X \leqslant \gamma\|y - F(\hat{x})\|_Y$, обеспечивает неравенство $\|\varphi(y) - \hat{x}\|_X \leqslant K\|y - F(\hat{x})\|_Y$ с $K = \gamma/(1-\theta)$. □

3. Начала теории экстремума

Введение начал теории экстремальных задач в математическое образование можно мотивировать тем, что вариационное исчисление (благодаря тому что большинство законов естествознания основывается на вариационных принципах) является базисом математического естествознания, оптимальное управление лежит в основе теории управления в производстве и движении динамических систем, а выпуклые проблемы оптимизации играют фундаментальную роль в математической экономике.

Теорема II.3 (теорема Ферма и правило множителей Лагранжа). *Пусть X и Y — пространства, $X = Y = \mathbb{R}^2$ (для* (A)*), $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ (для* (B)*), $(X, \|\cdot\|_X)$ и $(Y, \|\cdot\|_Y)$ — банаховы пространства (для* (C)*), V — окрестность точки $\hat{x} \in X$, $f_0\colon V \to \mathbb{R}$, $F\colon V \to Y$, при этом функция f_0 и отображение F строго дифференцируемы в точке $\hat{x}$ и образ оператора $F'(\hat{x})$ замкнут. Тогда если в точке $\hat{x}$ достигается локальный экстремум в задаче*

$$f_0(x) \to \min, \quad F(x) = 0. \qquad \text{(P)}$$

то

а) *необходимым условием локального экстремума в точке $\hat{x}$ в задаче без ограничений (когда в задаче* P *отображение F отсутствует; здесь достаточно простой дифференцируемости f_0) является условие стационарности: $f_0'(\hat{x}) = 0$ (в одномерном случае для полупрямой $x \geqslant 0$ $f_0'(0) \geqslant 0$);*

б) *необходимое условие локального экстремума в задаче* P *находится в соответствии с принципом Лагранжа. Это означает, что найдутся число $\lambda_0 > 0$ и элемент $\lambda \in Y^*$, не равные одновременно нулю, такие что для функции*

Лагранжа $\mathcal{L}(x,\bar{\lambda}) = \lambda_0 f_0(x) + \langle\lambda, F(x)\rangle$ $(\bar{\lambda} = (\lambda_0,\lambda))$ *выполнено условие стационарности* $\mathcal{L}_x(\hat{x},\bar{\lambda}) = 0$.

Этот результат был доказан профессором Московского университета Л. А. Люстерником в 1934 г.

Доказательство. Утверждение а) следует из определения дифференцируемости. Докажем б).

Лемма о нетривиальности аннулятора. *Если* X *— пространство* $\mathbb{R}^2$ (*для* (A)), *пространство* $\mathbb{R}^n$ (*для* (B)), *локально выпуклое пространство* (*для* (C)) *и* L *— замкнутое собственное подпространство, то аннулятор*

$$L^\perp = \{x^* \in X^* \mid \langle x^*, x\rangle = 0 \text{ для всех } x \in L\}$$

содержит ненулевой элемент.

Доказательство. Надо взять элемент, не принадлежащий L, и строго отделить его от L по второй теореме отделимости. □

Лемма об аннуляторе ядра сюръективного линейного оператора. *Если линейный оператор из одного банахова пространства в другое сюръективен, то аннулятор его ядра есть образ сопряжённого оператора.*

Доказательство. Пусть пространства X и Y банаховы, $\Lambda X = Y$ и $x^* \in \operatorname{Ker}\Lambda^\perp$. Покажем сначала, что образ X при отображении $Mx = (\langle x^*, x\rangle, Ax)$ замкнут в $\mathbb{R}\times Y$. Действительно, если (α, y) принадлежит замыканию MX, то существует последовательность $\{x_n\}_{n\in\mathbb{N}}$, такая что $\langle x^*, x_n\rangle \to \alpha$, $Ax_n \to y$. Положим $h_n = R(Ax_n - y)$, где R — правый обратный для оператора Λ. Из свойств оператора R следует, что h_n стремится к нулю, а значит, $\langle x^*, h_n\rangle \to 0$. Тогда, положив $\xi_n = x_n - h_n$, убеждаемся в том, что $M(\xi_n) \to (\alpha, y)$, т. е. (α, y) принадлежит замыканию MX. Остаётся применить вторую теорему отделимости, отделив точку $(1, 0)$ (не принадлежащую MX, так как $x^* \in \operatorname{Ker}\Lambda^\perp$) от замкнутого множества MX. □

Перейдём к доказательству правила множителей Лагранжа. Если образ оператора $F'(\hat{x})$ — собственное подпространство, надо применить лемму о нетривиальности аннулятора. В случае когда оператор $F'(\hat{x})$ сюръективен, рассмотрим отображение $\mathcal{F} = (f_0, F)\colon V \to \mathbb{R}\times Y$. Имеется альтернатива:

1) $\mathcal{F}'(\hat{x})$ — сюръективное отображение,
2) $\mathcal{F}'(\hat{x})$ не является сюръективным.

В первом случае, применив теорему о правом обратном, приходим к противоречию, во втором случае, используя лемму об аннуляторе ядра сюръективного линейного оператора, приходим к утверждению теоремы. □

4. Теория квадрик

Решению квадратных уравнений учат школьников с давних времён. Приведение квадратичных форм к каноническому виду — обязательная часть институтских и университетских образовательных программ всех уровней, от аналитической геометрии до дифференциальных уравнений с частными производными. Доказываемая в этом разделе теорема Гильберта—Шмидта имеет фундаментальные приложения и в самой математике, и в математической физике, и в квантовой механике.

Квадрикой называют множество уровня квадратичной формы $Q(x) = \langle Ax, x\rangle$, где A — симметрический оператор в евклидовом или гильбертовом пространстве.

Теорема II.4 (теорема Гильберта—Шмидта о диагонализации симметрического оператора, 1903—1906 гг.). *Пусть даны двумерное евклидово пространство* (*для* (A)), n-*мерное евклидово пространство* (*для* (B)), *сепарабельное гильбертово пространство* (*для* (C)) *и линейные операторы, отображающие эти пространства в себя. Тогда если операторы являются симметрическими, а в бесконечномерном случае оператор является ещё и компактным, то в некотором ортонормированном базисе эти операторы диагонализируемы.*

Доказательство. В случае (A) предположим, что квадратичная форма $Q(x) = \langle Ax, x\rangle$, где

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

принимает положительные значения. Рассмотрим тогда задачу на максимум при наличии ограничения типа равенства (если форма принимала бы только отрицательные значения, мы рассмотрели бы задачу на минимум):

$$f_0(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \to \max, \quad f_1(x) = x_1^2 + x_2^2 = 1,$$

Единичный круг $\mathbb{S}^1 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$ — компакт в $\mathbb{E}^2$, функция f_0 непрерывна всюду в $\mathbb{E}^2$, значит, по теореме Вейерштрасса о непрерывной функции на компакте решение поставленной задачи существует. Обозначим его f^1. Из правила множителей Лагранжа следует, что найдётся число λ_1, для которого $Af^1 = \lambda_1 f^1$ (вектор f^1 называется *собственным вектором матрицы* A, а число λ_1 называется *собственным значением этой матрицы*). Мы доказали существование первого собственного вектора матрицы A. Пусть f^2 — единичный вектор, ортогональный f^1. Имеем

$$\langle f^1, Af^2\rangle = \langle Af^1, f^2\rangle = \lambda_1\langle f^1, f^2\rangle = 0.$$

Итак, вектор Af^2 ортогонален f^1, как и f^2, и значит, они пропорциональны: $Af^2 = \lambda_2 f^2$. Разложим вектор x по единичным ортогональным друг другу векторам f^1 и f^2:

$$x = \langle f^1, x\rangle f^1 + \langle f^2, x\rangle f^2.$$

Тогда получим, что

$$\langle Ax, x\rangle = \lambda_1\langle f^1, x\rangle^2 + \lambda_2\langle f^2, x\rangle^2.$$

Доказано, что в базисе (f^1, f^2) матрица A диагональна.

