В. М. Тихомиров

КРАТКИЙ КУРС МАТЕМАТИКИ

Москва, 2007

В цикле из двух лекций я задумал рассказать о самом важном из того, чему меня учили в школе и на первых двух курсах университета. Мне (по многим причинам) очень существенно знать ваше мнение о том, в какой мере мне это удастся.

В этом конспекте изложение разбито на три главы, названия которых соответствует курсам, которые мне читали в первые два года студенчества: линейная алгебра, аналитическая геометрия, математический анализ и дифференциальные уравнения (последние два раздела я объединяю в одной третьей главе). При этом мне хотелось бы показать, что математика — единая наука, и все темы и разделы переплетены и взаимосвязаны. Размеры текста таковы, чтобы, кое-что сократив, сделать попытку рассказать обо всем за время, предоставляемое в Дубне для двух лекций. В каждой главе делается всего три шага — от одного к двум, затем к произвольному (натуральному) числу, и в конце — к бесконечности. Первые два параграфа посвящены математике, с которой я знакомился в школе и на кружках, изучая функции одного переменного и плоские фигуры, третий посвящен, в основном, институтской математике, где рассматриваются уравнения со многими переменными, и некоторые трехмерные фигуры, четвертый — математике университетской, где число переменных бывает и бесконечным. И всюду мы будем двигаться от самых истоков к тому, что было осознано в конце девятнадцатого и в первой половине двадцатого века. Постараюсь изложить доказательства важнейших результатов.

Данная краткая брошюра представляет собой нечто вроде “книги для чтения”, и читатель может в каждой главе, начав с того, что он заведомо знает, продвигаться дальше и дальше к финалу, имея , разумеется, возможность в любой момент “с достоинством выйти из игры” (как изящно сказал как-то по сходному поводу мой старинный приятель Роберт Адольфович Минлос).

Готовясь к этим лекциям, я советовался со своими друзьями и коллегами. Хочу выразить всем им свою благодарность. Особую признательность мне хотелось бы высказать М. Л. Гольдману, Г. Г. Магарил-Ильяеву, М. А. Мейснер, К. Ю. Осипенко и Е. О. Сивковой, дружеская критика которых была для меня весьма значимой. Надеюсь на критику и диалог с моими слушателями и читателями.

Содержание

Глава 1. Линейная алгебра (теория линейных уравнений)

Алгебра (от арабского слова ALD$\widehat{Z}$ABR (ал-джебр))[1] — часть математики, развившаяся в связи с потребностью решать уравнения. Уравнением называется равенство с одним или несколькими неизвестными. Если неизвестное или неизвестные входят в первой степени, уравнения называются *линейными.* Фрагмент алгебры, в которой изучают линейные уравнения, называют *линейной алгеброй.* Необходимость решать линейные уравнения возникает в огромном числе случаев, когда ищется ответ в конкретной прикладной задаче; в некотором отношении — это база прикладной математики. Данная глава посвящена *теории* линейных уравнений, ее цель — **научить методам решения этих уравнений и условиям их разрешимости**. Исторические рамки этой главы простираются от истоков (от Древнего Египта) до начала двадцатого века.

§ 1.1. Одно линейное уравнение с одним неизвестным. Пусть a и b — числа (совокупность всех вещественных чисел обозначают буквой $\mathbb{R}$; вещественным числам посвящен § 1 Дополнения). *Требуется найти все такие числа x, чтобы $ax = b$* или иначе — требуется решить уравнение

$$ax = b, \tag{1}$$

т. е. указать метод отыскания решений и описать условия разрешимости уравнения. Имеет место

Теорема 1.1 (о разрешимости уравнения (1) **с одним неизвестным).** *Имеются три возможности: 1) $a \neq 0$ и тогда уравнение однозначно разрешимо: $x = \frac{b}{a}$; 2) $a = 0$ и $b = 0$; тогда решением является любое число; наконец, 3) $a = 0$, а $b \neq 0$; тогда решений нет.*

Задачи, сводящиеся к (1) содержатся в древнейшем из дошедших до нас текстов, где обсуждаются математические сюжеты, — так называемом папирусе Райнда, созданном в Древнем Египте где-то около четырех тысяч лет тому назад. Одна из задач из этого папируса такова: *к числу прибавлена его седьмая часть и получилось девятнадцать; каково это число?* В этой задаче $x + \frac{x}{7} = \frac{8}{7}x = 19$, откуда, согласно "теории", $x = \frac{19\cdot 7}{8} = \frac{133}{8} = 16\frac{5}{8}$. Это и есть искомое число.

[1] В своем учебнике по алгебре для школьников П. С. Александров и А. Н. Колмогоров пишут, что этим словом арабские ученые называли операцию прибавления одних и тех же выражений к обеим частям уравнения, не меняя его корней.

§ 1.2. Система двух линейных уравнений с двумя неизвестными. Это, пожалуй, центральный параграф всей главы: здесь будет рассказано об основном методе решения систем линейных уравнений (методе Гаусса), приведены формулы Крамера и доказаны начальные варианты альтернативы Фредгольма и теоремы Кронекера-Капелли. Формулировки этих результатов сохранятся и в конечномерном и в бесконечномерном случаях.

Начнем, как всегда, с постановки задачи. Пусть заданы 6 чисел $a_{11}, a_{12}, a_{21}, a_{22}, b_1$ и b_2. Требуется найти такие два числа x_1 и x_2, чтобы выполнялась равенства:

$$a_{11}x_1 + a_{12}x_2 = b_1, \quad a_{21}x_1 + a_{22}x_2 = b_2. \tag{2}$$

В эти уравнения неизвестные x_1 и x_2 входят в первой степени, и потому система (2) называется *системой двух линейных уравнений с двумя неизвестными.* Числа $a_{11}, a_{12}, a_{21}, a_{22}$ называются *коэффициентами уравнения*, пара чисел b_1 и b_2 называется *правой частью выписанной системы.* Если $b_1 = b_2 = 0$, система (2) называется *однородной.*

Решать системы двух уравнений с двумя неизвестными также научились давным-давно. В китайском трактате "Семь частей искусства математики" (его относят ко второму веку нашей эры, к периоду династии Хань) обсуждается такая задача:

В клетке фазаны и кролики. У них вместе 35 голов и 94 ноги. Сколько в клетке фазанов и сколько кроликов?

Если число фазанов обозначить через x_1, а число кроликов через x_2, мы приходим к уравнениям: $x_1 + x_2 = 35,\ 2x_1 + 4x_2 = 94$.

Опишем *метод* решения любой системы двух уравнений с двумя неизвестными. Этот метод называют *методом Гаусса* (Карл Фридрих Гаусс (1777–1855) — один из величайших математиков всех времен). Метод Гаусса — это *метод исключения неизвестных.* При рассмотрении выписанной выше системы возможен вырожденный случай, когда все коэффициенты нули (и тогда разобраться очень просто: если правые части тоже нули, то любая пара чисел x_1 и x_2 — решение, а если b_1 и/или b_2 не равны нулю, то решения системы нет или, как говорят, *система несовместна*). Остается рассмотреть случай, когда какой-то коэффициент системы (2) отличен от нуля. Тогда соответствующее переменное можно выразить через другое и подставить в другое уравнение. В итоге мы приходим к одному уравнению с одним неизвестным, которое уже научились исследовать. В задаче о кроликах и фазанах из первого уравнения получаем $x_2 = 35 - x_1$, откуда из второго уравнения вытекает, что $2x_1 + 140 - 4x_1 = 94$ или $2x_1 = 46$, а значит, фазанов 23, кроликов $35 - 23 = 12$ — задача решена.

Скажем о методе Гаусса и решении системы уравнений (2) чуть подробнее и "математичнее". Таблицу $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ из коэффициентов уравнения назовем *матрицей коэффициентов* (второго порядка). Расположим пару чисел b_1 и b_2 в столбец, назовем вектором-столбцом и обозначим его $\bar{b}$, аналогично искомую пару x_1 и x_2 расположим в столбец и обозначим $\bar{x}$. Определим произведение матрицы A на вектор $\bar{x}$ следующим образом. Произведение $A\bar{x}$ — это вектор-столбец, верхнее число которого равно $a_{11}x_1 + a_{12}x_2$, а нижнее — $a_{21}x_1 + a_{22}x_2$, т. е. $A\bar{x} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$, и уравнение (2) может быть теперь кратко записано, как $A\bar{x} = \bar{b}$.

Выражение $\Delta = a_{11}a_{22} - a_{12}a_{21}$ называют *определителем (или детерминантом) матрицы* A и обозначают $\det A$. Равенство нулю детерминанта означает, что строки матрицы A (и ее столбцы) пропорциональны (если $\det A = 0$, то $a_{11}a_{22} = a_{12}a_{21}$ или $\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} = \alpha$, т. е. $a_{11} = \alpha a_{21}, a_{12} = \alpha a_{22}$). Значит, если определитель не равен нулю, строки матрицы A (и ее столбцы) *не пропорциональны*. Введем еще две матрицы:$A_1 = \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}$ и $A_2 = \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}$. Их определители обозначим, соответственно, Δ_1 и Δ_2. Теперь можно сформулировать основной результат этого параграфа, описывающий все решения системы двух линейных уравнений с двумя неизвестными.

Теорема 1.2 (о свойствах и разрешимости системы (2) **двух линейных уравнений с двумя неизвестными)** *a) Для системы* (2) *(или $A\bar{x} = \bar{b}$, $\bar{b}, \bar{x} \in \mathbb{R}^2$) имеет место альтернатива: либо эта система разрешима для любого $\bar{b}$, либо однородное уравнение имеет ненулевое решение. b) Система* (2) *разрешима тогда и только тогда, когда для любых (y_1, y_2), для которых $y_1a_{11}+y_2a_{21} = 0, y_1a_{12}+y_2a_{22} = 0$, выполняется равенство $y_1b_1 + y_2b_2 = 0$. c) Уравнение* (2) *разрешимо тогда и только тогда, когда* $\det A \neq 0$*; решением являются числа* $x_1 = \frac{\det A_1}{\det A}$, $x_2 = \frac{\det A_2}{\det A}$.

Мы видим, что и здесь имеются те же возможности, что и в теореме 1.1, когда система разрешается однозначно (будем называть такую систему *невырожденной*), когда она разрешается неоднозначно и когда она вообще не может быть разрешена.

Формулировки a) и b) теоремы (с естественными уточнениями) сохранятся и в общем конечномерном случае (это будет доказано здесь), и в бесконечномерном случае, который обсуждается позже. Результат

c) также верен в общем конечномерном случае. Этот случай будет рассмотрен в следующей главе.

Доказательство. Начнем с формального описания метода Гаусса. Случай нулевой матрицы A был разобран — там никакого "метода не потребовалось". Пусть матрица A ненулевая и значит, какой-то этой элемент матрицы не равен нулю. Переименованием координат и переменой порядка уравнений всегда можно добиться, чтобы $a_{22} \neq 0$. Действительно, если, скажем, в матрице A не равен нулю элемент a_{11}, переобозначим переменные (x_1 на x_2 и наоборот) и первое уравнение сделаем вторым, а второе — первым. Итак, пусть $a_{22} \neq 0$. Выразим тогда x_2 через x_1: $x_2 = \frac{b_1 - a_{21}x_1}{a_{22}}$, и, подставив это в первое уравнение и совершив тождественные преобразования (проделайте их самостоятельно), получим равенство: $\frac{\Delta}{a_{22}}x_2 = \frac{\Delta_1}{a_{22}}$. Тогда, если $\Delta = 0$, приходим к равенству $0x_2 = \frac{\Delta_1}{a_{22}}$. Это равенство возможно лишь если $\Delta_1 = 0$, а тогда, как было объяснено выше $\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} = \frac{b_1}{b_2}$, т. е. имеет место случай 2). Неравенство нулю Δ_1 равносильно тому, что тройки (a_{11}, a_{12}, b_1) и (a_{21}, a_{22}, b_2) непропорциональны, а невозможность равенства $0x_2 = \frac{\Delta_1}{a_{22}}$ означает несовместность системы (2). Если же $\Delta \neq 0$, получаем, что $x_1 = \frac{\Delta_1}{\Delta}$, и теперь, если подставить это число в выражение для x_2, приходим к равенству $x_2 = \frac{\Delta_2}{\Delta}$. Нетрудно понять, что формулировка пункта 2) теоремы равносильна совокупности доказанных условий. Доказательство закончено. □

Отметим еще, что одно однородное уравнение $a_1x_1 + a_2x_2 = 0$ с двумя неизвестными всегда имеет ненулевое решение и два уравнения $a_{11}x_1+a_{12}x_2+a_{13}x_3 = 0$, $a_{11}x_1+a_{12}x_2+a_{13}x_3 = 0$ с тремя неизвестными также имеют нетривиальное решение. Действительно, в первом случае если $a_1 = a_2 = 0$, любые x_1 и x_2 — решения, а если, скажем $a_1 \neq 0$, то решением будем $x_2 = a_1$, $x_1 = -a_2$. Во втором случае, если все нули, все просто, а если, скажем, $a_{11} \neq 0$, то исключим x_1 и дело сведется к уже решенной задаче.

Мы совершили переход от одного к двум. Наш следующий шаг — от двух к произвольному числу n.

§ 1.3. Система n линейных уравнений с n неизвестными.

Пусть заданы $n^2 + n$ чисел $a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{n1}, a_{n2}, \ldots, a_{nn}$,
$b_1, b_2, \ldots, b_n$. Требуется найти такие n чисел $x_1, x_2 \ldots, x_n$, чтобы выполнялись равенства:

$$a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = b_1, \ \ldots, \ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n = b_n. \quad (3)$$

Переведем разговор на язык векторов и матриц. n-ку чисел, расположенных в столбец называют *вектором-столбцом.* Совокупность векторов-столбцов $\bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ обозначают $\mathbb{R}^n$. Наряду с пространством векторов-столбцов оказывается естественным и плодотворным рассматривать "дуальное" пространство векторов-строк $\bar{y} = (y_1, \ldots, y_n)$. Обозначим его $(\mathbb{R}^n)'$. Эти два пространства соединены *внутренним произведением*: если $\bar{y} \in (\mathbb{R}^n)', \bar{x} \in \mathbb{R}^n$, то $\bar{y} \cdot \bar{x} = y_1x_1 + \ldots + y_nx_n$.

Матрица $A = (a_{ij})_{1 \le i,j \le n}$ состоит из n столбцов (обозначим их $\bar{a}^1, \ldots, \bar{a}^n$) и n строк (обозначим их $\bar{\alpha}^1, \ldots, \bar{\alpha}^n$). Произведением $A\bar{x}$ матрицы A на вектор-столбец $\bar{x}$ называется вектор-столбец $\bar{x} = \begin{pmatrix} \bar{\alpha}^1 \cdot \bar{x} \\ \vdots \\ \bar{\alpha}^n \cdot \bar{x} \end{pmatrix}$.

Таким образом, изучавшаяся нами система (3), кратко записывается $A\bar{x} = \bar{b}$ $(\bar{x}, \bar{b} \in \mathbb{R}^n)$.

Замечание. Матрицы можно перемножать. Рассмотрим две матрицы $A = (a_{ij})_{1 \le i,j \le n}$ и $B = (b_{ij})_{1 \le i,j \le n}$ Строки матрицы A обозначим $\bar{\alpha}^1, \ldots, \bar{\alpha}^n$, столбцы матрицы B — $\bar{b}^1, \ldots, \bar{b}^n$. Произведением $C = AB$ матриц A и B называется матрица $c_{ij} = \bar{\alpha}^i \cdot \bar{b}^j$. Отметим, что если у матрицы второго порядка $\det A \neq 0$, матрица $A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$ обладает тем свойством, что $A^{-1}A = AA^{-1} = I$, где $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ — единичная матрица. Такая матрица называется *обратной*.

Произведением $\bar{y}A$ вектора-строки $\bar{y}$ на матрицу A называется вектор-строка $(y_1a_{11} + \ldots + y_na_{n1}, \ldots, y_1a_{1n} + \ldots + y_na_{nn})$. Уравнения $A\bar{x} = \bar{b}$, $\bar{x}, \bar{b} \in \mathbb{R}^n$ и $\bar{y}A = \bar{c}$, $\bar{y}$, $\bar{c} \in (\mathbb{R}^n)'$ называются *сопряженными друг с другом;* если одно из уравнений объявлено *основным*, второе называют *сопряженным* ему.

Теорема 1.3 (о свойствах и разрешимости системы n линейных уравнений с n неизвестными) *a) Для системы* (3) *(или $A\bar{x} = \bar{b}$, $\bar{x}, \bar{b} \in \mathbb{R}^n$) имеет место альтернатива: либо эта система разрешима для любого $\bar{b}$, либо однородное уравнение имеет ненулевое решение. b) Система* (3) *разрешима тогда и только тогда, когда внутреннее произведение решения сопряженного однородного уравнения на правую часть основного уравнения равно нулю.*

Доказательство проведем на примере трех уравнений с тремя неизвестными:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{aligned} \qquad (3')$$

Мы так поступаем, во-первых, потому, что это нам понадобится в следующей главе, а во вторых, потому, что общий случай точно такой же. Сначала доказывается (точно также, как в конце § 1.2) утверждение, согласно которому три уравнения с четырьмя неизвестными имеют нетривиальное решение. Снова расположим коэффициенты уравнений в таблицу (матрицу третьего порядка)$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Числа b_i в правой части тоже расположим в столбец $\bar{b}$. Обозначим первый столбец A через $\bar{a}^1$, второй через $\bar{a}^2$, третий через $\bar{a}^3$. Решение — три числа x_1, x_2, x_3, расположенные в столбец будем обозначать единым символом $\bar{x}$, а саму систему $(3')$ кратко запишем в виде $A\bar{x} = \bar{b}$. Пусть однородное уравнение имеет только нулевое решение. Согласно утверждению найдутся такие четыре числа ξ_1, ξ_2, ξ_3 и ξ_4 не в равные одновременно нулю (далее НЕРОН), что $\bar{a}^1\xi_1 + \bar{a}^2\xi_2 + \bar{a}^3\xi_3 + \bar{b}\xi_4 = 0$ (мы кратко записали, что ξ_i, $1 \le i \le 4$ — решение трех уравнений с четырьмя неизвестными, а еще короче могли бы написать так: $A\bar{\xi} = 0$). Но **число ξ_4 не может быть нулем!** (ибо мы предположили, что однородная система не имеет ненулевого решения). Поделив на $-\xi_4$ и обозначив $x_i = \frac{\xi_i}{-\xi_4}, i = 1, 2, 3,$ получим решение нашей системы. В одну сторону мы теорему доказали.

Теперь в другую. Допустим, что нашу систему мы можем решить для любой правой части, а однородное уравнение имеет ненулевое решение и попробуем придти к противоречию. По предположению существует ненулевое решение $\bar{e}^1$ уравнения $A\bar{x} = 0$ ($\bar{e}^1$ — ненулевое решение однородного уравнения). По условию же можно решить уравнение $A\bar{x} = \bar{e}^1$. Обозначим его $\bar{e}^2$, аналогично построим $\bar{e}^3$ и $\bar{e}^4$. И снова по утверждению о трех уравнениях с четырьмя неизвестными найдем четыре числа z_1, z_2, z_3 и z_4 (НЕРОН), что $\bar{e} = \bar{e}^1z_1 + \bar{e}_2z_2 + \bar{e}_3z_3 + \bar{e}_4z_4 = 0$. Имеем $\bar{e} = 0$, значит, $A\bar{e} = 0$, откуда (воспользовавшись тем, что если $\bar{x}$ и $\bar{y}$ — два набора по три числа, то $A(\bar{x} + \bar{y}) = A\bar{x} + A\bar{y}$) получим: $A(\bar{e}^1z_1 + \bar{e}^2z_2 + \bar{e}^3z_3 + \bar{e}^4z_4) = A\bar{e}^1z_1 + A\bar{e}^2z_2 + A\bar{e}^3z_3 + A\bar{e}^4z_4)$. Но $A\bar{e}^1 = 0, A\bar{e}^2 = e^1, A\bar{e}^3 = e^2, A\bar{e}^4 = e^3$, значит, $\bar{e}_1z_2 + \bar{e}_2z_3 + \bar{e}_3z_4 = 0$. Проделав еще два раза этот трюк, придем к тому, что $e^1z_4 = 0$ откуда $z_4 = 0$. Аналогично докажем, что и z_3 и z_2 и z_1 равны нулю. Вошли в противоречие с "НЕРОНом". Теорема полностью доказана. □

Метод решения трех уравнений — метод Гаусса — с тремя неизвестными проходит по той же схеме: если матрица коэффициентов не нулевая, можно считать не ограничившись в общности, что $a_{33} \neq 0$. Выразим x_3 из последнего уравнения через x_1 и x_2 и, подставив в первые два, получим два уравнения с двумя неизвестными. А их мы решать умеем. Итак, мы научились решать систему из трех уравнений с тремя неизвестными, а значит, проведя совершенно аналогичный прием, решим систему четырех уравнений с четырьмя неизвестными и так далее.

Выведем из доказанной альтернативы условие разрешимости уравнения (3). Пусть задана матрица $A = (a_{ij})_{1\le i,j\le n}$ размеров n на n. Столбцы этой матрицы обозначим $\bar{a}^1, \ldots, \bar{a}^n$. Говорят, что столбцы $\bar{a}^j$, $1 \le j \le n$ *линейно независимы,* если не существует n чисел $x_1, \ldots, x_n$ НЕРОН таких, что $x_1\bar{a}^1 + \ldots + x_n\bar{a}^n = \bar{0}$ ($\bar{0}$ — столбец из n нулей. Мы видим, что сказать:"столбцы матрицы A линейно независимы" — то же самое, что сказать: "однородное уравнение (3) имеет только нулевое решение." Но по теореме 1.3, если однородная система имеет только нулевое решение, то система разрешима. Отсюда следует, что и строки $\alpha^i = (a_{i1}, \ldots, a_{in})$, $1 \le i \le n$ матрицы A также линейно независимы. Действительно, если это не так, найдутся числа $y_1, \ldots, y_n$ НЕРОН такие, что $y_1\bar{\alpha}^1 + \ldots y_n\bar{\alpha}^n = 0$. Пусть, скажем, $y_n \neq 0$. Тогда, разделив на $-y_n$, получим, что α^n "линейно выражается" через остальные: $\alpha^n = x_1\alpha^1 + \ldots + x_{n-1}\alpha_{n-1}$ и значит, уравнение $A\bar{x} = \bar{b}$, где столбец $\bar{b}$ состоит из всех нулей, кроме n-того числа, равного единице, решить невозможно. Противоречие.

Пусть теперь $B = (b_{ij})$, $1 \le j \le n, 1 \le i \le m$ — матрица размеров n (столбцов) на m (строк). *Рангом* этой матрицы (обозначаемым $\mathrm{rank}B$) назовем максимальный размер квадратной подматрицы матрицы B, у которой ее строки (и столбцы) линейно независимы. Если A — матрица системы (3), то матрица $\widetilde{A}$, получаемая присоединением к A столбца $\bar{b}$ правых частей, называется *расширенной матрицей.*

Теорема Кронекера–Капелли *Для того, чтобы уравнение* (3) *было разрешимым, необходимо и достаточно, чтобы ранг матрицы* A *был бы равен рангу матрицы* $\widetilde{A}$.

Доказательство. Если ранг матрицы A системы (3) равен n, система разрешима, ранг матрицы $\widetilde{A}$ тоже равен n, и все доказано. Допустим, что $\mathrm{rank}A = n-1$ (общий случай доказывается аналогично). Не ограничив общности можно считать, что матрица с независимыми строками и столбцами размера $(n-1)\times(n-1)$ расположена в левом верхнем углу. Тогда последний столбец линейно выражается через первые $n-1$ столбец (иначе $\mathrm{rank}A$ был бы равен n). Если система разрешима, то число линейно независимых строк остается равным $n-1$, т.е. $\mathrm{rank}\widetilde{A} = n-1$.

Если $\mathrm{rank}\widetilde{A} = n-1$, то последняя строка матрицы $\widetilde{A}$ выражается через остальные, а это означает разрешимость системы (3). □

§ 1.4. Система бесконечного числа линейных уравнений с бесконечным числом неизвестных.

В §1.3 изучались уравнения $A\bar{x} = \bar{b}$, где $\bar{x}, \bar{b} \in \mathbb{R}^n$. Здесь в бесконечномерном случае векторы $\bar{x} = (x_1, x_2, \ldots, x_k \ldots)$ расположим в строку и будем накладывать такое требование: $|\bar{x}|^2 = \sum_{k=1}^{\infty} x_k^2 < \infty$. Совокупность таких векторов обозначают l_2. В l_2 вводится скалярное произведение $(\bar{x}, \bar{y}) = \sum_{k=1}^{\infty} x_k y_k$. То, что из $\bar{x} \in l_2$ и $\bar{y} \in l_2$ следует, что $\sum_{k=1}^{\infty} x_k y_k < \infty$, вытекает из неравенства $(\bar{x}, \bar{y}) \le |\bar{x}||\bar{y}|$, а отсюда сразу следует, что векторы из l_2 можно складывать; разумеется, их можно умножать на вещественные числа, т. е. они образуют векторное пространство (см. Дополнение §2).) Не все, но некоторые матрицы $A = (a_{ij})_{1\le i,j<\infty}$ порождают отображение из l_2 в l_2, сопоставляя вектору $\bar{x} \in l_2$ вектор $\bar{y} = A\bar{x}$, $\bar{y} = (\sum_{i=1}^{\infty} a_{1i}x_i, \ldots, \sum_{i=1}^{\infty} a_{1i}x_i, \ldots)$ из l_2, но заведомо так будет в случае матриц, у которых $a_{ij} \neq 0$ лишь если $1 \le i, j \le n$. Операторы, порожденные такими матрицами, называют вырожденными, а матрицы, порождающие оператор, действующий из l_2 в l_2 будем называть допустимыми. Таким образом, нам предстоит исследовать уравнения

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{1k}x_k + \ldots = b_i, \ 1 \le i < \infty, \ \Leftrightarrow A\bar{x} = \bar{b}, \ \bar{x}, \bar{b} \in l_2, \quad (4)$$

где A — допустимый оператор. Матрицу A^T, в которой i-тый столбец занимает положение i-той строки (для всех i) называют транспонированной. Систему уравнений

$$A^T y = 0 \ \Leftrightarrow \ y_1 a_{1j} + y_2 a_{2j} + \ldots + y_k a_{kl} + \ldots = 0, \ 1 \le j < \infty,$$

назовем сопряженным однородным уравнением. Будем изучать уравнения (4) не с любыми допустимыми операторами, а с представимыми в виде $A = I + B$, где I — единичный оператор, а B — вырожденный или являющийся пределом вырожденных (такие операторы называют компактными). Не будем определять этого здесь, сделаем это в конце следующей главы, когда будем проводить доказательство. А здесь лишь приведем формулировку теоремы об альтернативе Фредгольма:

Теорема 1.4. (Альтернатива Фредгольма) *Для системы* (4) *(или* $A\bar{x} = \bar{b}$, $\bar{x}, \bar{b} \in \mathbb{R}^n$*), в которой* A *— сумма единичного и компактного оператора, имеет место альтернатива: либо эта система разрешима для любого* $\bar{b}$*, либо однородное уравнение имеет ненулевое решение. .*

Эта теорема (доказанная шведским математиком Эриком Иваром Фредгольмом (1866–1927) в 1900 году) — одна из вершин университетского математического образования (ее проходят на третьем курсе). Мы дадим эскиз доказательства этой теоремы в следующей главе — удобнее проводить его в главе, посвященной геометрии, ибо в доказательство во многом основано на геометрических соображениях. Тогда и будет совершен переход от папируса Райнда к началу 20 века.

Глава 2. Аналитическая геометрия и визуализация[2] линейной алгебры, а также кривых и поверхностей второго порядка

Вопрос: "Что такое геометрия?" не не имеет однозначного ответа. В учебниках пишут: *"Геометрия — наука о свойствах геометрических фигур. «Геометрия» слово греческое, означает землемерие."* С моей точки зрения одно из важнейших назначений геометрии состоит в поверке математической гармонии не алгеброй, как у Сальери, а воображением: возможность при исследовании обратиться к мысленным образам — необыкновенный дар интеллекта. Геометрия располагается посередине между алгеброй и анализом и украшает их обоих. Кроме всего этого, геометрия дает бесценный материал для тренировки мышления.