Случаи (B) и (C) рассматриваются аналогично с небольшим различием. В конечномерном случае в качестве первого шага берётся задача $\langle Ax, x\rangle \to \max$,

$\langle x,x\rangle = 1$, а в бесконечномерном случае ограничение надо рассматривать в виде неравенства: $\langle x,x\rangle \leqslant 1$. Дело в том, что в $\mathbb{R}^n$ сфера компактна, а в l_2 нет, а шар $\langle x,x\rangle \leqslant 1$ слабо компактен. Теорема Вейерштрасса приводит к существованию первого собственного вектора f^1 и старшего собственного числа λ_1. Если наложить дополнительное ограничение в виде равенства $\langle f^1,x\rangle = 0$ и применить сначала теорему Вейерштрасса и затем правило множителей Лагранжа, мы придём ко второму собственному вектору и второму собственному числу. А затем надо повторять процедуру. В конечномерном случае всё закончится за конечное число шагов, в бесконечномерном случае процедура может длиться бесконечно. В итоге приходим к представлению

$$\langle Ax,x\rangle = \sum_{k=1}^{M} \lambda_k \langle f^k,x\rangle^2,$$

где M либо конечное число, либо бесконечность, и при этом $|\lambda_k| \downarrow 0$. □

5. Дифференциалы и интегралы

Взаимная обратимость интегрирования и дифференцирования — один из самых фундаментальных результатов и в математике, и в математическом естествознании. Для того чтобы сформулировать общий результат на эту тему, необходимо освоиться с понятием дифференциальной формы.

(A) *Дифференциальные формы от одного переменного*. Таковых две: $\omega_{10} = f(x)$ и $\omega_{11} = f_1(x)\,dx$ (функции f и f_1 предполагаются гладкими). Формы ω_{1j}, $j = 0,1$, нужно будет дифференцировать и интегрировать. Областью интегрирования будет здесь отрезок $\Omega^1 = [a,b]$, граница Ω^1 состоит из точек $\{-a,+b\}$. Вот формулы для дифференциалов: $d\omega_{10} = f'(x)\,dx$, $d\omega_{11} = 0$, а формулы интегрирования таковы:

$$\int_{\partial\Omega^1} \omega_{10} = f(b) - f(a), \quad \int_{\Omega^1} \omega_{11} = \int_a^b f_1(x)\,dx.$$

Нами была доказана формула Ньютона—Лейбница на отрезке $[-1,1]$. Она даёт возможность доказать основную теорему в одномерном случае (по ходу дела совершаем замену переменных CV (Change of Variables)) $x = y(t)$, $y(-1) = a$, $y(1) = b$:

$$\int_{\Omega^1} d\omega_{10} \overset{\text{def}}{=} \int_{\Omega^1} f'(x)\,dx \overset{\text{CV}}{=} \int_{-1}^{1} f'(y)\,dy \overset{\text{NL}}{=} f(b) - f(a) \overset{\text{def}}{=} \int_{\partial\Omega^1} f(x) \iff$$
$$\iff \int_{\Omega^1} d\omega_{10} = \int_{\partial\Omega^1} \omega_{10}.$$

Теперь возможно в первом приближении дать общее определение дифференциальной формы и доказать две формулы институтского уровня.

Дифференциальные формы ω_{nm} от n переменных порядка m — это суммы одночленов вида

$$f_{i_1 \ldots i_m}(x)\, dx_{i_1} \wedge \ldots \wedge dx_{i_m},$$

где $x \in \mathbb{R}^n$, $m \leqslant n$, $dx_{i_1} \wedge \ldots \wedge dx_{i_m}$ — определитель матрицы $\{dx_{i_1} e_{i_1}, \ldots, dx_{i_m} e_{i_m}\}$, где dx_i — числа, а $\{e_i\}_{i=1}^n$ — канонический базис в $\mathbb{R}^n$. Дифференциал одночленной формы определяется следующим образом:

$$df_{i_1 \ldots i_m}(x)\, dx_{i_1} \wedge \ldots \wedge dx_{i_m} = \sum_{j=1}^{m} \frac{\partial f_{i_1 \ldots i_m}(x)}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_m}$$

(если в каком-то члене дифференциалы совпадают, этот член отбрасывается).

Переходим к доказательству обобщений формулы Ньютона—Лейбница.

(B) *Дифференциальные формы от двух и трёх переменных*, когда $n = 2$ или $n = 3$, $m = n - 1$. При $n = 2$, $m = 1$ дифференциальные формы ω_{2i}, $i = 1, 2$, имеют вид

$$\omega_{21} = f_1(x_1, x_2)\, dx_1 + f_2(x_1, x_2)\, dx_2, \quad \omega_{22} = f(x_1, x_2)\, dx_1 \wedge dx_2.$$

Областью интегрирования Ω^2 будет гладкий взаимно-однозначный образ квадрата $B^2 = [-1, 1]^2$, граница Ω^2 состоит из образа границы квадрата (т. е. из его рёбер, проходимых, скажем, против часовой стрелки). Формулы для дифференциалов имеют вид

$$d\omega_{21} = \left(\frac{\partial f_2(x_1, x_2)}{\partial x_1} - \frac{\partial f_1(x_1, x_2)}{\partial x_2} \right) dx_1 \wedge dx_2, \quad d\omega_{22} = 0,$$

а формулы интегрирования — обычные, преподающиеся в курсах анализа интегралы:

$$\int\limits_{\partial \Omega^2} f_1(x)\, dx_1 + f_2(x)\, dx_2 \;—$$

интеграл по границе области Ω^2,

$$\int\limits_{\Omega^2} f(x_1, x_2)\, dx_1 \wedge dx_2 \;—$$

это обычный интеграл

$$\int\limits_{\Omega^2} f(x_1, x_2)\, dx_1\, dx_2$$

по области Ω^2. Двумерное обобщение формулы Ньютона—Лейбница проводится почти дословно как доказательство самой формулы Ньютона—Лейбница

(«CV» — замена переменных $x = y(t)$, отображающая Ω^2 на B^2):

$$\int\limits_{\Omega^2} d\omega_{21} \stackrel{\text{def}}{=} \int\limits_{\Omega^2} d(f_1(x)\,dx_1 + f_2(x)\,dx_2) \stackrel{\text{CV}}{=}$$

$$= \int\limits_{B^2} d(f_1(y)dy_1 + f_2(y)dy_2) \stackrel{\text{def}}{=} \int\limits_{B^2} \left(\frac{\partial f_2(y)}{\partial y_1} - \frac{\partial f_1(y)}{\partial y_2}) \, dy_1 \, dy_2 \right) \stackrel{\text{(NL)}}{=}$$

(эта импликация доказывается с помощью формулы Ньютона—Лейбница после перестановки порядка интегрирования)

$$= \int\limits_{\partial B^2} f_1(y)\,dy_1 + f_2(y)\,dy_2 \stackrel{\text{CV}}{=} \int\limits_{\partial \Omega^2} \omega_{21}.$$

Эта формула называется *формулой Грина*.

При $n = 3$, $m = 2$ дифференциальные формы ω_{3i}, $i = 2, 3$, имеют вид

$$\omega_{32} = f_1(x_1, x_2, x_3)\,dx_2 \wedge dx_3 + f_2(x_1, x_2, x_3)\,dx_3 \wedge dx_1 + f_3(x_1, x_2, x_3)\,dx_1 \wedge dx_2,$$
$$\omega_{33} = f(x_1, x_2, x_3)\,dx_1 \wedge dx_2 \wedge dx_3.$$

Областью интегрирования Ω^3 будет гладкий взаимно-однозначный образ куба $B^3 = [-1, 1]^3$, граница Ω^3 состоит из образа ориентированной границы этого куба. Формулы для дифференциалов имеют вид

$$d\omega_{32} = \left(\frac{\partial f_1(x)}{\partial x_1} + \frac{\partial f_2(x)}{\partial x_2} + \frac{\partial f_3(x)}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3, \quad d\omega_{33} = 0,$$

а формулы интегрирования — обычные, преподающиеся в курсах анализа интегралы:

$$\int\limits_{\partial \Omega^3} f_1(x)\,dx_2 \wedge dx_3 + f_2(x)\,dx_3 \wedge dx_1 + f_3(x)\,dx_1 \wedge dx_2 \,—$$

интеграл по границе области Ω^3,

$$\int\limits_{\partial \Omega^3} f(x_1, x_2, x_3)\,dx_1 \wedge dx_2 \wedge dx_3 \,—$$

это обычный интеграл

$$\int\limits_{\Omega} f(x_1, x_2, x_3)\,dx_1 dx_2 dx_3$$

по области Ω^3. Трёхмерное обобщение формулы Ньютона—Лейбница проводится почти дословно как доказательство самой формулы Ньютона—Лейбница или формулы Грина («CV» — замена переменных $x = y(t)$, отображающая Ω^3 на B^3):

$$\int\limits_{\Omega^3} d\omega_{32} \overset{\text{def}}{=} \int\limits_{\Omega^3} d(f_1(x)\,dx_2 \wedge dx_3 + f_2(x)\,dx_3 \wedge dx_1 + f_3(x)\,dx_1 \wedge dx_2) \overset{\text{CV}}{=}$$

$$= \int\limits_{B^3} d(f_1(y)\,dy_2 \wedge dx_3 + f_2(y)\,dy_3 \wedge dx_2) + f_3(y)\,dy_1 \wedge dx_2 \overset{\text{def}}{=}$$

$$= \int\limits_{B^3} \left(\frac{\partial f_1(y)}{\partial y_1} + \frac{\partial f_2(y)}{\partial y_2}\right) dy_1\,dy_2 + \frac{\partial f_3(y)}{\partial y_3})\,dy_1 \wedge dy_2\,dy_3 \overset{\text{(NL)}}{=}$$

(эта импликация доказывается с помощью формулы Ньютона—Лейбница после перестановки порядков интегрирования)

$$= \int\limits_{\partial B^3} f_1(y)\,dy_2 \wedge dy_3 + f_2(y)\,dy_3 \wedge dy_1 + f_3(y)\,dy_1 \wedge dy_2 = \int\limits_{\partial\Omega^3} \omega_{32}.$$

Эта формула называется *формулой Остроградского—Гаусса*.