§ 2.1. Одномерный случай. В школе на уроках алгебры, изучая функции, строят прямоугольную систему координат Oxy и рисуют графики функций $y = f(x)$, например, линейных функций $y = kx + b$. Решение уравнения $f(x) = 0$ называется *корнем* уравнения $y = f(x)$. Для нахождения корня уравнения $y = kx + b$, надо решить линейное уравнение с одним неизвестным $kx+b = 0$. Такие уравнения мы решали в § 1.1. В школе изучают также и квадратичную функцию $y = ax^2+2bx+c$ (в школе коэффициент при x обычно равен b, но нам удобен коэффициент именно $2b$). Рис. 1 представляет собой "визуализацию" следующего результата, который оставим без доказательства (кто его не знает ?):

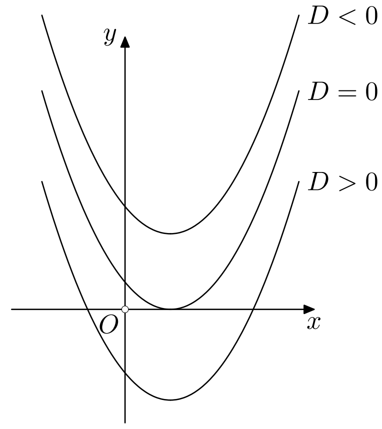


Рис. 1.

Теорема 2.1. *Совокупность корней уравнения $y = ax^2 + 2bx + c$, $a \neq 0$ это либо две точки прямой $\mathbb{R}$ $x_{1,2} = \frac{-b\pm\sqrt{b^2-ac}}{a}$ (если $D = b^2 - ac > 0$), либо одна $x = -b$ (если $b^2 = ac$), либо ни одной (если $D = b^2 - ac < 0$).*

Далее нам предстоит изучать и изображать нули линейной функции двух переменных, а также описывать и

[2]Визуализация — представление явления или процесса в форме удобной для зрительного восприятия.

изображать нули квадратичной функций двух переменных, затем сделать то же самое с совокупностью решений некоторого числа линейных уравнений и одной квадратичной функции n переменных (иллюстрироваться это будет на примере $n = 3$) и, наконец, описать один специальный тип квадратичных функций бесконечного числа переменных.

§ 2.2. Аналитическая планиметрия.

2.2.1.Двумерное евклидово векторное пространство.

В конце прошлой главы мы уже изучали векторное евклидово пространство l_2 бесконечного числа переменных. Но нельзя исключать, что к § 1.4 прошлой главы читатель-таки “с достоинством вышел из игры”. Здесь мы опять будем подбираться к l_2, начиная с двух, затем трех, затем любого числа переменных.

Лист бумаги, лежащий на плоском столе, представляет собой фрагмент евклидовой плоскости, которую изучают в школе. Если на нашем листе бумаги провести две перпендикулярные прямые и назвать их осями координат Ox_1 и Ox_2, выбрать на них одинаковый масштаб и точке X на листе бумаги поставить в соответствие вектор $\bar{x} = (x_1, x_2)$, где x_1 — это расстояние до нуля точки X_1, получающейся пересечением с осью Ox_1 прямой, проходящей через точку X параллельно оси Ox_2, а x_2 определяется аналогично (рис. 2), получится отображение точек листа бумаги (а в воображении — всей евклидовой плоскости) на совокупность пар (x_1, x_2), которые будем называть *векторами* (или тоже точками) и обозначать $\bar{x}$. Числа x_1 и x_2 называются (в честь французского ученого и философа Р. Декарта (1596–1650)) *декартовыми координатами* точки $\bar{x}$. Соответствие $X \leftrightarrow \bar{x} = (x_1, x_2)$ взаимно однозначно, и описание фигур на плоскости (нарисованных на листе бумаги), можно записать в виде каких-то уравнений или неравенств в от двух переменных (x_1, x_2).

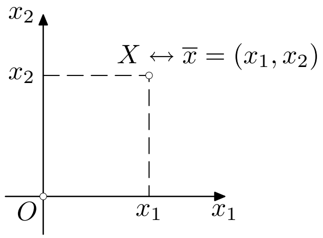


Рис. 2.

Евклидово двумерное векторное пространство $\mathbf{E}^2$ (напомним, что о векторных пространствах можно прочитать в § 2 Дополнения) образовано векторами $\bar{x} = (x_1, x_2)$, снабженными *скалярным произведением* $(\bar{x}, \bar{y}) = x_1y_1 + x_2y_2$ для $\bar{x} = (x_1, x_2)$, $\bar{y} = (y_1, y_2)$. Скалярное произведение характеризуется свойствами: 1) скалярное произведение вектора с самим собой неотрицательно и равно нулю, лишь если вектор равен нулю; 2) $(\alpha\bar{x}, y) = \alpha(\bar{x}, \bar{y})$; 3)$(\bar{x}' + \bar{x}'', \bar{y}) = (\bar{x}', \bar{y}) + (\bar{x}'', \bar{y})$; 4)$(\bar{x}, \bar{y}) = (\bar{y}, \bar{x})$. Число $|\bar{x}| = \sqrt{x_1^2 + x_2^2}$ называется *модулем вектора* $\bar{x} \in \mathbf{E}^2$. Пусть векторы $\bar{x} = (x_1, x_2)$ и $\bar{y} = (y_1, y_2)$ наклонены к оси Ox_1 под углами α и β.

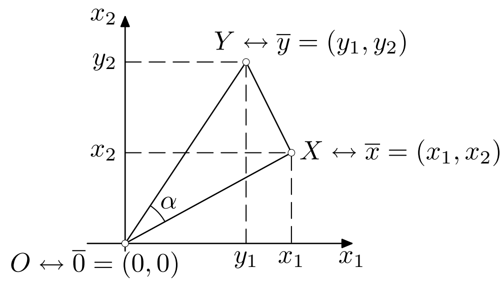
Это означает, что $x_1 = |\bar{x}| \cos\alpha$, $x_2 = |\bar{x}| \sin\alpha$, $y_1 = |\bar{y}| \cos\beta$, $y_2 = |\bar{y}| \sin\alpha$. Тогда

$$(\bar{x}, \bar{y}) = x_1y_1 + x_2y_2 = |\bar{x}| \cos\alpha |\bar{y}| \cos\beta + |\bar{x}| \sin\alpha |\bar{y}| \sin\beta = |\bar{x}||\bar{y}| \cos(\alpha - \beta). \tag{1}$$

Таким образом, *скалярное произведение двух векторов $\bar{x}$ и $\bar{y}$ равно произведению их модулей на косинус угла между ними* — такова геометрическая интерпретация скалярного произведения (рис. 3). Точка $(0, 0)$ называется *началом координат.*

Евклидовы двумерное пространство это модель — евклидовой плоскости, которая изучались нами в школе на уроках планиметрии В принципе любая геометрическая задача может быть переведена на язык алгебры, а многие алгебраические факты возможно изобразить или вообразить. Продемонстрируем это на трех примерах.

2.2.2. Определители. Если читатель подсчитает площадь треугольника OXY (или “по векторному” $\bar{o} = (0, 0), \bar{x} = (x_1, x_2), \bar{y} = (y_1, y_2)$), то он обнаружит, что эта площадь равна половине модуля определителя матрицы $\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$ (см. рис. 3). А сам определитель матрицы размеров два на два есть *ориентированная* площадь параллелограмма, натянутого на векторы-стоки матрицы (площадь надо брать со знаком +, если от первого вектора ко второму можно пройти против часовой стрелке, повернувшись на угол, меньший π, и − в противоположном случае.



$a)(\bar{x}, \bar{y}) = |\bar{x}||\bar{y}| \cos\alpha$

$b)S_{OXY} = \left|\det\left(\begin{smallmatrix} x_1 & x_2 \\ y_1 & y_2 \end{smallmatrix}\right)\right|$

Рис. 3.

2.2.3. Прямые.

В школьной геометрии прямые не определяются: это неопределяемый основной объект, а в аналитической геометрии прямая определяется: прямой в $\mathbf{E}^2$ называется множество векторов $\bar{x} \in \mathbf{E}^2$, удовлетворяющих одному линейному уравнению $(\bar{a}, \bar{x}) = b$, где $\bar{a} \in \mathbf{E}^2$, $\bar{a} \neq 0$, а b — некоторое число. При этом пары $(\bar{a}, b)$ и $(\lambda\bar{a}, \lambda b)$, $\lambda \neq 0$ определяют одну и ту же прямую. На рис. 4 изображена прямая $2x_1 + 3x_2 = 2$. Иначе говоря, прямая на евклидовой плоскости — это *совокупность решений одного линейного уравнения с двумя неизвестными.* Приведем это уравнение в развернутом виде:

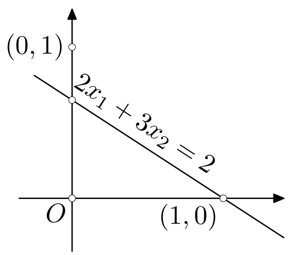


Рис. 4.

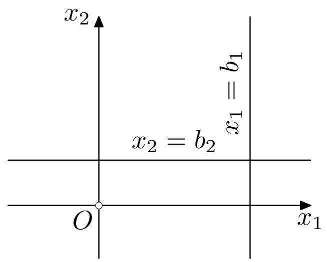


Рис. 5.

$$a_1x_1 + a_2x_2 = b \;\Leftrightarrow\; (\bar{a}, \bar{x}) = 0,\, \bar{a}, \bar{x} \in \mathbf{E}^2, \tag{2}$$

и назовем *общим уравнением* прямой на плоскости. Рассмотрим частные случаи общего уравнения (см. рисунки):

Неполные уравнения (см. рис. 5); *уравнения в отрезках* (рис. 6); *уравнение с угловым коэффициентом* (рис. 7); *уравнение в параметрической форме* (рис. 8); *каноническое уравнение* (рис. 9).

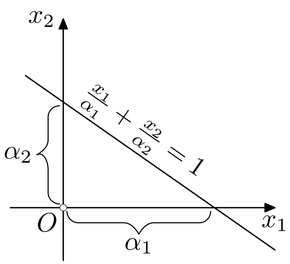


Рис. 6.

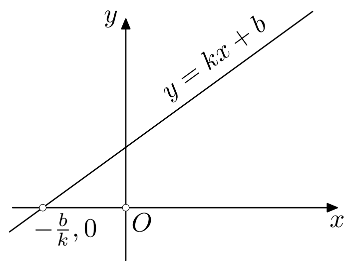


Рис. 7.

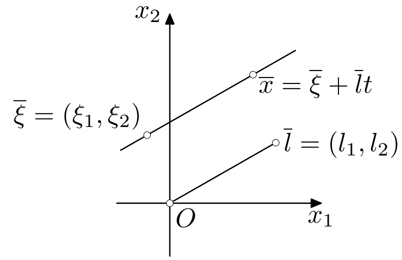


Рис. 8.

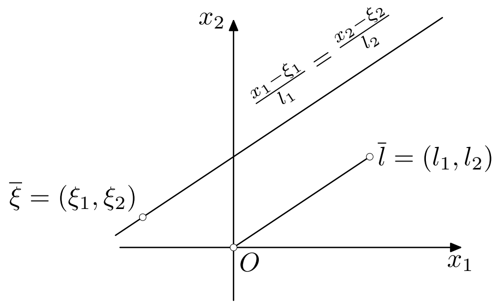


Рис. 9.

В геометрии Евклида есть две всем известные аксиомы: 1)через две различные точки плоскости проходит прямая и притом только одна, 2) через точку вне прямой можно провести параллельную и притом только одну. В аналитической геометрии плоскости это теоремы. Докажем их. Пусть $\bar{x}^1$ и $\bar{x}^2$ — две точки. Ищем пару $(\bar{a}, b)$, чтобы $(\bar{a}, \bar{x}^i) = b,\ i = 1, 2$. Вычитая одно уравнение из другого, приходим к одному однородному уравнению с двумя неизвестными a_1, a_2: $(\bar{a}, \bar{x}^2 - \bar{x}^1) = 0$. Но это мы научились решать в гл.1. Пусть теперь $(\bar{a}, \bar{x}) = b$ — прямая и $\bar{c}$ — точка, не принадлежащая этой прямой (т. е. $(\bar{a}, \bar{c}) \neq b$). Тогда прямая $(\bar{a}, \bar{x} - \bar{c}) = 0$ не пересекается с изначальной прямой и проходит через точку $\bar{c}$, что и требовалось. Единственность в обоих случаях докажите самостоятельно.

В качестве упражнений можете порешать такие задачи: 1) Прямые $(\bar{a}^1, \bar{x}) = b^1$ и $(\bar{a}^2, \bar{x}) = b^2$ параллельны (перпендикулярны) тогда и только тогда, когда $\bar{a}^1 = \lambda\bar{a}^2,\ \lambda \neq 0$ $((\bar{a}^1, \bar{a}^2) = 0)$; 2) Расстояние от точки $\bar{c}$ до прямой $\bar{a}, \bar{x} = b$ равно $|(\frac{\bar{a}}{|\bar{a}|}, \bar{c}) - \frac{b}{|a|}|$.

2.2.4. Коники и квадрики.

Коники появились в геометрии, как *плоские сечения (двуполого) конуса.* Так они определялись в трудах одного из трех — наряду с Евклидом и Архимедом — величайших ученых древности Аполлония (ок. 262–ок. 190 до н. э.)). Плоскость может пересекать все образующие конуса, такую кривую Аполлоний назвал *эллипсом*; она может быть параллельна одной образующей, такую кривую он назвал *параболой*; она может пересечь обе полы конуса, такую (двусвязную) кривую он назвал гиперболой (см. рис. 10). Но плоскость может проходить и через

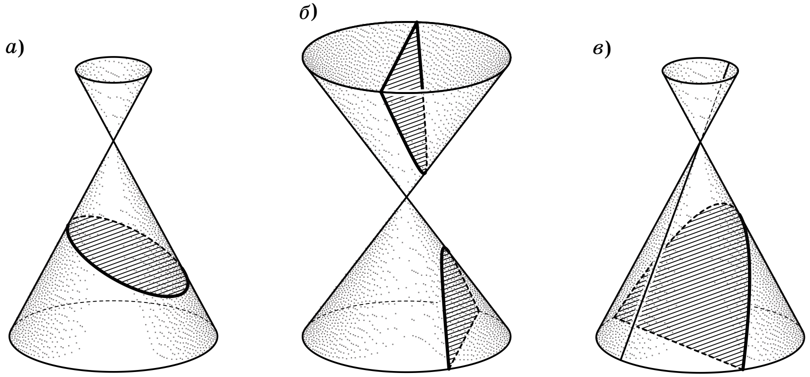


Рис. 10.