(C) *Дифференциальные формы, где n и m произвольны.*

Доказанные формулы Ньютона—Лейбница, Грина и Гаусса—Остроградского позволяют высказать следующее общее утверждение об обратимости операторов дифференцирования и интегрирования дифференциальных форм: интеграл от дифференциала дифференциальной формы по области равен интегралу от самой формы по границе области. А именно, имеет место формула

$$\int\limits_{\Omega^n} d\omega_{nm} = \int\limits_{\partial\Omega^m} \omega_{nm}.$$

Эту формулу иногда называют *формулой Стокса* (Стокс опубликовал формулу с $n = 2$, $m = 3$), иногда *формулой Стокса—Пуанкаре*, иногда *формулой Ньютона—Лейбница—Грина—Гаусса—Остроградского—Стокса—Пуанкаре*. Доказательство, которое можно провести по применённым выше схемам, предоставляется читателю.

III. Приложения

Обзор математических курсов мехмата

Здесь мы касаемся курсов мехмата, содержание которых отражено в программе государственного экзамена.

1. Аналитическая геометрия

Основное содержание курса составляют геометрическая интерпретация теории линейных уравнений и теории квадрик в двумерном и трёхмерном пространствах и исчисление этих объектов. В государственный экзамен включён единственный вопрос.

Аффинная и метрическая классификация кривых и поверхностей второго порядка. Проективная классификация кривых.

Теория этого вопроса заключена в теореме II.4 о диагонализации симметрического оператора, причём сам вопрос относится к частным случаям: двумерному и трёхмерному. Далее идёт классификация возникающих при этом объектов. Продемонстрируем всё это на примере аффинной классификации в двумерном случае.

Теорема III.1 (аффинная классификация кривых второго порядка). *Плоская кривая второго порядка аффинным преобразованием приводится к одному из следующих типов: окружность, гипербола, парабола, пара прямых* (*параллельных или пересекающихся*), *точка и пустое множество.*

Доказательство. Плоская кривая второго порядка (плоская квадрика, а по Аполлонию — коника) — это множество уровня квадратной функции, т. е. множество

$$\mathcal{K} = \{(x_1, x_2) \mid a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0\}.$$

Наша цель — аффинной заменой переменных привести уравнение к наиболее простой форме. Ортогональная замена переменных (с применением теоремы о диагонализации с последующей гомотетией) приведёт квадратичную форму к виду $y_1^2 + \varepsilon y_2^2$, где $\varepsilon = \{1, 0, -1\}$ плюс аффинные члены. Если $\varepsilon = 1$ или $\varepsilon = -1$, то сдвиг $y_i \mapsto y_i + \alpha_i$, $i = 1, 2$, с последующей гомотетией приведут уравнение квадрики к виду $z_1^2 + z_2^2 = \{-1, 0, 1\}$ или $z_1^2 - z_2^2 = \{1, 0\}$, а в случае $\varepsilon = 0$ — к одному из двух видов $z_1^2 = \{1, 0, -1\}$ или $z_2 = z_1^2$, что и приводит к ответу (скажем, если выражение $z_1^2 + z_2^2$ равно единице — это окружность, нулю — это точка (начало координат), минус единица — пустое множество и т. д.). □

2. Высшая алгебра

Теорема Кронекера—Капелли и основная теорема алгебры — два важнейших результата курса высшей алгебры, проходимого на первом курсе мехмата.

Рангом матрицы называют максимальное число линейно независимых столбцов этой матрицы. Квадратную матрицу A порядка n назовём *разрешимой*, если уравнения $Ax = y$ разрешимы для любого вектора $y \in \mathbb{N}$.

Лемма о ранге. *Ранг матрицы равен наивысшему порядку разрешимой квадратной подматрицы данной матрицы.*

Доказательство. Ранг матрицы A обозначим через r, а наивысший порядок разрешимой квадратной матрицы A' через r'. Без ограничения общности считаем, что столбцы матрицы A' имеют первые r' номеров. Из теоремы Фредгольма следует, что у разрешимой матрицы столбцы линейно независимы, и следовательно, $r \geqslant r'$. Допустим, что $r > r'$ и $(r' + 1)$-й столбец матрицы A линейно независим с первыми её r' столбцами. Опять-таки по теореме Фредгольма найдётся линейная комбинация столбцов матрицы A', для которой первые r' координат совпадут с первыми r' координатами $(r'+1)$-го столбца матрицы A. Тогда из определения $(r'+1)$-го столбца следует, что найдётся элемент $a_{i(r'+1)}$, $i > r'+1$, не совпадающий с элементом с теми же номерами построенной линейной комбинации. Но это означает, что квадратная матрица размера $r'+1$, «окаймлённая» столбцом $a_{1(r'+1)}, \ldots, a_{r'(r'+1)}, a_{i(r'+1)}$ и строкой $a_{i1}, \ldots, a_{ir'}, a_{i(r'+1)}$ матрицы A', будет (снова по теореме Фредгольма) разрешимой, что противоречит определению матрицы A'. □

Теорема III.2 (теорема Кронекера—Капелли). *Система линейных уравнений разрешима тогда и только тогда, когда ранг матрицы системы совпадает с рангом матрицы, получаемой из матрицы системы присоединением к ней вектора правых частей.*

Доказательство. Если система $Ax = y$ с матрицей A из m строк и n столбцов разрешима, это означает, что вектор y свободных членов выражается через столбцы матрицы, а значит, в силу леммы о ранге он выражается через r столбцов матрицы A, где r — её ранг. Таким образом, ранг расширенной матрицы (см. определения ранга) равен r. А если система не является разрешимой, то это значит, что вектор y не выражается через столбцы матрицы A, т. е. ранг расширенной матрицы больше ранга A. □

Теорема III.3 (теорема Даламбера—Гаусса, основная теорема алгебры). *Многочлен с комплексными коэффициентами, степень которого больше нуля, имеет комплексный корень.*

Доказательство. Пусть $p(\cdot)$ — многочлен степени n. Рассмотрим задачу без ограничений $|p(z)|^2 \to \min$. Модуль многочлена растёт на бесконечности:

$$|p(z)| = \left| a_n z^n \left(1 + \sum_{k=0}^{n-1} \frac{a_k}{a_n} z^{k-n} \right) \right| = |a_n|\,|z|^n \big(1 + o(1)\big) \to \infty,$$

когда $|z| \to \infty$. В силу теоремы Вейерштрасса в задаче достигается минимум. Не ограничивая себя в общности, можно считать, что минимум достигается в нуле, и тогда $p(z) = a_0 + a_k z^k + \ldots$, где $a_1 = \ldots = a_{k-1} = 0$, $a_k \neq 0$. Имеем

$$f(t) = |p(t^{1/k} e^{i\theta})|^2 = |a_0|^2 + 2|a_0|\,|a_k| t \cos(k\theta + \beta) + o(t), \quad t \geqslant 0,$$

откуда (по минимальности f в нуле) следует, что $|a_k|\,|a_0| \cos(k\theta + \beta) \geqslant 0$ для всех θ (теорема Ферма для полупрямой), что возможно, лишь если $a_0 = 0$. □

3. Линейная алгебра и геометрия

Без сомнения, к числу наиболее существенных результатов, изучаемых в курсе линейной алгебры и геометрии, относятся теоремы о приведении квадратичной формы к главным осям и к нормальному виду.

Следующие результаты дают ответы на вопросы госэкзамена.

Теорема III.4 (о приведении квадратичной формы к диагональному виду). *Для всякой симметричной матрицы A найдётся ортогональная матрица, которая приводит матрицу A к диагональному виду (или, что то же самое, конечномерный симметрический оператор ортогональным преобразованием приводится к диагональному виду.*

Это не что иное, как конечномерная теорема о диагонализации, доказанная во второй части.

Теорема III.5 (о приведении квадратичной формы к нормальному виду). *Квадратичная форма линейной заменой переменных $x = Cy$ (где C — обратимая матрица) приводится к нормальному виду:*

$$Q(x) = \sum_{j=1}^{n} \varepsilon_j y_j,$$

где $\varepsilon_j = \{-1, 0, 1\}$.

Доказательство. По теореме о диагонализации в некотором ортонормированном базисе $\{f_j\}$ квадратичная форма запишется в виде

$$\sum_{j=1}^{m} \lambda_j z_j^2,$$

где $\lambda_j \neq 0$ и $m \leqslant n$. Остаётся сделать простейшие замены переменных $y_j = \sqrt{\lambda_j} z_j$, если $\lambda_j > 0$, и $y_j = -\sqrt{-\lambda_j} z_j$, если $\lambda_j < 0$, и мы приходим к нужному результату. □

К последнему результату примыкает следующая теорема.

Теорема III.6 (о законе инерции). *Число положительных и отрицательных квадратов, к которым приводится квадратичная форма, не зависит от выбора преобразования.*

Доказательство. Пусть преобразования $x = By$ и $x = Cz$ с невырожденными матрицами B и C привели форму $x \mapsto Q(x)$ к двум разным нормальным видам, скажем,

$$Q = y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = z_1^2 + \ldots + z_l^2 - z_{l+1}^2 - \ldots - z_r^2, \qquad (**)$$

причём $k < l$. Выразив все y и z через x, составим систему из $n + k - l < n$ однородных уравнений относительно $x \in \mathbb{R}^n$:

$$y_1(x) = \ldots = y_k(x) = z_{l+1}(x) = \ldots = z_n(x) = 0.$$

Из теоремы Кронекера—Капелли следует, что существует решение $\hat{x} \neq 0$ этой системы. Подставляя эти решения в $(**)$, приходим к тому, что

$$-y_{k+1}^2(\hat{x}) - \ldots - y_r^2(\hat{x}) = z_1^2(\hat{x}) + \ldots + z_l^2(\hat{x}),$$

откуда следует, что $z_j(\hat{x}) = 0$, $1 \leqslant j \leqslant l$, т. е. система $z = C^{-1}x$ имеет ненулевое решение, что невозможно. Теорема доказана. □

4. Математический анализ

Курс математического анализа — это огромный двухгодичный курс. Посмотрим на него с позиции государственного экзамена. К разделу «Математический анализ» относятся 11 вопросов.

1. *Непрерывность функций одной переменной. Свойства непрерывных функций.*
2. *Функции многих переменных, полный дифференциал и его геометрический смысл. Достаточные условия дифференцируемости. Градиент.*
3. *Определённый интеграл. Интегрируемость непрерывной функции. Первообразная непрерывной функции.*
4. *Неявные функции. Существование, непрерывность и дифференцируемость неявных функций.*
5. *Числовые ряды. Сходимость рядов. Критерий Коши сходимости ряда. Достаточные условия сходимости.*
6. *Абсолютная и условная сходимость ряда. Свойства абсолютно сходящихся рядов. Умножение рядов.*
7. *Ряды функций. Равномерная сходимость. Признак Вейерштрасса. Свойства равномерно сходящихся рядов (непрерывность суммы, почленное дифференцирование и интегрирование).*
8. *Степенные ряды в действительной и комплексной области. Радиус сходимости, свойства степенных рядов (почленное дифференцирование, интегрирование). Разложение элементарных функций.*
9. *Несобственные интегралы и их сходимость. Равномерная сходимость интегралов, зависящих от параметра. Свойства равномерно сходящихся интегралов.*
10. *Ряды Фурье. Достаточные условия представимости функции рядом Фурье.*
11. *Теоремы Остроградского и Стокса.*

Многие вопросы из этого списка — о непрерывных функциях, дифференцируемости, о неявных функциях — были изложены нами на более высоком уровне. Вопросы с пятого по седьмой и девятый относятся к исчислениям. Одиннадцатый вопрос был нами освещён.

Формула Тейлора с интегральным остаточным членом. Пусть функция $f\colon [a,b] \to \mathbb{R}$ (для определённости) непрерывна на $[a,b]$. Любую функцию семейства

$$x \mapsto F(x) = \int\limits_a^x f(t)\,dt + C$$

называют *первообразной функции* f, а само семейство — *неопределённым интегралом от* f и обозначают

$$\int f(x)\,dx.$$

Из определения производной легко выводится, что $F'(x) = f(x)$. Это равенство иногда принимают за определение первообразной (для доказательства равносильности определений надо уметь доказывать, что уравнению $F'(x) = 0$ удовлетворяют только константы). Если f непрерывно дифференцируема, то из сказанного вытекает формула:

$$f(x) = f(a) + \int\limits_a^x f'(t)\,dt.$$

Если предположить, что f дважды непрерывно дифференцируема и первая производная представима формулой

$$f'(t) = f'(a) + \int\limits_a^t f''(\tau)\,d\tau,$$

то после подстановки второй формулы в первую и после перемены порядка интегрирования приходим к формуле

$$f(x) = f(a) + f'(a)(x-a) + \int\limits_a^x (t-a) f''(t)\,dt.$$

Повторив эту процедуру ещё $n-1$ раз, приходим к формуле Тейлора с интегральным остаточным членом:

$$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{1}{(n-1)!} f^{(n-1)}(a)(x-a)^{n-1} + \\ + \frac{1}{(n-1)!} \int\limits_a^x (x-t)^{n-1} f^{(n)}(t)\,dt.$$

Элементарные функции и их разложения в ряды Тейлора. К элементарным функциям относят полиномы, экспоненты, тригонометрические функции и обратные к ним. Полиномы сами по себе представляют собой тейлоровские суммы. Важнейшая из экспонент — функция

$$x \mapsto e^x \text{ —}$$

есть решение задачи Коши

$$\dot{x} = x, \quad x(0) = 1$$

или в интегральной форме

$$x(t) = 1 + \int_0^t x(\tau)\,d\tau.$$

Итеративная последовательность

$$x_{n+1}(t) = 1 + \int_0^t x_n(\tau)\,d\tau, \quad x_0(t) = 1$$

приводит к ряду

$$e^t = \sum_{k\geqslant 0} \frac{t^k}{k!},$$

сходимость которого легко обосновывается.

Аналогично функция

$$x \mapsto \sin t$$

есть решение задачи Коши

$$\ddot{x} + x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1$$

или в интегральной форме

$$x(t) = t + \int_0^t (t-\tau)x(\tau)\,d\tau.$$

Итеративная последовательность

$$x_{n+1}(t) = t + \int_0^t (t-\tau)x_n(\tau)\,d\tau, \quad x_0(t) = t$$

приводит к ряду

$$\sin t = \sum_{k\in\mathbb{N}} (-1)^{k-1} \frac{t^{2k-1}}{(2k-1)!}.$$

Ряды для обратных функций (логарифма, арктангенса и др.) получаются интегрированием рядов:

$$\ln'(1+x) = \frac{1}{1+x}$$

следовательно,

$$\ln(1+x) = \sum_{k\in\mathbb{N}} (-1)^{k+1} \frac{x^k}{k}.$$

Ряды Фурье и специальные функции. Окружность — простейшее компактное однородное пространство, т. е. компактное многообразие, на котором транзитивно действует группа сдвигов, сохраняющих расстояние. Реализуем это многообразие в виде отрезка $[-\pi, \pi]$ с идентифицированными концами.

На компактных однородных пространствах существуют дифференциальные операторы второго порядка, инвариантные относительно сдвига. На окружности (в нашей реализации) таким оператором является оператор

$$\Delta = \frac{d^2}{dt^2}$$

с периодическими краевыми условиями: $x(-\pi) = x(\pi)$, $\dot{x}(-\pi) = \dot{x}(\pi)$. Этот оператор не является ограниченным, однако обратный к нему — интегральный оператор — является симметричным и компактным оператором. Для того чтобы выписать его, надо решить уравнение

$$\frac{d^2}{dt^2} x(t) = \delta(t),$$

где $\delta(\cdot)$ — δ-функция Дирака, с периодическими условиями на концах и при равенстве нулю от решения уравнения. Проинтегрировав «эвристически» δ-функцию, приходим к функции

$$B_1(t) = \begin{cases} -\dfrac{t+\pi}{2\pi} & \text{для } -\pi \leqslant t < 0; \\ -\dfrac{t-\pi}{2\pi} & \text{для } 0 < t \leqslant \pi. \end{cases}$$

Повторное интегрирование приводит к формуле

$$B_2(t) = c + \begin{cases} -\dfrac{(t+\pi)^2}{4\pi} & \text{для } -\pi \leqslant t < 0; \\ -\dfrac{(t-\pi)^2}{4\pi} & \text{для } 0 < t \leqslant \pi. \end{cases}$$