вершину конуса. при этом получаются две пересекающиеся прямые, одна прямая (если плоскость касается конуса) и, наконец, пересечением может быть лишь только сама вершина. Всего оказалось шесть возможностей. На рис. 11 приведено на самом деле геометрическое доказательство того, что эллипс — это множество точек плоскости, сумма расстояний от которых до двух данных точек (на рис 11 — от точек F_1 и F_2), в которых "верхняя" и "нижняя" сферы, вписанные в конус, касаются секущей плоскости. На рис. 12 доказывается, что гипербола — это множество точек плоскости, разность от которых до двух данных точек постоянна, наконец, на рис. 13 доказывается, что парабола — это множество точек плоскости, расстояние которых до некоторой точки и некоторой прямой постоянно. Переведем приведенные геометрические определения на язык алгебры.

Начнем с простейшего случая — параболы. Пусть на листе бумаги отмечена точка F и проведена прямая l. Проведем ось Ox_1 через F перпендикулярно l. Точку пересечения прямых l и Ox_1 обозначим

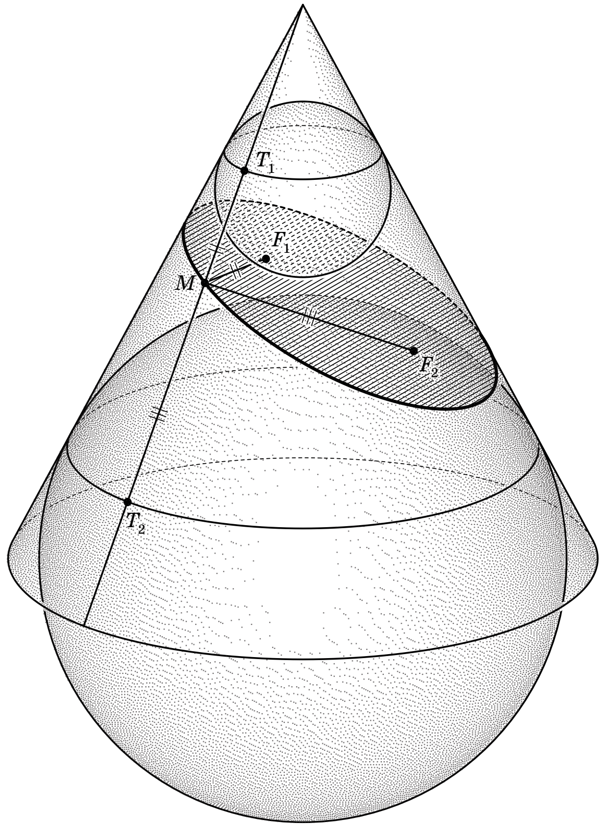


Рис. 11.

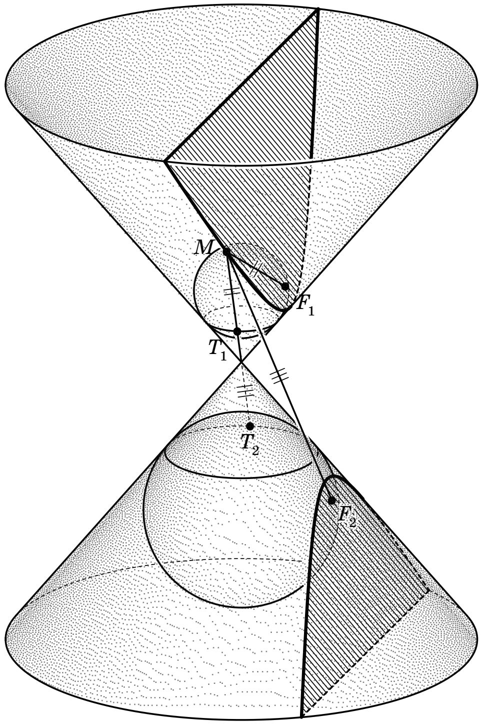


Рис. 12.

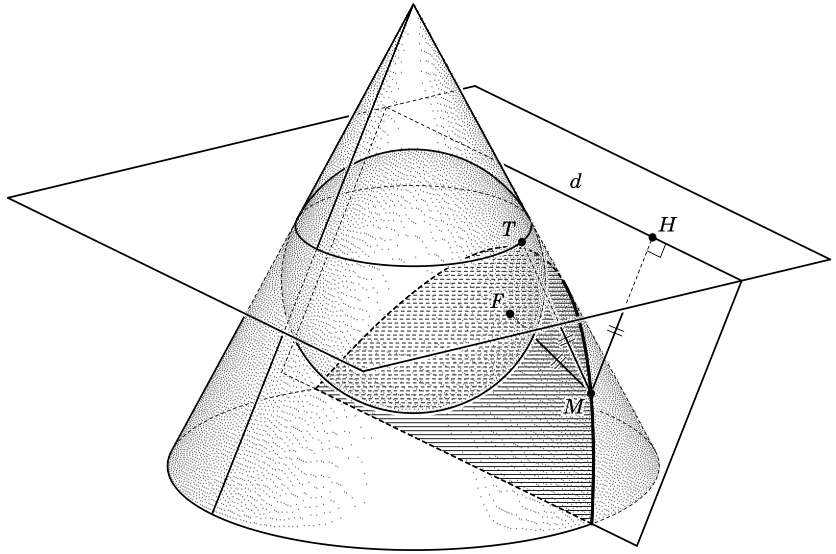


Рис. 13.

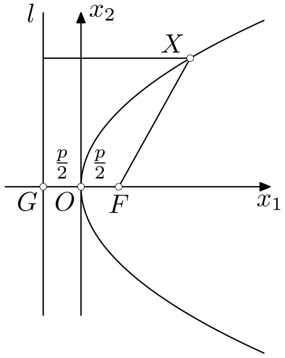


Рис. 14.

G (рис. 14). Тогда $|FG|$ (длина отрезка $[FG]$) — это расстояние от F до l. Пусть оно равно числу p. Середину отрезка $[FG]$ обозначим O и проведем через O прямую, параллельную l. В декартовой модели нашей плоскости (нашего листа бумаги) прямая l имеет уравнение $x_1 = -\frac{p}{2}$, а точка F имеет координаты $(\frac{p}{2}, 0)$. Пусть X — точка на параболе, т. е. $|XF| = d(X, l)$ (ее расстояния до F до l, согласно определению, совпадают). В декартовой модели это равенство запишется в виде: $x_1 + \frac{p}{2} = \sqrt{(x_1 - \frac{p}{2})^2 + x_2^2}$. Возводя в квадрат и приводя подобные, приходим к *каноническому уравнению параболы:* $x_2^2 = 2px_1$. (Это нулевая линия уровня квадратичной функции $F(x_1, x_2) = x_2^2 - 2px_1$).

В случае эллипса и гиперболы ось Ox_1 проводится через фокусы, начало координат выбирается посередине между фокусами, а ось Ox_2 проводится через точку O перпендикулярно оси Ox_1. Тогда фокусы будут иметь координаты $(\pm c, 0)$ и точка на эллипсе (гиперболе) характеризуется равенством $|XF_1| + |XF_2| = a$ ($|XF_1| - |XF_2| = a$), где a — некоторое число. В декартовой модели это запишется как $\sqrt{(x_1 + c)^2 + x_2^2} \pm \sqrt{(x_1 - c)^2 + x_2^2} = 2a$. Переведя один из радикалов в правую сторону, возводя в квадрат, приводя подобные и потом снова возводя в квадрат, приходим к каноническим уравнениям эллипса $((\frac{x_1}{a})^2 + (\frac{x_2}{b})^2 = 1,\ b = \sqrt{a^2 - c^2}$ (рис. 15) и гиперболы: $((\frac{x_1}{a})^2 + (\frac{x_2}{b})^2 = 1,\ b = \sqrt{c^2 - a^2}$ (рис. 16).

А теперь будем решать обратную задачу — пытаться визуализировать алгебраические уравнения, точнее — рисовать на бумаге или в воображении нулевые линии уровня квадратической функции. Для этого придется сначала выбирать удачную систему координат.

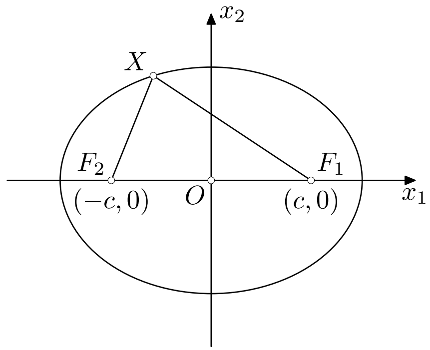


Рис. 15.

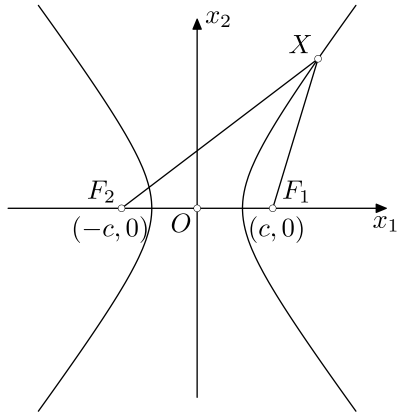


Рис. 16.

Рассмотрим многочлен второго порядка от двух переменных: $f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c$. Выразим его в матрично-векторной форме $f(\bar{x}) = (A\bar{x}, \bar{x}) + 2(\bar{b}, \bar{x}) + c$. В этом выражении симметричная матрица $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $a_{12} = a_{21}$ порождает отображение E^2 в E^2 по правилу $A\bar{x} = (a_{11}x_1 + a_{21}x_2, a_{12}x_1 + a_{22}x_2)$, а $\bar{b}$ — вектор (b_1, b_2). Наша задача теперь, расположить (на листе бумаги, где нарисована кривая $f(\bar{x}) = 0$) декартову систему координат так, чтобы уравнение кривой стало попроще.

Сначала попробуем сдвинуть систему координат параллельно на вектор a, перейдя к координатам $y_1 = x_1 + a_1, y_2 = x_2 + a_2$ или $\bar{x} = \bar{y} - \bar{a}$. Подставив в f вместо $\bar{x}$ вектор $\bar{y} - \bar{a}$ и обозначив $\varphi(y) = f(\bar{y} - \bar{a})$, получим $\varphi(y) = (A(\bar{y} - \bar{a}), (\bar{y} - \bar{a})) + 2(\bar{b}, (\bar{y} - \bar{a})) + c$. После несложных преобразований будем иметь: $\varphi(y) = (A\bar{y}, \bar{y}) + 2(\bar{b} - A\bar{a}, \bar{y}) + c'$, где $c' = c + (\bar{a} - 2\bar{b}, \bar{a})$. Если A невырожденная матрица (если, ее определитель не нуль), то можно решить уравнение $A\bar{a} = \bar{b}$, что приведет квадратичную форму к более простому виду $(A\bar{y}, \bar{y}) + c'$.

Отметим, что замену переменных можно трактовать, как переход к другой декартовой системе координат, в которой начало координат выбрано в точке $\bar{a}$, а оси координат проведены параллельно старым. Но есть еще возможность, не меняя масштаба повернуть ортогональную систему, не меняя масштаба. Так вот оказавается, что тогда возможно избавиться от удвоенных произведений и оставить одни квадраты. И действительно, совершим переход от координат $\bar{y} = (y_1, y_2)$ к координатам

$\bar{z} = (z_1, z_2)$ по формулам: $z_1 = y_1 \cos\alpha + y_2 \sin\alpha, z_2 = -y_1 \sin\alpha + y_2 \cos\alpha$ или, выражая игреки через зеты: $y_1 = z_1 \cos\alpha - z_2 \sin\alpha,\ y_2 = z_1 \sin\alpha + z_2 \cos\alpha$. Подставив в уравнение для коники зеты вместо игреков, приходим к выражению $c_{11}(\alpha)z_1^2 + 2z_1z_2((a_{22}-a_{11})\sin 2\alpha + a_{12}\cos 2\alpha) + c_{22}(\alpha)z_2^2$. Таким образом, если положить $\alpha = \frac{1}{2}\operatorname{arcctg}\frac{(a_{22}-a_{11})}{a_{12}}$, мы приведем форму к виду $\lambda_1 z_1^2 + \lambda_2 z_2^2 + c_1 = 0$, а это каноничесие уравнения эллипся, гиперболы и пары прямых. Возможны также лишь начало координат или пустое множество. А если через $\bar{f}_1$ и $\bar{f}_2$ обозначить образы векторов $\bar{e}_1 = (1,0)$ и $\bar{e}_2 = (1,0)$ при наших преобразованиях — сначала сдвиге, а потом повороте, то окажется, что наша квадратичная функция f приобрела такой вид: $f(\bar{x}) = \lambda_1(\bar{f}_1, \bar{x})^2 + \lambda_2(\bar{f}_2, \bar{x})^2$

Сформулируем доказанное в виде теоремы:

Теорема 2.2 (описание невырожденных квадрик). *Если* $\det A \neq 0$, *то найдутся два единичных взаимно ортогональных вектора* $\{\bar{f}_i\}_{i=1,2}$ *при которых функция* $f(\bar{x}) = (A\bar{x}, \bar{x}) + 2(\bar{b}, \bar{x})$ *приобретает вид:* $f(\bar{x}) = \sum_{i=1}^{2} \lambda_i(\bar{f}_i, \bar{x})$, При этом совокупность решений уравнения $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0$, есть либо пустое множество, либо точка, либо пара пересекающихся прямых, либо эллипс, либо гипербола.

§ 2.3. Аналитическая многомерная геометрия. Многомерный случай почти во всем повторяет двумерный, и потому будем кратки.

2.3.1.n-мерное евклидово векторное пространство. Евклидово n-мерное векторное пространство $\mathbf{E}^n$ образовано векторами $\bar{x} = (x_1, x_2, \ldots, x_n)$, снабженными *скалярным произведением* $(\bar{x}, \bar{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$. Число $|\bar{x}| = \sqrt{x_1^2 + \ldots + x_n^2}$ называется *модулем вектора* $\bar{x} \in \mathbf{E}^n$. Евклидовы трехмерное векторное пространство $\mathbf{E}^3$ — это модель евклидова трехмерного пространства, которые изучалось нами в школе на уроках стереометрии.

Замечание. Скалярное произведение позволяет идентифицировать пространство $\mathbb{R}^n$ и пространство линейных функционалов на нем, обозначенное нами в гл. 1 через $(\mathbb{R}^n)'$.

Двинемся дальше. Пусть $A = (a_{ij})_{1\le i,j\le n}$ — квадратная матрица порядка n. Ей соответствует отображение (оператор) из $\mathbf{E}^n$ в $\mathbf{E}^n$, задаваемое формулами $A\bar{x} = (\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{nj}x_j$. Это отображение обладает линейным свойством: $A(\alpha\bar{x} + \beta\bar{y}) = \alpha A\bar{x} + \beta A\bar{y}$. Теперь нам предстоит увидеть или вообразить прямые, плоскости и т. п.