Здесь константа подбирается так, чтобы интеграл по периоду от функции $B_2(\cdot)$ равнялся нулю. В итоге мы приходим к формуле обращения:

$$\xi(t) = \int\limits_{\mathbb{T}} B_2(t+\tau)\ddot{\xi}(\tau)\, d\tau,$$

которая проверяется непосредственным вычислением для функций, скажем, из $C^2(\mathbb{T})$ (интеграл которых по периоду равен нулю), а значит, формула обращения верна и в пополнении этого пространства по среднеквадратичной норме. Такое пополнение обозначается $L_2(\mathbb{T})$. Про него доказывается, что оно изометрично l_2. Оператор

$$\mathcal{B}x(t) = \int\limits_{\mathbb{T}} B_2(t+\tau)x(\tau)\, d\tau$$

таков, что к оператору

$$x(\cdot) \mapsto x(\cdot) + \mathcal{B}x(\cdot)$$

применима теорема Гильберта—Шмидта. Применение этой теоремы приводит к тому, что $L_2(\mathbb{T})$ разложится в ряд по собственным функциям оператора $\mathcal{B}$ или, что то же самое, по собственным функциям оператора

$$\frac{d^2}{dt^2}$$

(с периодическими краевыми условиями). Так мы приходим к базису в $L_2(\mathbb{T})$:

$$e_0(t) = \frac{1}{2\pi}, \quad e_{2k-1}(t) = \frac{1}{\pi}\sin kt, \quad e_{2k}(t) = \frac{1}{\pi}\cos kt, \quad k \in \mathbb{N}.$$

Разложение, описанное в теореме Гильберта—Шмидта — это не что иное, как разложение в обычный ряд Фурье.

На этом пути можно построить огромное число специальных функций, начиная со сферических функций на $\mathbb{S}^2$, выражающихся через основные и присоединённые полиномы Лежандра.

5. Обыкновенные дифференциальные уравнения

В государственный экзамен включены две важнейших теоремы о существовании и единственности.

Теорема III.7 (локальная теорема о существовании и единственности решения задачи Коши). *Пусть V — окрестность точки $(\tau, \xi) \in \mathbb{R} \times \mathbb{R}^n$ и функция $f\colon V \to \mathbb{R}^n$ определена и непрерывна в V и удовлетворяет там условию Липшица по x. Тогда на некотором отрезке $\Delta = [\tau - \alpha, \tau + \alpha]$ определено единственное решение задачи Коши*

$$\dot{x} = f(t, x), \quad x(\tau) = \xi, \tag{III.1}$$

Доказательство. Пусть в V лежит (при некоторых положительных a и b) множество $D = [\tau - a, \tau + a] \times B_{\mathbb{R}^n}(\xi, b)$ и для любых $(t, x_i) \in D$, $i = 1, 2$ выполняется неравенство $|f(t, x_2) - f(t, x_1)| \leqslant L|x_2 - x_1|$. Применим следствие теоремы о правом обратном — теорему Грейвса — к случаю, когда $X = Y = C([\tau - \alpha, \tau + \alpha], \mathbb{R}^n)$,

$$F\big(x(\cdot)\big)(t) = x(t) - \xi - \int\limits_{\tau}^{t} f(x, s)\, ds,$$

$\Lambda = \mathrm{Id}$ (тождественный оператор),

$$\alpha = \min\left\{a, \frac{1}{2L}, \frac{b}{2M}\right\},$$

где $M = \max\limits_{t \in [\tau - a, \tau + a]} f(t, \xi)$. Читатель без труда убедится в том, что при этом условие $(**)$ выполняется, и значит, применение теоремы Грейвса приводит к цели. □

Теорема III.8 (глобальная теорема существования и единственности для линейных систем). *Пусть в дифференциальном уравнении из* (III.1) $f(t,x) = A(t)x + b(t)$, *где* $A(\cdot)$ — *непрерывная на отрезке* $[t_0, t_1]$ *матричная функция, а* $b(\cdot)$ — *непрерывная векторная функция. Тогда для любых* $\tau \in [t_0, t_1]$ *и* $\xi \in \mathbb{R}^n$ *существует единственное решение уравнения* (III.1), *определённое на всём отрезке* $[t_0, t_1]$.

Доказательство. Применим только что доказанную локальную теорему к нашему случаю. В нашем случае число α, определённое в локальной теореме, не зависит от положения начальной точки. Решение $\hat{x}(\cdot)$ задачи (III.1) для линейной системы будет определено на отрезке $[\tau - \alpha, \tau + \alpha]$. Теперь продолжим решение на отрезок $[\tau + \alpha, \tau + 2\alpha]$, решив задачу Коши с начальными данными $x(\tau + \alpha) = \hat{x}(\tau + \alpha)$. За несколько шагов мы достигнем правого конца, а затем, двигаясь «влево», достигнем и левого конца, т. е. построим решение на всём отрезке $[t_0, t_1]$. $\square$

6. Комплексный анализ

Центральными результатами этого раздела являются теорема и формула Коши.

Определение производной комплексной функции. Пусть функция f определена в окрестности точки $\hat{z}$. Говорят, что она *дифференцируема в этой точке*, если существует число $a = \alpha + i\beta \in \mathbb{C}$, такое что $f(\hat{z} + z) = f(\hat{z}) + az + r(z)$, где $r(z) = o(|z|)$, т. е. $\lim\limits_{z \to 0} \frac{r(z)|}{|z|} = 0$. Число a, определяемое написанным соотношением однозначно, называется *производной* f *в точке* $\hat{z}$ и обозначается $f'(\hat{z})$. Если f дифференцируема в окрестности точки $\hat{z}$ и производная непрерывна как функция z в окрестности $\hat{z}$, говорят, что f там *голоморфна*.

Лемма III.1 (условия Коши—Римана). *Пусть функция* f *голоморфна в окрестности точки* $\hat{z} = (\hat{x}, \hat{y})$. *Тогда для* $u(\hat{x}, \hat{y}) = \operatorname{Re} f(\hat{z})$ *и* $v(\hat{x}, \hat{y}) = \operatorname{Im} f(\hat{z})$ *выполнены соотношения*

$$\frac{\partial u(\hat{x}, \hat{y})}{\partial x} = \frac{\partial v(\hat{x}, \hat{y})}{\partial y}, \quad \frac{\partial v(\hat{x}, \hat{y})}{\partial x} = -\frac{\partial u(\hat{x}, \hat{y})}{\partial y}$$

(*называемые условиями Коши—Римана*).

Доказательство. Имеем

$$f(\hat{z} + z) = f(\hat{z}) + f'(\hat{z})z + o(|z|) = u(\hat{x} + x, \hat{y} + y) + iv(\hat{x} + x, \hat{y} + y) + o(|z|) \overset{\mathrm{Id}}{=}$$
$$= f(\hat{z}) + \frac{\partial u}{\partial x}x + \frac{\partial v}{\partial y}y + i\left(\frac{\partial v}{\partial x}x + \frac{\partial v}{\partial y}y\right) + o(|z|).$$

Положив $f'(\hat{z}) = \xi + i\eta$, приравняв вещественную и мнимую компоненты левой и правой частей равенства, приходим к соотношению

$$\xi x - \eta y + i(\xi y + \eta x) = \frac{\partial u}{\partial x}x + \frac{\partial v}{\partial x}y + i\left(\frac{\partial v}{\partial x}x + \frac{\partial v}{\partial y}y\right).$$

Тогда
$$\frac{\partial u}{\partial x} = \xi, \quad \frac{\partial u}{\partial y} = -\eta, \quad \frac{\partial v}{\partial x} = \eta, \quad \frac{\partial v}{\partial y} = \xi,$$
следовательно,
$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \qquad \square$$

Теорема III.9 (теорема Коши об интеграле по замкнутому контуру). *Пусть U — открытая односвязная область в $\mathbb{C}$, $f\colon U \to \mathbb{C}$ — голоморфная функция в U, γ — спрямляемый замкнутый путь в U. Тогда*
$$\int\limits_{\gamma} f(z)\,dz = 0$$
(*иначе говоря, интеграл по гладкому замкнутому контуру, охватывающему односвязную область, от аналитической функции равен нулю*).