2.3.2. Прямые и плоскости в $\mathbf{E}^3$, линейные многообразия в $\mathbf{E}^n$. В аналитической стереометрии плоскость определяется как множество уровня линейной функции — это множество векторов $\bar{x} \in \mathbf{E}^3$, удовлетворяющих одному линейному уравнению $(\bar{a}, \bar{x}) = b$, где $\bar{a} \in \mathbf{E}^3$, $\bar{a} \neq$

0, а b — некоторое число. При этом пары $(\bar{a}, b)$ и $(\lambda\bar{a}, \lambda b)$, $\lambda \neq 0$ определяют одну и ту же плоскость. Множество $\{\bar{x} \in \mathbf{E}^n \mid (\bar{a}, \bar{x}) = b,\ \bar{a} \in \mathbf{E}^n,\ b \in \mathbb{R}\}$ (множество уровня линейной функции в $\mathbf{E}^n$) называется *гиперплоскостью.*

Замечание. Отметим, что вектор $\bar{a}$ перпендикулярен прямой $(\bar{a}, \bar{x}) = b$, и это можно нарисовать. Вернемся на секунду к двум уравнениям с двумя неизвестными: $A\bar{x} = \bar{b}$ (ситуация чуть изменилась, мы "жили" в $\mathbb{R}^2$, а "переехали" в $\mathbf{E}^2$, где прямое и дуальное пространства как бы сливаются, и можно все на свете изображать). Так вот, если $A\mathbf{E}^2 \neq \mathbf{E}^2$, (и $A \neq 0$), то образ оператора A — прямая, так что необходимым и достаточным условием разрешимости, является ортогональность $\bar{b}$ перпендикулярной прямой к $A\mathbf{E}^2$, т. е. $(\bar{b}, \bar{y}) = 0$ для любого $\bar{y}$, такого что $(\bar{y}, A\bar{x}) = 0\ \forall \bar{x}$. Но это в точности то, что было выше сформулировано нами, если учесть, что сопряженное уравнение в $\mathbf{E}^2$ задается транспонированной матрицей.

Пересечение гиперплоскостей называется *линейным многообразием.* Пересечения двух несовпадающих плоскостей в $\mathbf{E}^3$ — это прямые в $\mathbf{E}^3$. Если $(\bar{a}, \bar{x}) = b$ — уравнение гиперплоскости, то, поделив на $|\bar{a}|$ и обозначив $\frac{\bar{a}}{|\bar{a}|} = \bar{e}$, а $\frac{b}{|a|} = p$, приходим к *нормальному уравнению* гиперплоскости: $(\bar{e}, \bar{x}) = p$. Такую гиперплоскость будем обозначать $\Gamma(\bar{e}, p)$. Пусть прямая содержит точки $\bar{x}_0$ и $\bar{x}_0 + \bar{l}$ (в этом случае $\bar{l}$ называют *направляющим вектором* прямой). Тогда прямую можно представить в *параметрическом виде*, как $\{\bar{x} = \bar{x}_0 + t\bar{l}\ |,\ t \in \mathbb{R}\}$. Будем обозначать ее $L(\bar{x}_0.\bar{l})$. Если гиперплоскость задавать нормальным уравнением, а прямую представлять в параметрическом виде, очень легко отвечать на разнообразные вопросы об их расположении: a) параллельность гиперплоскостей: $\Gamma(\bar{e}^1, p_1)||\Gamma(\bar{e}^2, p_2) \Leftrightarrow \bar{e}^1 = \bar{e}^2, p_1 \neq p_2$; b) параллельность гиперплоскости $\Gamma(\bar{e}, p)$ и прямой $L(\bar{x}_0.\bar{l}) : \Gamma(\bar{e}, p)||L(\bar{x}_0.\bar{l}) \Leftrightarrow \bar{x}_0 \notin \Gamma(\bar{e}, p), (\bar{l}, \bar{e}) = 0$; c) расстояние $d(\bar{x}_0, \Gamma(\bar{e}, p)$ от точки $\bar{x}_0$ до гиперплоскости $\Gamma(\bar{e}, p) : d(\bar{x}_0, \Gamma(\bar{e}, p)) = |(\bar{e}, \bar{x}_0) - p|$ и т. д. и т. п.

2.3.3. Определители и правило Крамера.

Системы линейных уравнений с помощью определителей не решают (такие решения требуют слишком большого числа операций), но само понятие определителя играет выдающуюся роль и в алгебре, и в геометрии, и в анализе.

Геометрический смысл определителя, порожденного n векторами в n-иерном пространстве — это *ориентированный объем параллелепипеда, порожденного этими векторами.* Отметим свойства такой характеристики, ясной на примерах $n = 2, 3$. Введем функцию $V : (\mathbb{R}^n)^n \to \mathbb{R}$, $V = V(\bar{a}^1, \ldots, \bar{a}^n)$ системы векторов $\{\bar{a}^j\}_{j=1}^n$. и потребуем от этой функции выполнения таких условий: a) условия нормировки, согласно которому объем единичного куба должен равняться единице: $V(\bar{e}^1,\ldots,\bar{e}^n)=1$, где $\bar{e}^1=\begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix},\ldots,\bar{e}^n=\begin{pmatrix}0\\ \vdots\\ 1\end{pmatrix}$ стандартный базис в $\mathbb{R}^n$,

b) условия "антисимметричности":
$V(\bar{e}^1,\ldots,\bar{e}^{i-1},\bar{e}^i,\bar{e}^{i+1},\bar{e}^{i+2},\ldots,\bar{e}^n)=-V(\bar{e}^1,\ldots,\bar{e}^{i-1},\bar{e}^{i+1},\bar{e}^i,\bar{e}^{i+2},\ldots,\bar{e}^n)$ и

c) условия линейности по каждому векторному аргументу (для этого достаточно потребовать, чтобы $V(\alpha\bar{a}^1+\alpha'\bar{a}^{1'},\bar{a}^2,\ldots,\bar{a}^n)=\alpha V(\bar{a}^1,\bar{a}^2,\ldots,\bar{a}^n)+\alpha' V(\bar{a}^{1'},\bar{a}^2,\ldots,\bar{a}^n)$).

Выведем из этих свойств выражение для определителя в двумерном случае. Положим $\bar{e}^1=\begin{pmatrix}1\\0\end{pmatrix}$, $\bar{e}^2=\begin{pmatrix}0\\1\end{pmatrix}$. Имеем: $V=V(\bar{a}^1,\bar{a}^2)=V(a_{11}\bar{e}^1+a_{21}\bar{e}^2,a_{12}\bar{e}^1+a_{22}\bar{e}^2)=(a_{11}a_{22}-a_{12}a_{21})V(e^1,e^2)=a_{11}a_{22}-a_{12}a_{21}$.

Аналогично из приведенных аксиом однозначно выводится выражение для объема параллелепипеда, порожденного векторами $\bar{a}^1,\bar{a}^2$ и $\bar{a}^3$ в трехмерном случае: $V(\bar{a}^1,\bar{a}^2,\bar{a}^3)=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{21}a_{32}a_{13}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{32}a_{23}$.

По индукции приходим к такому результату: *существует единственная функция $V=V(\bar{a}^1,\ldots,\bar{a}^n)$, которая определяется аксиомами a) - c).* Эта функция равна сумме $\sum_P(-1)^{\mathrm{sign}P}\Pi_{i=1}^n a_{iP(i)}$, взятой по всем перестановкам P первых n чисел $\{1,2,\ldots,n\}$ со знаком, равным $(-1)^{\mathrm{sign}P}$, где число $\mathrm{sign}P$ равно нулю, если переход от $\{1,2,\ldots,n\}$ к $\{P(1),P(2),\ldots,P(n)\}$ требует четного числа транспозиций соседних элементов, и единице, если требуется нечетное число транспозиций (скажем, член $a_{21}a_{32}a_{13}$ имеет знак $+$, ибо переход от (312) к (123) требует двух транспозиций:(312) $\to$ (132) $\to$ (323)). Выписанная функция удовлетворяет аксиомам a) - c). Она и называется *детерминантом или определителем* матрицы A и обозначается $\det A$. Для тренировки читатель, который сталкивается с этим впервые, может сосчитать какой-нибудь определитель четвертого порядка (придется просуммировать выражение из двадцати четырех слагаемых).

Из условия антисимметрии вытекает, что если у матрицы имеется два одинаковых столбца, то ее определитель равен нулю; а если соединить это утверждение с линейностью определителя по столбцу, то обнаружится, что если столбцы матрицы линейно зависимы, то определитель также равен нулю.

Правило Крамера. *Если определитель* $\mathrm{det}A$ *системы* (3) *отличен от нуля, тогда система однозначно разрешима для любой правой части, и решение представимо формулой Крамера:*

$$x_i=\frac{\mathrm{det}A_i}{\mathrm{det}A},\ \mathrm{det}A_i=V(\bar{a}^1,\ldots,\bar{a}^{i-1},\bar{b},\bar{a}^{i+1}\ldots,\bar{a}^n),\ 1\le i\le n. \qquad (i)$$

Доказательство. Пусть $\mathrm{det}A\neq 0$. Тогда по сказанному выше, его столбцы линейно независимы и по теореме 1.3 система (3) разрешима. Пусть $\bar{b}\in\mathbb{R}^n$. В силу разрешимости системы $A\bar{x}=\bar{b}$, найдутся x_i, $1\le i\le n$, такие, что $x_1\bar{a}^1+\ldots+x_n\bar{a}^n=\bar{b}$. Подставим вектор $x_1\bar{a}^1+\ldots+x_n\bar{a}^n$ вместо вектора $\bar{a}^i$. Определитель полученной матрицы в силу свойств определителя равен $x_i\mathrm{det}A$. А если мы подставим вместо $\bar{a}^i$ равный вектору $x_1\bar{a}^1+\ldots+x_n\bar{a}^n$ вектор $\bar{b}$, мы получим матрицу A_i. Значит, $x_i=\frac{\mathrm{det}A_i}{\mathrm{det}A}$. □

2.3.4. Классификация невырожденных квадрик в $\mathbf{E}^n$.

Пусть $A=\begin{pmatrix}a_{11}&\ldots&a_{1n}\\ \ldots&\ldots&\ldots\\ a_{n1}&\ldots&a_{22}\end{pmatrix}$, $a_{ij}=a_{ji}$ — симметричная матрица и $\bar{b}\in\mathbf{E}^n$.

Эта пара порождает квадратичную функцию $f(x)=(A\bar{x},\bar{x})+2(\bar{b},\bar{x})+c$ одномерный и двумерный вариант которой рассматривался выше. Будем рассматривать эту форму в другом "ортогональном базисе". Сделав параллельный переход, иначе говоря, замену переменных $\bar{y}=\bar{x}+\bar{a}$, преобразуем форму к виду: $\varphi(y)=(A\bar{y},\bar{y})+2(\bar{b}-A\bar{a},\bar{y})+c'$, где $c'=c+(\bar{a}-2\bar{b},\bar{a})$. Далее будем предполагать, что матрица A невырожденная. Это значит, что можно решить уравнение $A\bar{a}=\bar{b}$, что приведет квадратичную форму к более простому виду $(A\bar{y},\bar{y})+c'$. Оказывается возможным совершить переход от полученной сдвигом системы координат к новой декартовой системе, при которой удвоенные произведения исчезнут. А именно, имеет место

Теорема 2.3 (описание невырожденных квадрик). *Существует набор $\{\bar{f}_i\}_{i=1}^n$ единичных взаимно ортогональных векторов при которых форма $(A\bar{y},\bar{y})$ приобретает вид: $\sum_{i=1}^m\lambda_i(\bar{f}_i,\bar{y})^2$.*

В следующем пункте мы докажем этот результат одновременно с эскизом доказательства его бесконечномерного обобщения

§2.4. Эскизы доказательств альтернативы Фредгольма и приведения квадратичной формы к главным осям в конечномерном и бесконечномерном случае.

2.4.1. Эскиз доказательства альтернативы Фредгольма. Итак, пусть $X=l_2$, $A=I+B$, где B — компактный оператор из X в X. Вот, что это значит. В l_2 естественно вводится понятие модуля вектора

$\bar{x} : |x| = \sqrt{(\bar{x}, \bar{x})}$, а это в свою очередь, дает возможность определить метрику в l_2, положив $d(\bar{x}, \bar{y}) = |\bar{x} - \bar{y}|$ (о метрических пространствах см. §3 Дополнения). Множество $C \subset l_2$ называется *ограниченным*, если существует число $a > 0$, для которого $|x| < a \; \forall x \in C$. Оператор B называется компактным, если он ограниченное множество переводит в компактное (определение компактного множества см. в §3 Дополнения), если из любой последовательности в нем можно выбрать сходящуюся подпоследовательность). Доказывается, что компактный оператор является пределом вырожденных, но мы на этом не остановимся. Совокупность решений однородного уравнения $A\bar{x} = 0$ называется ядром оператора A, которое обозначается KerA. Нам надлежит доказать альтернативу: $AX = X$ или $\text{KerA} \neq \bar{0}$.

Доказательство проходит в значительной мере по схеме конечномерной теоремы 1.3. А) Допустим, что нашу систему $A\bar{x} = \bar{b}$ можно решить для любой правой части, а однородное уравнение имеет ненулевое решение и придем к противоречию. По предположению существует ненулевое решение $\bar{f}^1$ уравнения $A\bar{x} = 0, |\bar{f}^1| = 1$. По условию же можно решить уравнение $A\bar{x} = \bar{f}^1$. Обозначим его $\bar{e}^2$. В плоскости, натянутой на $\bar{f}_1$ и $\bar{e}_2$ найдем вектор единичный по модулю и ортогональный $\bar{f}_1$ $(|\bar{f}_2| = 1, (\bar{f}_2, \bar{f}_1) = 0$. Далее найдем решение $\bar{e}_3$ уравнения $A\bar{e}_3 = f_2$ и в трехмерном пространстве, порожденном векторами $\bar{f}_1, \bar{f}_2, \bar{e}_3$ найдем вектор f_3 такой, что $|\bar{f}_3| = 1, (\bar{f}_3, \bar{f}_1) = ((\bar{f}_3, \bar{f}_2) = 0$ и далее, поступая аналогично, построим бесконечную систему векторов $\{\bar{f}_n\}_{n\in\mathbb{N}}$. Пусть $m < n$, тогда в силу равенства $B\bar{f}_n - B\bar{f}_m = \bar{f}_n - A\bar{f}_n - f_m + Af_m$ видно, что векторы $\bar{f}_m, A\bar{f}_n, A\bar{f}_m$ ортогональны $\bar{f}_n$, и значит, $|Bf_n - Bf_m| \geq 1$ (наклонная не меньше катета). Пришли к противоречию: единичные векторы отображаются в постедовательность, из которой нельзя выбрать сходящуюся.