Доказательство. Действительно,
$$\int\limits_{\partial\Omega} f(z)\,dz \overset{\text{def}}{=} \int\limits_{\partial\Omega} (u+iv)(dx + i\,dy) \overset{\text{Id}}{=} \int\limits_{\partial\Omega} (u\,dx - v\,dy) + i(v\,dx + u\,dy) \overset{\text{формула Грина}}{=}$$
$$= \int\limits_{\Omega} \left(-\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right) dx\,dy + i \int\limits_{\Omega} \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy \overset{\text{теорема Коши—Римана}}{=} 0. \qquad \square$$

Теорема III.10 (формула Коши). *Пусть f голоморфна в некоторой области, содержащей точку z, и γ — замкнутая кривая, охватывающая односвязную область, включающую точку z. Имеет место формула Коши*
$$f(z) = \frac{1}{2\pi i} \int\limits_{\gamma} \frac{f(\zeta)}{\zeta - z}\,d\zeta.$$

Доказательство. Проведём окружность $\gamma_\varepsilon = \{\zeta \in \mathbb{C} \mid |\zeta - z| = \varepsilon\}$. Если из области, ограниченной кривой γ, исключить круг, ограниченный окружностью γ_ε, останется двусвязная область, ограниченная кривыми γ и γ_ε. По теореме Коши сумма интегралов, проходимых по γ в положительном направлении (против часовой стрелки) и по γ_ε в отрицательном направлении, равна нулю. Значит, интеграл по γ равен
$$J_\varepsilon = \int\limits_{|\gamma - z| = \varepsilon} \frac{f(\zeta)d\zeta}{\zeta - z}.$$
Положим $\zeta = z + \varepsilon e^{it}$, $0 \leqslant t \leqslant 2\pi$, и получим, что
$$J = \int\limits_0^{2\pi} \frac{f(\varepsilon e^{it})\varepsilon i e^{it}\,dt}{\varepsilon e^{it}} = i \int\limits_0^{2\pi} f(\varepsilon e^{it})\,dt = 2\pi i\bigl(f(\varepsilon e^{it}) + o(\varepsilon)\bigr).$$
Переходя к пределу при $\varepsilon \to 0$, получаем формулу Коши. $\square$

7. Дифференциальная геометрия

В государственный экзамен включены два результата по дифференциальной геометрии: теорема Менье и формула Эйлера. Мы расскажем о них, но сначала определим понятия, нужные для их формулировок.

Гладкая плоская кривая локально есть образ на евклидовой плоскости гладкого отображения единичного интервала вещественной прямой $\mathbb{R}$. Рассмотрим плоскую кривую γ, точку ξ на этой кривой, а также $T_\xi\gamma$ и $N_\xi\gamma$ — касательную и нормаль к γ в точке ξ. Примем $T_\xi\gamma$ и $N_\xi\gamma$ за координатные оси евклидовой плоскости. В этих осях уравнение кривой запишется в виде

$$x_2 = f_\gamma(x_1) = \frac{kx_1^2}{2} + o(x_1^2), \quad k \geqslant 0.$$

Если $k > 0$, это уравнение означает, что в окрестности точки ξ с точностью до второго порядка кривая аппроксимируется параболой. При этом найдётся единственная точка $(0, r)$, такая что окружность с центром в $(0, r)$ радиуса r аппроксимирует γ в точке $\xi \leftrightarrow (0,0)$ с точностью выше второго порядка. Нетрудно показать, что $r = k^{-1}$. Число k называют *кривизной γ в ξ*, а точку $(0, k^{-1})$ — *центром кривизны γ в ξ*.

Рассмотрим теперь кривую ℓ в трёхмерном евклидовом пространстве (локально это образ при гладком отображении в это пространство единичного интервала из $\mathbb{R}$). Пусть p — точка на ℓ, $T_p\ell$ — касательная к этой кривой в точке p. Среди плоскостей, содержащих $T_p\ell$, существует та, которая локально, в окрестности точки p, наилучшим образом аппроксимирует кривую ℓ. Либо такая плоскость единственна (невырожденный случай), либо любая плоскость, проходящая через $T_p\ell$, обладает свойством наилучшей аппроксимации (вырожденный случай). В первом случае единственная аппроксимирующая плоскость называется *сопровождающей плоскостью*, обозначим её $\mathcal{P}$. Пусть ℓ' — это образ кривой ℓ при проектировании на $\mathcal{P}$. *Центром кривизны пространственной кривой ℓ в точке $\hat{p}$* называется центр кривизны плоской кривой ℓ'.

Локально гладкая поверхность в трёхмерном евклидовом пространстве есть образ в этом пространстве гладкого отображения единичного диска из $\mathbb{R}^2$. Рассмотрим гладкую поверхность S, выберем точку P на этой поверхности и построим T_PS и N_PS — касательное пространство и нормаль к S в точке P. Пусть e_3 — единичный вектор вдоль N_PS. При выборе любого базиса в T_PS поверхность S в окрестности P запишется в виде

$$x_3 = Q(x_1, x_2) + o(x_1^2 + x_2^2),$$

где Q — квадратичная форма. Если выбрать в T_PS ортонормированный базис e_1, e_2, в соответствии с теоремой II.4 о диагонализации уравнение поверхности S запишется в виде

$$x_3 = \frac{k_1x_1^2 + k_2x_2^2}{2} + o(x_1^2 + x_2^2).$$

Направления вдоль e_1 и e_2 называются *главными направлениями* на поверхности S, а величины k_1 и k_2 — *главными кривизнами*; сечение поверхности S

плоскостями, содержащими нормаль N_pS, называют *нормальными сечениями*. Имеют место следующие результаты.

Теорема III.11 (теорема Менье). *Центр кривизны любой кривой на поверхности есть проекция центра кривизны нормального сечения с той же касательной на главную нормаль этой кривой.*

Теорема III.12 (формула Эйлера). *Если отличные от нуля числа k_1 и k_2 — главные кривизны поверхности S, то кривизна k_α нормального сечения, угол которого с вектором e_1 равен α, вычисляется по формуле $k_\alpha = k_1 \cos\alpha + k_2 \sin\alpha$.*

Обе теоремы фактически следуют из определений.

В заключение скажем несколько слов о двух курсах, не отражённых в госэкзамене.

В курсе «Функциональный анализ» существенную роль играют три принципа линейного анализа: принцип Банаха об открытости отображения, теорема Хана—Банаха и принцип равномерной ограниченности (теорема Банаха—Штейнгауза). Все они (иногда в иной форме) были нами доказаны. Действительно, лемма о правом обратном для линейного сюръективного оператора, доказанная выше, равносильна принципу открытости; теорема Хана—Банаха, по сути дела, равносильна теореме отделимости, доказанной нами, а доказательство теоремы Банаха—Штейнгауза аналогично доказательству леммы II.2 об ограниченности слабо сходящейся последовательности.

Основные задачи математической физики, изучаемые в курсе «Дифференциальные уравнения с частными производными», решаются методом Фурье, редуцирующим задачи математической физики к теореме Гильберта—Шмидта. Методом Фурье решаются и классические задачи квантовой механики.

Продемонстрируем это на примере описания колебания струны. Пусть $u(t,x)$ — положение в момент t колеблющейся закреплённой в концах отрезка $[0,1]$ струны, начальное положение которой в точке x было равно $\varphi(x)$, а скорость в точке x в начальный момент времени была равна $\psi(x)$. Функция $(t,x) \mapsto u(t,x)$ удовлетворяет уравнению струны (в этом уравнении «проглядывает» второй закон Ньютона, соединённый с законом Гука):

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad (*)$$

краевым условиям $u(t,0) = u(t,1) = 0$ и начальным условиям $u(0,x) = \varphi(x)$, $u_t(0,x) = \psi(x)$.

Уравнение $(*)$ относят к числу *гиперболических уравнений*. Начальные условия в задаче о колебании струны содержат положение струны и её начальную скорость, как в законе Ньютона.

Решение. Решим задачу $(*)$ методом разделения переменных. Ищем решение уравнения струны в виде произведения $T(t)X(x)$. Тогда уравнение

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$$

приводит к равенству
$$\frac{T''(t)}{T(t)}=\frac{X''(x)}{X(x)}.$$
Значит,
$$\frac{T''(t)}{T(t)}=\frac{X''(x)}{X(x)}=\mathrm{const}=\lambda.$$
Краевые условия ведут к равенствам $X(0)=X(1)=0$. Таким образом, функция $X(\cdot)$ является решением задачи Коши
$$X''=\lambda X,\quad X(0)=X(1)=0.$$
Отсюда следует, что $\lambda_k=-k^2$, и решениями задачи Коши являются функции $c_k=\sin k\pi x$, $k\in\mathbb{N}$. Остаётся применить теорему Гильберта—Шмидта о приведении квадратичной формы к главным осям. Положим $G(x,\xi)=(1-\xi)x$ при $0\leqslant\xi\leqslant x$ и $G(x,\xi)=(1-x)\xi$ при $x\leqslant\xi\leqslant 1$. Дважды продифференцировав по x функцию
$$z(x)=\int\limits_0^1 G(x,\xi)y(\xi)\,d\xi,$$
убеждаемся в том, что $z''(x)=y(x)$, и при этом $z(0)=z(1)=0$. Система функций $\{2^{-1/2}\sin k\pi x\}_{k\in\mathbb{N}}$ является полной ортонормированной системой в $L_2([0,1])$. Остаётся решить уравнения $T''(t)+n^2T(t)=0$, и тогда общее решение уравнения струны оказывается следующим:
$$u(t,x)=\sum_{k\in\mathbb{N}}(c_k\sin nt+d_k\cos nt)\sin k\pi x,$$
где c_k и d_k определяются однозначно из начальных данных. □

Аналогичным образом исследуются уравнения Шрёдингера, описывающее эволюцию систем квантовой механики.