В) Пусть $AX \neq X$ Для доказательства, утверждения что тогда $\text{Ker}A \neq \bar{0}$, придется использовать три факта: что AX — замкнутое подпространство, что к замкнутому подпространству l_2 можно провести перпендикуляр и что оператор, сопряженный к компактному компактен (в конечномерном случае все это тривиальности, так что проводимое нами далее рассуждение дает еще одно доказательство в конечномерном случае). Итак, AX — замкнутое подпространство. Пусть $\bar{y}$ — перпендикуляр к нему, т. е. $(\bar{y}, A\bar{x}) = (A^T\bar{y}, \bar{x}) = 0 \; \forall \bar{x} \in X$. Это означает, что $A^T\bar{y} = \bar{0}$, т. е. $\text{Ker}A^T \neq \bar{0}$. По доказанному в А) отсюда следует (ввиду того, что A^T компактен), что $A^TX \neq X$. А по только что доказанному, отсюда следует, что $\text{Ker}A^{TT} \neq \bar{0}$. Но $A^{TT} = A$ (двукратная смена строк и столбцов все возвращает наместо). Вот мы и доказали, что хотели: $AX \neq X \Rightarrow \text{Ker}A \neq \bar{0}$. □

2.2.4. Теорема Гильберта и эскиз ее доказательства.

Теорема 2.4 (о приведении компактной квадратичной формы к главным осям) а) (конечномерный случай) *Пусть A — симметричная матрица, порождающая отображение из $\mathbf{E}^n$ в $\mathbf{E}^n$. Тогда существует набор $\{\bar{f}_i\}_{i=1}^n$ единичных взаимно ортогональных векторов, при которых форма $(A\bar{y}, \bar{y})$ приобретает вид: $\sum_{i=1}^m \lambda_i(\bar{f}_i, \bar{y}),\ m \leq n$.*

b) В бесконечномерном случае матрица предполагается компактной и тогда существует набор $\{\bar{f}_i\}_{i=1}^\infty$ единичных взаимно ортогональных векторов, при которых форма $(A\bar{y}, \bar{y})$ приобретает вид: $\sum_{i=1}^M \lambda_i(\bar{f}_i, \bar{y}),\ m \leq n$, где M — либо натуральное число, либо бесконечность; при этом $|\lambda_i|$ монотонно стремятся к нулю.

Эскиз доказательства. Доказательство а) основывается на правиле множителей Лагранжа. Проведем полное доказательство в двумерном случае, а потом сразу перейдем к бесконечномерному.

Рассмотрим задачу на максимум функции $g_0(\bar{x}) = (A\bar{x}, \bar{x})$ (предполагая, что квадратичная форма принимает положительные значения, иначе рассмотрим задачу на минимум) при ограничении $g_1(\bar{x}) = (\bar{x}, \bar{x}) = 1$. Здесь ищется максимум непрерывной функции g_0 на окружности единичного радиуса, которая является компактом в $\mathbf{E}^2$. Такая функция достигает своего максимума (см. §3 Дополнения). Обозначим вектор, на котором достигается этот максимум $\bar{f}_1$.

Согласно правилу множителей Лагранжа (которое обсуждается в следующей главе) найдется число λ_1 такое, что производная функции $g_0 + \lambda_1 g_1$ в точке $\bar{f}_1$ равна нулю. Это приводит к равенству $A\bar{f}_1 = \lambda_1\bar{f}_1$. (Такие векторы называются собственными). Пусть $\bar{f}_2$ — единичный вектор, ортогональный $\bar{f}_1$. Воспользовавшись симметрией матрицы A, мы поучим $(A\bar{f}_2, \bar{f}_1) = (\bar{f}_2, A\bar{f}_1) = \lambda_1(\bar{f}_2, \bar{f}_1) = 0$, т. е. $A\bar{f}_2$ тоже ортогонален $\bar{f}_1$, значит $\bar{f}_2$ и $A\bar{f}_2$ пропорциональны: $A\bar{f}_2 = \lambda_2\bar{f}_2$. Но тогда $\bar{x} = (\bar{x}, \bar{f}_1)f_1 + (\bar{x}, \bar{f}_2)f_2$. Подставляя это выражение вместо $\bar{x}$ в $(A\bar{x}, \bar{x})$ и пользуясь тем, что $\bar{f}_i$ ортогональны и единичны $(|f_1| = |f_2| = 1,\ (f_1, f_2) - 0$, приходим к требуемому равенству $(A\bar{x}, \bar{x}) = \lambda_1(\bar{x}, \bar{f}_1)^2 + \lambda_2(\bar{x}, \bar{f}_2)^2$.

Что же касается бесконечномерного случая, то тут надо рассмотреть задачу о максимизации $(\bar{x}, \bar{x})$ (как функции на l_2) при условии, что $(A\bar{x}, \bar{x}) \leq 1$. Из компактности A извлекается существование решения. Применение правила множителей Лагранжа (которое не будет доказано здесь, но мы остановимся совсем близко от него), позволит найти первый собственный вектор $\bar{f}_1$. Затем надо рассмотреть ту же задачу с дополнительным ограничением $(x, \bar{f}_1) = 0$. из нее извлекается второй собственный вектор, затем накладывается ограничение $(x, \bar{f}_2) = 0$, находится третий собственный вектор и т.п.

Глава 3. Математический анализ и дифференциальные уравнения

Эта глава посвящена математическому анализу и фрагменту теории дифференциальных уравнений. Математический анализ зародился в трудах И. Ньютона (1643–1727) и Г. Лейбница (1646–1716). “ Поворот в математике [вызванный зарождением математического анализа], — писал А. Н. Колмогоров, — произошел в 17 веке одновременно с созданием основ математического естествознания. Значение этого поворота настолько велико, что образовавшиеся в результате этого поворота разделы объединяют под названием *высшей* математики в отличие от сложившейся ранее *элементарной* математики”. Суть связи естествознания и математического анализа можно выразить так: **динамические процессы, происходящие в природе, описываются дифференциальными уравнениями.** Цель главы — **ознакомить с основными понятиями математического анализа, научить решать нелинейные (в частности, дифференциалные) уравнения, научить исследовать экстремальные задачи и простейшие задачи математической физики.** Основное место в этой главе уделяется *методу Ньютона решения нелинейных уравнений и исчислениям*[3] — дифференциальному и интегральному (которые изучаются совместно с дифференциальными уравнениями). Обоснованиям анализа в тексте также уделяется некоторое место. Исторические рамки этой главы — от семнадцатого до двадцатого века.

§ 3.1. Дифференциальное и интегральное исчисления, теорема о корне (одно переменное).

3.1.1. Производная и интеграл.

Основные понятия математического анализа очень естественны. В подтверждение этого приведем слова Ньютона, в которых он объясняет суть создаваемой им теории: “Для пояснения искусства анализа, — писал Ньютон, — нужно привести некоторые примеры задач. 1)Пусть длина пути известна. Нужно узнать скорость в данный момент времени. 2)Пусть известна скорость движения. Надо узнать длину пройденного пути.”

[3]Исчисления — это набор правил действия с вводимыми понятиями математического анализа, которые позволяют пользоваться в приложениях математическим языком также, как мы пользуемся нашим родным языком — не проверяя всякий раз свою речь правилами грамматики.

Пусть к моменту t был преодолен путь $s(t)$ (т. е. функция $t \mapsto s(t)$ считается известной). Продумывание первого вопроса естественно приводит тому, что скорость $v(\tau)$ в данный момент τ 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* (рис. 17). [4] Математики пишут при этом так: $v(\tau)=\lim_{\Delta t\to 0}\frac{s(\tau+\Delta t)-s(\tau)}{\Delta t}$ и говорят, что скорость — это *производная*

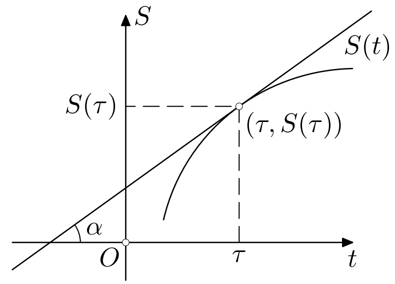


Рис. 17.

пути по времени и обозначают $v(\tau)=s'(\tau)$ или $v(\tau)=\frac{ds(\tau)}{dt}$. Из рис. 17 видно, что $s'(\tau)=\operatorname{tg}\alpha$, где α — угол, образованный касательной к графику функции $t\mapsto s(t)$, проведенной в точке $(\tau, s(\tau))$.

И совершенно также как осмысление первой задачи привело нас к понятию производной, попытка ответить на второй вопрос Ньютона — как узнать длину пройденного пути, если известна скорость движения, — естественно приводит к понятию определенного интеграла. Пусть нам известна скорость $v(t)$ в любой момент времени t на отрезке времени $[t_0,t_1]$, и мы хотели бы найти длину пути, преодоленного за этот период времени. Естественно тогда разбить отрезок $[t_0,t_1]$ на большое число n частей точками $t_0=\tau_0<\tau_1<\ldots<\tau_n=t_1$ и считать, что на участке $[\tau_{k-1},\tau_k]$ (длину которого $\tau_k-\tau_{k-1}$ обозначим Δ_k) скорость примерно постоянна и равна $v(\theta_k)$, где θ_k какая-то промежуточная точка на отрезке $[\tau_{k-1},\tau_k]$. Таким образом, преодоленный путь, равный $s(t_1)-s(t_0)$ оказывается примерно равным сумме $v(\theta_1)\Delta_1+\ldots+v(\theta_n)\Delta_n$. (Эта сумма называется *римановой суммой* (рис. 18); ее обозначают $\mathcal{R}(v(\cdot),D,S)$, где $D=t_0=\tau_0<\tau_1<\ldots<\tau_n=t_1$ — это разбиение

[4]Вот что это значит. Число $v=v(\tau)$ называется пределом функции $\Delta t\mapsto\frac{s(\tau+\Delta t)-s(\tau)}{\Delta t}$ при Δt стремящемся к нулю, если для любого $\varepsilon>0$ можно найти такое число $\delta>0$, что если $|\Delta t|<\delta$, то $|v(\tau)-\frac{s(\tau+\Delta t)-s(\tau)}{\Delta t}|<\varepsilon$.

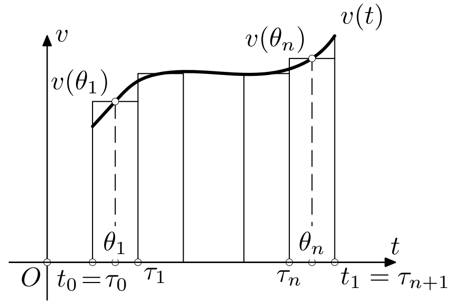


Рис. 18.

(D — от слова division — разбиение, S — от слова selection — выбор) — выбор точек θ_k на отрезке $[\tau_k, \tau_{k+1}]$. Эта сумма (если движение происходит только вперед и тем самым скорость неотрицательна) примерно равна *площади под графиком скорости*. Так вот, *определенный интеграл и есть эта самая площадь* или *предел римановых сумм*, когда максимальное из чисел Δ_k разбиения D (называемое *диаметром разбиения* D и обозначаемое $\mathrm{Diam}(D)$) стремится к нулю.[5] Эту площадь (или этот предел, если он существует) обозначают $\int_{t_0}^{t_1} v(t)dt$ и называют *(определенным) интегралом от функции* $v(\cdot)$ *от* t_0 *до* t_1.

Заодно мы пришли к знаменитой формуле Ньютона–Лейбница: $\int_{t_0}^{t_1} s'(t)dt = s(t_1) - s(t_0)$.

3.1.2. Функции, элементарные функции, пределы, непрерывность и дифференцируемость.

Далее рассматриваются вещественные функции f одного переменного (т. е. функции $f : X \to \mathbb{R}$, определенные на подмножествах X вещественной прямой $\mathbb{R}$), сопоставляющие числу x из X вещественное число $f(x)$). На первых порах мы ограничимся *элементарными функциями* . Для того, чтобы сделать это понятие определенным, введем *основные* элементарные функции. Таковы функции $f_1(x) = c$, где c — константа (постоянные функции), $f_2(x) = x^\alpha$ (степенные функции), $f_3(x) = a^x, a > 0$ (показательные функции), $f_4(x) = \log_a x, a > 0, a \neq 1$ (логарифмические функции), $f_5(x) = \sin x$, $f_6(x) = \cos x$, $f_7(x) = \operatorname{tg} x$, $f_8(x) = \operatorname{ctg} x$ (тригонометрические

[5] Вот что это означает: определенным интегралом функции f, определенной на отрезке $[a, b]$ называется число J такое, что для любого $\varepsilon > 0$ найдется $\delta > 0$, для которого при любом разбиении D отрезка $[a, b]$ диаметра меньшего δ и любом выборе S точек в элементах этого разбиения, выполняется неравенство $|R(f, D, S) - J| < \varepsilon$.

функции), $f_9(x) = \arcsin x$, $f_{10} = \arccos x$, $f_{11} = \operatorname{arctg} x$ $f_{12} = \operatorname{arcctg} x$ (обратные тригонометрические функции). Об этих функциях рассказывается в школе.

Элементарными функциями называют функции, явным образом получаемые из основных с помощью конечного числа арифметических операций и композиций. Таковы линейные, квадратичные функции, полиномы, рациональные функции (отношения полиномов, $x \to |x|(= \sqrt{x^2})$ и т. д. Позволяется доопределять функции в отдельных точках, скажем, единицей в нуле функцию $x \to \frac{\sin x}{x}$ или нулем в нуле функцию $x \to \frac{x}{|x|}$ (эту функцию обозначают $x \to \mathrm{sign} x$).

Функция $x \to \alpha(x)$, определенная на некотором интервале $(-c, c)$, $c > 0$ называется *бесконечно малой* , если для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0\,(\delta \leq c)$, что если $|x| < \delta$, то $|\alpha(x)| < \varepsilon$. В этом случае будем писать α б.м.

Функции, определенные на натуральном ряде ($(n \mapsto x_n, n \in \mathbb{N})$) или $\{x_n\}_{n\in\mathbb{N}}$ называются *последовательностями.* Последовательность $\{\alpha_n\}_{n\in\mathbb{N}}$ называется бесконечно малой (и в этом случае будем писать $\{\alpha_n\}$ б.м.), если для любого числа $\varepsilon > 0$ найдется такое число $N \in \mathbb{N}$, что если $n > N$, то $|\alpha_n| < \varepsilon$.

Говорят, что последовательность $\{x_n\}_{n\in\mathbb{N}}$ *имеет предел, равный* a *при* n, *стремящемся к* ∞, если последовательность $\alpha_n = x_n - a$ б.м.; при этом пишут $\lim_{n\to\infty} x_n = a$.

Пусть функция f определена на интервале $(x_0 - c, x_0 + c)$, $c > 0$ (за исключением, быть может, самой точки x_0). Говорят, что она *имеет предел, равный* a *при* x, *стремящемся к* x_0, если функция $\alpha(x) = f(x_0 + x) - a$ б.м.; при этом пишут $\lim_{x\to x_0} f(x) = a$; говорят, что эта функция *непрерывна в точке* x_0 (и пишут $f \in C(x_0)$) , если $\lim_{x\to x_0} f(x) = f(x_0)$. Если функция f определена на некотором множестве X и непрерывна в каждой точке множества X, говорят, что она *непрерывна на* X *и пишут* $f \in C(X)$.

Повторим определение производной для функций $y = f(x)$.

Определение производной функции одного переменного (по Коши; Огюстен Луи Коши (1789–1857) — ученый, внесший фундаментальный вклад в математический анализ и его обоснование) . Говорят, что функция f, определенная всюду на интервале $(x_0 - c, x_0 + c)$, $c > 0$, дифференцируема в точке x_0, если существует предел $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$, обозначаемый $f'(x_0)$ и называемый *производной* f *в точке* x_0. Линейная функция $dx \mapsto f'(x_0)dx$ называется дифференциалом f в точке x_0.

Если функция f дифференцируема в точке x_0, мы пишем $f \in D(x_0)$.

Очевидно, что если функция дифференцируема в точке, то она непрерывна в этой точке, но обратное неверно, как показывает функция $x \mapsto |x|$ в нуле.

Элементарные функции, как правило, непрерывны и дифференцируемы всюду в области их определения, мы вскоре в этом убедимся.

Функция F такая, что $F' = f$ называется *первообразной* f. Совокупность всех первообразных называется *неопределенным интегралом* f и обозначается $\int f\,dx$. Таким образом, мы решили простейшее дифференциальное уравнение: $y' = f(x)$. Его "общее решение" (совокупность всех решений) — это неопределенный интеграл $\int f\,dx$.

В одномерном случае существуют *дифференциальные формы* двух типов — функции $\omega_0 = f_0$ и выражения вида $\omega_1 = f_1(x)dx$ Формы ω_0 можно дифференцировать, и мы приходим к специальному виду формы ω_1: $df_0(x) = f_0'(x)dx$, вторые можно интегрировать: $\int_a^b f_1(x)dx$. Из формулы Ньютона–Лейбница получаем: $\int_a^b df_0(x) = f_0(b) - f_0(a)$. За этим простым фактом можно уже усмотреть замечательную формулу Пуанкаре $\int_\Omega d\omega = \int_{\partial\Omega} \omega$.

3.1.3. Исчисление бесконечно малых, пределов, непрерывных функций, производных и неопределенных интегралов.

Предложение 1. *a) 1)Произведение бесконечно малой на число, 2) и 3) сумма и произведение бесконечно малых — бесконечно малая; b) Пусть* $\lim_{x\to x_0} f_i(x) = a_i$, $i = 1,2$ *и* $a \in \mathbb{R}$. *Тогда 1)*$\lim_{x\to x_0}(af_1(x)) = aa_1$; *2)* $\lim_{x\to x_0}(f_1(x)+f_2(x)) = a_1+a_2$; 3) $\lim_{x\to x_0}(f_1(x)f_2(x)) = a_1a_2$; *4) если* $a_2 \neq 0$, *то* $\lim_{x\to x_0} \frac{f_1(x)}{f_2(x)} = \frac{a_1}{a_2}$. *c) 1)Произведение функции, непрерывной в точке на число, 2) и 3) сумма и произведение функций, непрерывных в этой точке и 4) если значение знаменателя в этой точке не равно нулю, то и частное функций, непрерывных функций в этой точке непрерывны в этой точке.*5) Суперпозиция непрерывных функций непрерывна.

Лемма (о разности первообразных). *Если функция f непрерывна, то производная разности двух ее первообразных равне нулю* (и значит, неопределенный интеграл $\int f(x)dx$, — это первообразная плюс произвольная константа).

Теорема Лейбница о производной сложной функции. *Если* $g \in D(x_0)$, *а* $f \in D(g(x_0))$, *то* $f(g) \in D(x_0)$ *и при этом* $(f(g))'(x_0) = f'(g(x_0))g'(x_0)$.

Следствие о производной обратной функции. *Если f и g взаимно обратны,* $g \in D(x_0)$, $f \in D(g(x_0))$, *то* $g'(x_0) = \frac{1}{f'(g(x_0))}$.

Основные соотношения (ОС) исчисления производных. a) $f \in D(x_0), a \in \mathbb{R} \Rightarrow (af) \in D(x_0)$ и $(af)'(x_0) = af'(x_0)$; b) $f_i \in D(x_0)$, $i = 1,2 \Rightarrow (f_1 + f_2) \in D(x_0)$ и $(f_1 + f_2)'(x_0) = f_1'(x_0) + f_2'(x_0)$; c) $f_i \in D(x_0)$, $i = 1,2 \Rightarrow (f_1f_2) \in D(x_0)$ и $(f_1f_2)'(x_0) = f_1'(x_0)f_2(x_0) + f_1(x_0)f_2'(x_0)$; d) $f_i \in D(x_0)$, $i = 1,2$, $f_2'(x_0) \neq 0 \Rightarrow \frac{f_1}{f_2} \in D(x_0)$ и $(\frac{f_1}{f_2})'(x_0) = \frac{f_1'(x_0)f_2(x_0)-f_1(x_0)f_2'(x_0)}{f_2^2(x_0)}$.

Все факты и равенства доказываются очень несложно, фактически из определений. Продемонстрируем пару доказательств ($\Rightarrow$ значит "следовательно", $\forall$ — "для всякого", $\exists$ — "найдется").

Предложение 1 а2). α_i, $i = 1,2$ б.м., следовательно, по определению (далее: $\overset{\text{def}}{\Rightarrow}$) $\forall\varepsilon > 0\ \exists\delta_i$ такие, что (далее :) $|x| < \delta_i \Rightarrow |\alpha_i| < \frac{\varepsilon}{2} \Rightarrow$ если $|x| < \min(\delta_1, \delta_2)$, то $\alpha_1(x) + \alpha_2(x) < \varepsilon$, что и требовалось (далее $\square$).

Предложение 1 в3). Дано: $\lim_{x\to x_0} f_i(x) = a_i$, $i = 1,2 \overset{\text{def}}{\Rightarrow} f_i(x) = a_i + \alpha_i$, α_i б.м. $\Rightarrow f_1(x)f_2(x) = (a_1 + \alpha_1(x))(a_2 + \alpha_2(x)) \overset{1a1),a2)}{=} a_1a_2+$б.м. $\square$

Лемма о разности первообразных следует из интегрируемости непрерывной функции и формулы Ньютона–Лейбница, интегрируемость непрерывной функции вытекает из теоремы Кантора (см. §3 Дополнения). Теорема Кантора влечет за собой также формулу Ньютона–Лейбница (все это хорошие упражнения).

Основные производные (ОП) элементарных функций. a) $\sin' x = \cos x$; b) $(e^x)' = e^x$. Эти формулы выводятся в любом курсе анализа.

Из теоремы 1, ее следствия, и формул ОС и ОП вытекают формулы исчисления производных и неопределенных интегралов.

1. $y = e^x \overset{\text{ОП}}{\Rightarrow} y' = e^x; y' = e^x \Rightarrow y = e^x + C$.

2. $y = \sin x \overset{\text{ОП}}{\Rightarrow} y' = \cos x; y' = \cos x \Rightarrow y = \sin x + C$.

3. $y = \cos x \overset{\text{def}}{=} \sin(x + \frac{\pi}{2}) \overset{\text{ОП,Thm1}}{\Rightarrow} y' = \sin(x + \frac{\pi}{2}) = -\sin x; y' = \sin x \Rightarrow y = -\cos x + C$.

4.$y = \operatorname{tg} x \overset{\text{ОСd)}}{\Rightarrow} y' = \frac{1}{\cos^2 x}; y' = \frac{1}{\cos^2 x} \Rightarrow y = \operatorname{tg} x + C$.

5. $y = \operatorname{ctg} x \overset{\text{ОСd)}}{\Rightarrow} y' = -\frac{1}{\sin^2 x}; y' = \frac{1}{\sin^2 x} \Rightarrow y = -\operatorname{ctg} x + C$

Подобные формулы можно достаточно долго выписывать. Ввиду недостатка времени не буду продолжать (будучи уверенным, что почти все мои слушатели все это так или иначе проходили). Ограничусь лишь двумя замечаниями: 1) формула для интегрирования $dg(t) = g'(t)dt$ дает возможность применять при интегрировании *замену переменных* ($\int f(x)dx \overset{x=g(t)}{=} \int f(g(t))g'(t)dt$); 2) формула $d(uv) = du v + u\,dv$ дает возможность *интегрировать по частям*: $\int u\,dv = uv - \int v\,du$.

Сделаю еще несколько замечаний о рядах. Пусть дана последовательность $\{a_n\}_{n=1}^{\infty}$. Выражение $a_1 + a_2 + \ldots + a_n + \ldots$ называют *рядом*. Говорят, что этот ряд *сходится (абсолютно сходится)*, если последовательность его частных сумм $a_1,\ a_1 + a_2, \ldots, a_1 + \ldots + a_n, \ldots$ сходится (последовательность сумм $|a_1|,\ |a_1| + |a_2|, \ldots, |a_1| + \ldots + |a_n|, \ldots$ сходится). Два абсолютно сходящихся ряда можно перемножать. Полнота $\mathbb{R}$ немедленно приводит к следующему критерию сходимости Коши: *ряд $a_1 + a_2 + \ldots + a_n + \ldots$ сходится тогда и только тогда, когда для любого $\varepsilon > 0$ найдется такое $N \in \mathbb{N}$, что $|a_n + a_{n+1} + \ldots + a_{n+m}| < \varepsilon\ \forall n \geq N$ и $m \geq 0$.*

Существует множество *признаков сходимости*, из которых укажем два: 1) Если $\theta = \lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$, то ряд сходится (признак Даламбера); если ряд знакопеременный ($a_n a_{n+1} < 0\, \forall n$) и $|a_n|$ монотонно стремятся к нулю, то ряд сходится (признак Лейбница). Первый верен из за того, что начиная с некоторого номера выполняется неравенство $|a_n| \leq C\theta^n$, во втором сегменты $[s_{2n}, s_{2n-1}]$, где $s_n = \sum_{k=1}^{n} a_k$ вложены друг в друга, а длины их стремятся к нулю.

3.1.4. Упрощенный метод Ньютона нахождения корня нелинейной функции одного переменного.

В гл. 1 мы учились решать линейные уравнения сначала одного, потом двух, потом многих и, наконец, бесконечного числа переменных. Здесь я собирался придерживаться той же схемы и идти от одного нелинейного уравнения с одним переменным, к одному уравнению и системе двух уравнений с двумя переменными, затем к системе m уравнений с n переменными и, наконец, к бесконечномерному случаю, но проделаю только часть пути. Параллельно будем входить в круг математического анализа сначала функций одного, потом двух, потом конечного числа переменных.

В начале § 2.1 был продемонстрирован на графике корень линейного уравнения с одним неизвестным $y = kx + b$. А теперь пусть задана функция $y = f(x)$, изображенная на рис. 19. Идея, как находить корни уравнения $f(x) = 0$, восходит к И. Ньютону. В письме секретарю Королевского общества Ольденбургу от 24.10. 1676 г., на примере решения уравнения $f(x) = x^3 - 2x - 5 = 0$, Ньютон изложил метод его решения, который в современных обозначениях представляет собой такую итеративную процедуру $x_n = x_{n-1} - (f'(x_{n-1}))^{-1} f(x_{n-1}),\ n \in \mathbb{Z}$.

Применив упрощенный метод, в котором производные не участвуют, докажем следующий результат:

Теорема 3.1 (о корне нелинейного уравнения в одномерном случае). *Пусть $B_{\mathbb{R}}(x_0, \delta) = \{x \in \mathbb{R} \mid |x - x_0| \leq \delta\}$ — отрезок числовой прямой с центром в x_0 ширины 2δ, F — функция на $B(x_0, \delta)$, и при*

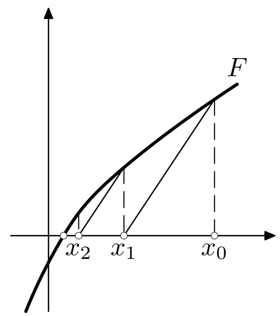


Рис. 19.

этом существует число $\Lambda \neq 0$ такое, что

$$|F(\xi) - F(x) - \Lambda(\xi - x)| \leq \theta |\Lambda| |\xi - x|, \tag{a}$$

$0 < \theta < 1, \forall \xi, x \in B(x_0, \delta)$. Тогда, если $|F(x_0)| \leq \delta(1 - \theta)|\Lambda|$, последовательность x_n, определенная соотношением

$$x_n = x_{n-1} - \Lambda^{-1} F(x_{n-1}),\ n \in \mathbb{Z} \tag{1}$$

сходится к числу $\widehat{x}$, являющемуся нулем функции F и при этом $|\widehat{x} - x_0| \leq K|F(x_0)|$, где $K = \frac{|\Lambda^{-1}|}{(1-\theta)}$.

В доказательстве используется полнота вещественной прямой (о полноте прямой можно прочитать в §1 Дополнения). Суть доказательства очень проста; она изображена на рис. 19. Проведем формальное доказательство.