Приложения к естествознанию

Выше было сказано о единстве математики и математического естествознания. Галилею (1564—1642) принадлежат слова: «Книга Природы написана на языке математики». Давайте читать Книгу Природы на каждом из уровней (A), (B) и (C).

Об одной особенности законов, записанных в Книге Природы, Л. Эйлер (1707—1783) сказал так: «В мире не происходит ничего, в чём не был бы виден смысл какого-нибудь максимума или минимума». Эту мысль иногда формулируют следующим образом: «законы природы в значительной своей доле описываются экстремальными принципами» (в этом одна из мотивировок теории экстремума в математике).

И здесь проведём разговор на «трёх уровнях».

(A) Знаний по теории экстремума, полученных выше, достаточно, чтобы даже на школьном уровне вывести закон преломления света при переходе из одной однородной среды с одной скоростью распространения света в другую с другой скоростью, экспериментально установленный В. Снеллиусом (1580—1626).

Для этого надо применить условие минимума в простой задаче

$$f(x) = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (l - x)^2}}{v_2} \to \min,$$

которая формализует первый экстремальный принцип в естествознании, открытый Ферма (1601—1665) в 1662 г. Согласно этому принципу свет, распространяющийся в неоднородной среде от одной точки до другой, избирает путь, при котором затрачивается наименьшее время. В поставленной задаче свет идёт из точки $(0, a)$ в верхней полуплоскости, где его скорость распространения равна v_1, в точку $(l, -b)$ в нижней полуплоскости, где скорость равна v_2.

В 1687 году вышла книга И. Ньютона (1643—1727) «Математические начала натуральной философии», величайшая в истории науки. В ней содержался целый свод законов природы, связанных единой концепцией. Размышляя над проблемами школьного математического образования, А. Н. Колмогоров писал: «Вряд ли нужно доказывать, насколько желательно с общеобразовательной точки зрения достигнуть того, чтобы все учащиеся могли вполне конкретно понять хотя бы ньютоновскую концепцию математического естествознания».

В применении к законам динамики ньютоновскую концепцию математического образования можно кратко выразить словами «законы динамики описываются дифференциальными уравнениями». Всё это мотивирует изучение теории обыкновенных дифференциальных уравнений и исчисление таких уравнений.

В частности, движение $t \mapsto x(t)$ частицы массы m вдоль прямой под воздействием силы $F(x)$ (зависящей от координаты частицы) удовлетворяет дифференциальному уравнению $m\ddot{x} = F(x)$. Этот закон при $F(x) = 0$ и $F(x) = \mathrm{const}$ описывает законы Галилея (первые законы динамики) — закон инерции и закон падения тел, а при законе Гука $F(x) = -\alpha x$ — закон гармонических колебаний. Если сила воздействия центра убывает пропорционально расстоянию от центра, положение частицы экспоненциально убывает со временем. Лишь одно это — исследование простейших одномерных движений — мотивирует целесообразность изучения элементарных функций (полиномов, тригонометрических функций и экспонент, а для решений уравнений с ними — радикалов, обратных тригонометрических функций и логарифмов) в школьных, институтских и университетских курсах.

(B) Большинство законов динамики выводятся из *принципа Гамильтона*, согласно которому каждая динамическая система характеризуется функцией $L = L(t, x, \dot{x})$, $L\colon \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$, называемой *лагранжианом*. Функционал

$$S(x(\cdot)) = \int\limits_{t_0}^{t_1} L\big(t, x(t), \dot{x}(t)\big)\, dt$$

называют *действием*. Подобно сформулированному в принципе Ферма относительно света, механическая система, движущаяся от одной точки (t_0, x_0) до другой (t_1, x_1), «избирает» траекторию, требующую (локально, при начальной и конечной точках, достаточно близко расположенных друг к другу) наименьшего действия на её преодоление, т. е., в частности, удовлетворяет *уравнению Эйлера*

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0,$$

о котором рассказывается на лекциях по теории экстремума. Этот принцип (после установления характера взаимодействия) сводит вывод закона природы к исчислению. Выведем, исходя из принципа Гамильтона, законы Кеплера о движении тел вокруг Солнца. Это уже университетский уровень.

(С) Сила F, притягивающая движущееся тело массы m (находящееся в точке плоскости, имеющей декартовы координаты $\big(x(t), y(t)\big)$ и полярные координаты $\big(r(t), \varphi(t)\big)$) к Солнцу, расположенному в начале координат, равна по закону всемирного тяготения

$$F(x,y) = -kr^{-2} = -\frac{k}{\sqrt{x^2+y^2}}.$$

Кинетическая энергия T тела равна

$$\frac{m}{2}\big(\dot{x}^2(t) + \dot{y}^2(t)\big) = \frac{m}{2}\left(\frac{d}{dt}\big(r(t)\cos\varphi(t)\big)^2 + \frac{d}{dt}\big(r(t)\sin\varphi(t)\big)^2\right) = \\ = \frac{m}{2}\big(\dot{r}^2(t) + r^2(t)\dot{\varphi}^2(t)\big),$$

потенциальная энергия U равна

$$\frac{k}{r(t)}.$$

Лагранжиан $T - U$ не зависит от времени, следовательно, уравнение Эйлера имеет интеграл энергии, выражающий *закон сохранения энергии*

$$\frac{m}{2}\big(\dot{r}^2(t) + r^2(t)\dot{\varphi}^2(t)\big) + \frac{k}{r(t)} = H.$$

Он не зависит также от φ, и потому имеет место *закон сохранения импульса*

$$mr^2(t)\dot{\varphi}(t) = mL,$$

который переписывается в виде второго кеплеровского закона площадей:

$$\frac{d\varphi}{dt} = \frac{L}{r^2}.$$

Закон сохранения энергии (с учётом закона сохранения импульса) перепишем в виде

$$\frac{m}{2}\left(\dot{r}^2(t) + \frac{L}{mr^2(t)}\right) + \frac{k}{r(t)} = H,$$

что равносильно дифференциальному уравнению

$$\frac{dr}{dt} = \sqrt{\frac{2}{m}\left(H - \frac{mk}{r}\right) - \frac{L^2}{m^2r^2}}.$$

Поделив $d\varphi/dt$ на dr/dt, приходим к интегрируемому дифференциальному уравнению

$$\frac{d\varphi}{dr} = \frac{L}{r^2\sqrt{\frac{2}{m}\left(H - \frac{mk}{r}\right) - \frac{L^2}{m^2r^2}}},$$

которое можно проинтегрировать подстановкой Эйлера. Произведя интегрирование, получим

$$r = \frac{p}{1 + e\cos\varphi}, \quad p = \frac{L^2}{|k|}, \quad e = \sqrt{\frac{2HL^2}{mk^2} + 1}, \quad L = \sqrt{|k|p}.$$

Итак, небесные тела под воздействием Солнца движутся по кривым второго порядка. При этом если $e < 1$, то движение проходит по эллипсу (этим доказан первый закон Кеплера). Закон сохранения импульса представляет собой второй закон Кеплера. Докажем третий закон. Интегрируя соотношение закона площадей по периоду и используя то, что площадь эллипса с полуосями a и b равна πab, приходим к равенству $\pi ab = LT$, откуда, используя то, что $b^2/a = p$, $L^2 = |k|p$, получаем, что

$$T^2 = \frac{\pi^2a^2b^2}{L^2} = \frac{a^3}{|k|}.$$

Это третий закон Кеплера.

Дополнение

Доказательство леммы о правом обратном. В доказательстве используются теорема Бэра о категории и модифицированный метод Ньютона.

Пусть $k \in \mathbb{N}$ и $U_k = U_X(0,k) = \{x \in X \mid \|x\|_X < n\}$ — открытый шар в X с центром в нуле радиуса n. Оператор A сюръективен, $\bigcup\limits_{k\in\mathbb{N}} AU_k = Y$. Из теоремы Бэра о категории следует, что существуют $N \in \mathbb{N}$, $y_0 \in Y$ и $r > 0$ такие, что $y_0 + U_Y(0,r) \subset \mathrm{cl}\,\Lambda U_N$, где cl обозначает замыкание. Поскольку множество AU_N выпукло и симметрично, то нетрудно проверить, что этими же свойствами обладает его замыкание. Отсюда следует, что $U_Y(0,r) \subset \mathrm{cl}\,AU_N$.

Пусть $\eta \in U_Y(0,r)$. Найдём элемент $x(\eta)$, такой что

$$\|\eta - Ax(\eta)\|_Y \leqslant \frac{r}{4}. \tag{i}$$

Теперь для произвольного $y \in Y \setminus \{0\}$ положим

$$\bar{R}(y) = 2r^{-1}\|y\|_Y x(r2^{-1}\|y\|_Y^{-1}y), \quad \bar{R}(0) = 0.$$

Тогда из (i) и непосредственно из определения $\bar{R}$ вытекает, что

$$\|y - A\bar{R}(y)\|_Y \leqslant \frac{\|y\|_Y}{2}, \quad \|\bar{R}(y)\|_X \leqslant C_1\|y\|_Y, \tag{ii}$$

где $C_1 = 2N/r$. Отображение $\bar{R}$ — «почти» правое обратное к A. Для построения правого обратного применяем модифицированный метод Ньютона:

$$x_k = x_{k-1} + \bar{R}(y - Ax_{k-1}), \quad k \in \mathbb{N}, \quad x_0 = 0. \tag{iii}$$

Имеем тогда

$$\|y - Ax_k\|_Y \overset{\text{(iii)}}{=} \|y - Ax_{k-1} - \Lambda\bar{R}(y - \Lambda x_{k-1})\|_Y \overset{\text{(i)}}{\leqslant} \frac{\|y - Ax_{k-1}\|_Y}{2} \leqslant \ldots \leqslant \frac{\|y\|_Y}{2^k}, \tag{iv}$$

следовательно,

$$\|x_{k+1} - x_k\|_X \overset{\text{(iii)}}{\leqslant} \|\bar{R}(y - Ax_k)\|_Y \overset{\text{(ii)}}{\leqslant} C_1\|y - Ax_k\|_Y \overset{\text{(iv)}}{\leqslant} \frac{C_1\|y\|_Y}{2^k}$$

и

$$\|x_{k+m} - x_k\|_X \leqslant \frac{C_1}{2^{k-1}}\|y\|_Y, \tag{v}$$

откуда следует, что последовательность $\{x_k\}_{k\in\mathbb{Z}_+}$ фундаментальна. Теперь, если положить $R(y) = \lim\limits_{k\to\infty} x_k$, то, переходя к пределу в (iv), получаем, что $\Lambda R(y) = y$, а переходя к пределу в неравенстве

$$\|x_k\|_X \leqslant \|x_k - x_{k-1}\|_X + \ldots + \|x_1 - x_0\|_X \overset{\text{(v)}}{\leqslant} 2C_1\|y\|_Y,$$

получаем $\|R(y)\|_Y \leqslant C\|y\|_Y$, где $C = 2C_1$, что и требуется. □

Доказательство леммы II.1. Пусть $\|Bf^i - Bf^j\|_X \geqslant \alpha$, $i \neq j$. Выберем N настолько большим, чтобы выполнялось

$$\|B - B_N\|_{\mathcal{L}(X,X)} \leqslant \frac{\alpha}{4}, \tag{vi}$$

где $B_N = (b_{ij})_{1\leqslant i,j\leqslant N}$. Тогда

$$\begin{aligned} \|B_Nf^i - B_Nf^j\|_X &\overset{\text{Id}}{=} \|Bf^i - Bf^j + B_Nf^i - Bf^i - B_Nf^j + Bf^j\|_X \overset{\text{TI}}{\geqslant} \\ &\geqslant \|Bf^i - Bf^j\|_X - \|B_Nf^i - Bf^i\|_X - \|B_Nf^j - Bf^j\|_X \end{aligned} \tag{vii}$$

(здесь «TI» означает «неравенство треугольника» (Triangle Inequality)). Но из определения нормы вытекает, что

$$\|B_Nf^i - Bf^i\|_X \leqslant \|B_N - B\|_{\mathcal{L}(X,X)}\|f^i\|_X \overset{\text{(vi)}}{\leqslant} \frac{\alpha}{4},$$

аналогично

$$\|B_Nf^j - Bf^j\|_X \leqslant \frac{\alpha}{4}.$$

Отсюда и из (vii) следует, что

$$\|B_Nf^i - B_Nf^j\|_X \geqslant \frac{\alpha}{2}.$$

Мы свели дело к конечномерному случаю. Поскольку N-мерный шар $B_{\mathbb{R}^N}(0,R)$ (с центром в нуле радиуса $R = \|B_N\|_{\mathcal{L}(X,X)}$) является компактным множеством в $\mathbb{R}^N$, в нём нельзя расположить бесконечно много точек x^i, находящихся друг от друга на расстоянии, большем $\alpha/2$. Лемма доказана. $\square$

Доказательство леммы II.2. Пусть X — пространство l_2, $B\colon X \to X$ — предкомпактный линейный оператор, y принадлежит замыканию образа оператора $A = \mathrm{Id} - B$ (Id — единичный оператор) и $Ax_n \to y$. При этом можно считать, что x_n ортогональны $\operatorname{Ker} A$ (ибо всегда можно вычесть из x_n проекцию на $\operatorname{Ker} A$). Но тогда нормы $\|x_n\|_X$ ограничены. Действительно, если $\|x_n\|_X \to \infty$, то из последовательности

$$\left\{B\frac{x_n}{\|x_n\|_X}\right\}_{n\in\mathbb{N}}$$

можно выбрать сходящуюся подпоследовательность

$$\left\{B\frac{x_{n_k}}{\|x_{n_k}\|_X}\right\}_{k\in\mathbb{N}},$$

а тогда и сама последовательность

$$\left\{\frac{x_{n_k}}{\|x_{n_k}\|_X}\right\}_{k\in\mathbb{N}}$$

сходится, так как

$$\frac{x_{n_k}}{\|x_{n_k}\|_X} = A\frac{x_{n_k}}{\|x_{n_k}\|_X} + B\frac{x_{n_k}}{\|x_{n_k}\|_X},$$

а первое слагаемое стремится к нулю. Предел подпоследовательности

$$\left\{\frac{x_{n_k}}{\|x_{n_k}\|_X}\right\}_{k\in\mathbb{N}}$$

обозначим через z. Тогда $\|z\|_X = 1$ и $Az = 0$. Мы пришли к противоречию и тем самым доказали, что нормы $\|x_n\|_X$ ограничены. Выбрав из последовательности $\{Bx_n\}_{n\in\mathbb{N}}$ сходящуюся подпоследовательность $\{Bx_{n_l}\}_{l\in\mathbb{N}}$ (B — компактный оператор), приходим к тому, что сама последовательность $\{x_{n_l}\}_{l\in\mathbb{N}} = \{y + Bx_{n_l}\}_{l\in\mathbb{N}}$ сходится. Если предел обозначить через ζ, то получится, что $y = \zeta - B\zeta$, что и требовалось. $\square$

Доказательство леммы II.3. Пусть $\bar{x}$ — точка, не принадлежащая замкнутому подпространству $L \subset X$, и $\{y_n\}_{n\in\mathbb{N}}$ — последовательность векторов из L, таких что

$$\|\bar{x} - y_n\|_X \to \inf_{y\in Y} \|\bar{x} - y\|_X =: d.$$

Если при $n \geqslant N$ выполняется неравенство $\|\bar{x} - y_n\|_X - d < \delta$, то для $m, n \geqslant N$ из равенства параллелограмма

$$\|y_n - y_m\|_X^2 = 2\|\bar{x} - y_m\|_X^2 + 2\|\bar{x} - y_n\|_X^2 - 4\left\|\bar{x} - \frac{y_1 + y_2}{2}\right\|_X^2$$

следует, что если $\|x - y_n\|_X$ и $\|x - y_m\|$ малы, то и норма разности $\|y_n - y_m\|_X$ мала, т. е. последовательность $\{y_n\}_{n\in\mathbb{N}}$ фундаментальна. Вследствие полноты X и замкнутости L последовательность y_n сходится к некоторому элементу $\bar{y} \in L$, и при этом $d = \|\bar{x} - \bar{y}\|_X$. Значит, для любого $y \in Y$ функция $f(t) = \langle \bar{x}-\bar{y}-ty, \bar{x}-\bar{y}-ty\rangle$ достигает минимума в нуле. Дифференцируя f и полагая $t = 0$, получаем, что $\langle \bar{x} - \bar{y}, y\rangle$, т. е. $\bar{x} - \bar{y}$ — перпендикуляр к L. $\square$