Доказательство. Рассмотрим последовательность (1). Очевидно, что $x_0 \in B_{\mathbb{R}}(x_0, \delta)$. Допустим, что $x_k \in B_{\mathbb{R}}(x_0, \delta),\ 0 \leq k \leq N$. Соотношение (1) можно переписать для $n = N$ как $F(x_{N-1}) + \Lambda(x_N - x_{N-1}) = 0$, откуда:

$$|x_{N+1} - x_N| \overset{(1)}{=} |\Lambda^{-1}||F(x_N)| \overset{\mathrm{Id}}{=} |\Lambda^{-1}||F(x_N)) - F(x_{N-1}) - \Lambda(x_N - x_{N-1})| \overset{(a)}{\leq}$$

$$\theta |x_N - x_{N-1}| \leq \ldots \leq \theta^N |x_1 - x_0|. \tag{i}$$

Отсюда получаем

$$|x_{N+1} - x_0| \leq \sum_{k=0}^{N} |x_{k+1} - x_k| \overset{(i)}{\leq} (\sum_{k=0}^{N} \theta^k)|x_1 - x_0| \overset{(1)}{\leq} \frac{|\Lambda^{-1}||F(x_0)|}{(1-\theta)} \leq \delta, \tag{ii}$$

т. е. $x_{N+1} \in B_{\mathbb{R}}(x_0, \delta)$ и значит, $x_n \in B_{\mathbb{R}}(x_0, \delta)$ для всех $n \in \mathbb{N}$. Тогда, проведя аналогичную выкладку, получим, что $|x_{n+k} - x_n| \le \delta\theta^n\ \forall k, n \in \mathbb{N}$, т. е. последовательность $\{x_n\}_{n\in\mathbb{N}}$ фундаментальна и, следовательно, сходится к числу $\widehat{x}$. Переход к пределу в (1) показывает, что $\widehat{x}$ есть нуль F, а переход к пределу в (ii) приводит к неравенству $|\widehat{x} - x_0| \le K|F(x_0)|$, $K = \frac{|\Lambda^{-1}|}{(1-\theta)}$. □

§ 3.2. Дифференциальное и интегральное исчисления, теорема о корне (два переменных).

3.2.1. Производные функций и отображений; дифференциальные формы, их дифференцирование и интегрирование. Формула Грина.

Определение производной в одномерном случае, согласно которому функция f, определенная всюду на интервале $(x_0 - c, x_0 + c)$, $c > 0$, называется дифференцируемой в точке x_0, если существует такое число a (называемое производной f в точке x_0 и обозначаемое $f'(x_0)$), что имеет место равенство: $f(x_0+x) = f(x_0)+ax+\alpha(x)x$, где α б.м., равносильно определению по Коши, приведенному выше. Оно, в отличие от определения Коши, допускает многомерное (и даже бесконечномерное) обобщение. Сделаем первый шаг — от одного к двум. В одномерном случае производной было число, в двумерном — вектор из $(\mathbb{R}^2)'$ (линейный функционал на $\mathbb{R}^2$).

Пусть f — функция двух переменных, определенная на множестве $\{\bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid |x_1 - x_{10}| < c, |x_2 - x_{20}| < c\}$. называется дифференцируемой в точке $\bar{x}_0$, если существует такой вектор $\bar{a} \in (\mathbb{R}^2)'$, что имеет место равенство: $f(\bar{x}_0 + \bar{x}) = f(\bar{x}_0) + \bar{a}\cdot\bar{x} + r(\bar{x})$, где $r(\bar{x}) = \alpha(|\bar{x}|)|\bar{x}|$, $|\bar{x}| = \sqrt{x_1^2 + x_2^2}$, а α б.м.) Вектор $\bar{a}$ называется *производной* f *в точке* $\bar{x}_0$; если функция f дифференцируема в точке $\bar{x}_0$, мы пишем $f \in D(\bar{x}_0)$.

Рассмотрим две функции одного переменного, связанные с f. Обозначим через $\bar{e}_1$ вектор$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, а через $\bar{e}_2$ вектор $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ и положим $\varphi_1(t) = f(\bar{x}_0 + t\bar{e}_1)$ и $\varphi_2(t) = f(\bar{x}_0 + t\bar{e}_2)$. Из дифференцируемости f в точке x_0 следует дифференцируемость φ_1 и φ_2 в нуле. Производные этих функций в нуле обозначают $\frac{\partial f(\bar{x}_0)}{\partial x_1}$ и $\frac{\partial f(\bar{x}_0)}{\partial x_2}$ соответственно и называют частными производными по x_1 и x_2. Простые примеры (скажем, функция равная нулю всюду, кроме, как на фрагменте параболы $x_2 = x_1^2,\ x_1 > 0$) показывает, что наличие частных производных в точке не гарантирует дифференцируемость функции в этой точке (если $f \in D(\bar{x}_0)$, она непрерывна в этой точке, а в приведенном примере этого нет).

Две функции f_1 и f_2 двух переменных задают отображение $F = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ из $\mathbb{R}^2$ в $\mathbb{R}^2$. Пусть функции f_1 и f_2, задающие это отображение, определены на множестве $\{\bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid |x_1 - x_{10}| < c, |x_2 - x_{20}| < c\}$. Отображение F называется дифференцируемым в точке $\bar{x}_0$, если существует такая матрица $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, что имеет место равенство: $F(\bar{x}_0+\bar{x}) = F(\bar{x}_0)+A\bar{x}+\bar{\alpha}(|x|)|\bar{x}|$, где $\bar{\alpha}$ — отображение с б. м. компонентами. Матрица A называется производной F в точке $\bar{x}_0$, обозначаемой

$$F'(x_0) = \begin{pmatrix} \frac{\partial f_1(\bar{x}_0)}{\partial x_1} & \frac{\partial f_1(\bar{x}_0)}{\partial x_2} \\ \frac{\partial f_2(\bar{x}_0)}{\partial x_1} & \frac{\partial f_2(\bar{x}_0)}{\partial x_2} \end{pmatrix}.$$

Понятие определенного интеграла $\int_\Omega f(\bar{x})dx_1dx_2$ имеет для положительной функции сходный геометрический смысл с определенным интегралом от функции одного переменного: это *объем тела под функцией* f *над* Ω. Дадим точное определение для одной из простейших плоских фигур — параллелограмма, понимаемого, как часть плоскости. Пусть Ω параллелограмм, на котором определена функция f, D — некоторое разбиение Ω на параллелограммы $\Omega_i,\ 1 \le i \le N$, имеющие площади Δ_i и S — некоторая выборка по точке $\bar{\theta}_i \in \Omega_i$. Сумма $f(\theta_1)\Delta_1 + \ldots + f(\theta_N)\Delta_N$ называется *римановой суммой*; ее обозначают $\mathcal{R}(f(\cdot), D, S)$ Эта сумма для $f \ge 0$ примерно равна *объему под графиком* f. Определенный интеграл и есть этот самый объем или *предел римановых сумм*, когда максимальный из диаметров Ω_k (называемый *диаметром разбиения* D и обозначаемый $\mathrm{Diam}(D)$) стремится к нулю. Это означает, что число J, обозначаемое $\int_\Omega f(\bar{x})dx_1dx_2$ (и называемое определенным интегралом от f по Ω) таково, что для любого $\varepsilon > 0$ найдется $\delta > 0$, для которого при любом разбиении D множества Ω диаметра меньшего δ и любом выборе S точек в элементах этого разбиения, выполняется неравенство $|R(f, D, S) - J| < \varepsilon$. Переходим к дифференциальным формам.

В двумерном случае существуют дифференциальные формы трех типов — функции $\omega_0 = f_0(\bar{x})$, а также выражения $\omega_1 = f_{11}(\bar{x})dx_1 + f_{12}(\bar{x})dx_2$ и $\omega_2 = f_2(\bar{x})dx_1 \wedge dx_2$. Формы ω_0 и ω_1 можно дифференцировать, и мы приходим к специальному виду форм ω_1 и ω_2: $df_0(\bar{x}) = \frac{\partial f(\bar{x})}{\partial x_1}dx_1 + \frac{\partial f(\bar{x})}{\partial x_2}dx_2$ и $d(f_{11}(\bar{x})dx_1 + f_{12}(\bar{x})dx_2) = df_{11}(\bar{x}) \wedge dx_1 + df_{12}(\bar{x}) \wedge dx_2 = (\frac{\partial f_{12}(\bar{x})}{\partial x_1} - \frac{\partial f_{11}(\bar{x})}{\partial x_2})dx_1 \wedge dx_2$. (В исчислении форм пользуются соотношениями: $dx_i \wedge dx_i = 0$ и $dx_i \wedge dx_j = -dx_j \wedge dx_i,\ i \ne j$).

Формы ω_1 и ω_2 можно интегрировать. Пусть $\bar{x}(t) = (x_1(t), x_2(t)),\ a \leq t \leq b$ — плоская кривая, которую обозначим Γ. Тогда $\int_\Gamma \omega_1 = \int_\Gamma f_{11}(\bar{x})dx_1 + f_{12}(\bar{x})dx_2 = \int_a^b (f_{11}(\bar{x}(t))x_1'(t) + f_{12}(\bar{x}(t))x_2'(t))dt$. Пусть Ω — плоское подмножество, для простоты — параллелограмм Ω. Тогда $\int_\Omega \omega_2 = \int_\Omega f_2(\bar{x})dx_1dx_2$. Границу Ω обозначим $\partial\Omega$. Пусть оси координат направлены параллельно сторонам параллелограмма. Из формулы Ньютона–Лейбница сразу выводится формула: $\int_\Omega d\omega_1 = \int_{\partial\Omega} \omega_1$, которая при расшифровке превращается в формулу Грина (известную еще Эйлеру):

$\int_{\partial\Omega} f_{11}(\bar{x})dx_1 + f_{12}(\bar{x})dx_2 = \int_\Omega (\frac{\partial f_{12}(\bar{x})}{\partial x_1} - \frac{\partial f_{11}(\bar{x})}{\partial x_2})dx_1 \wedge dx_2$.

3.2.2. Упрощенный метод Ньютона нахождения корня нелинейного отображения из $\mathbb{R}^2$ в $\mathbb{R}^2$.

Пусть задано отображение F из $\mathbb{R}^2$ в $\mathbb{R}^2$. Корнем такого отображения называется вектор $\bar{x}$, для которого $F(\bar{x}) = \bar{0}$. Если $\Lambda = (\lambda_{ij})_{1\leq i,j\leq 2}$ — матрица размеров два на два, то в качестве нормы $\|\Lambda\|$ этой матрицы возьмем число $\lambda_{11}^2 + \lambda_{12}^2 + \lambda_{21}^2 + \lambda_2^2$.

Теорема 3.2 (о корне нелинейного уравнения в двумерном случае). *Пусть $B_\mathbb{R}(\bar{x}_0, \delta) = \{\bar{x} \in \mathbb{R}^2 \mid |\bar{x} - \bar{x}_0| \leq \delta\}$ — круг с центром в $\bar{x}_0$ радиуса δ, $F = (f_1, f_2)$ — отображение из $B(\bar{x}_0, \delta)$ в $\mathbb{R}^2$, и при этом существует невырожденная матрица Λ такая, что*

$$|F(\bar{\xi}) - F(\bar{x}) - \Lambda(\bar{\xi} - \bar{x})| \leq \theta\|\Lambda\||\bar{\xi} - \bar{x}|, \qquad (b)$$

$0 < \theta < 1, \forall \bar{\xi}, \bar{x} \in B(\bar{x}_0, \delta)$. *Тогда, если* $|F(\bar{x}_0)| \leq \delta(1 - \theta)\|\Lambda\|$,*последовательность $\bar{x}_n$, определенная соотношением*

$$\bar{x}_n = \bar{x}_{n-1} - \Lambda^{-1}F(\bar{x}_{n-1}),\ n \in \mathbb{Z} \qquad (2)$$

сходится к числу $\widehat{x}$, являющемуся корнем отображения F и при этом $|\widehat{x} - x_0| \leq K|f(x_0)|$, где $K = \frac{\|\Lambda^{-1}\|}{(1-\theta)}$.

Доказательство этой теоремы фактически ничем не отличается от доказательства теоремы 3.1, убедиться в чем предоставляется читателю, тем более, что в следующем пункте будет приведено доказательство бесконечномерного варианта теоремы о корне.

Извлечем из теоремы 3. 2 одно важное следствие — о правиле множителей Лагранжа.Скажем, что функция f, определенная на $B_\mathbb{R}(\bar{x}_0, \delta)$, *строго дифференцируема в точке $\bar{x}_0$*, если она дифференцируема в $\bar{x}_0$ и для любого $\varepsilon > 0$ найдется такое число $\delta > 0$, что из $|\bar{\xi} - \bar{x}_0| < \delta$, $|\bar{x} - \bar{x}_0| < \delta$ последует $|f(\bar{\xi}) - f(\bar{x}) - f'(\bar{x}_0)(\bar{\xi} - \bar{x})| \leq \varepsilon|\bar{\xi} - \bar{x}|$.

Правило множителей Лагранжа). *Пусть в задаче на максимум или минимум функции $\bar{x} \mapsto f_0(\bar{x})$ при ограничении $f_1(\bar{x}) = 0$ функции f_0 и f_1 определены на $B_\mathbb{R}(\bar{x}_0, \delta)$, строго дифференцируемы в точке $\widehat{x}_0$ и при этом $f_1'(\bar{x}_0) \neq 0$. Тогда, если в $\bar{x}_0$ достигается локальный максимум или минимум в задаче то найдется такое число λ, что выполнено равенство $f_0'(\bar{x}_0) + \lambda f_1'(\bar{x}_0) = \bar{0}$.*

Действительно, если $f_0'(\bar{x}_0) = \bar{0}$, то надо положить $\lambda = 0$. Если $f_0'(\bar{x}_0) \neq \bar{0}$, то либо производные пропорциональны, и тогда все доказано, либо матрица из производных невырождена и тогда по теореме 3.2 при малом α можно найти $\bar{x}_\alpha$, для которого $f_0(\bar{x}_\alpha) = \alpha$, а $f_1(\bar{x}_\alpha) = 0$, откуда следует, что $\bar{x}_0$ не является ни локальным максимумом, ни локальным минимумом. Противоречие.

3.2.3. Дифференциальные уравнения.

Пусть $f = f(t, x) : \mathbb{R}^2 \to \mathbb{R}$ $(t, x \in \mathbb{R})$ — функция двух переменных. Далее рассматриваются дифференциальные уравнения вида $\dot{x} = f(t, x)$. Задача: *решить дифференциальное уравнение $\dot{x} = f(t, x)$ с начальным условием $x(t_0) = x_0$* называется *задачей Коши*.

Начнем с теоремы существования решения задачи Коши. Для этого докажем один бесконечномерный вариант применения метода Ньютона.

Теорема 3.2$'$ (о корне $F(x(\cdot)) = 0)$[6]. *Пусть $X = Y = C([t_0, t_1])$ — пространство непрерывных функций на $[t_0, t_1]$, где $\|x(\cdot)\|_C = \max_{t\in[t_0,t_1]} |x(t)|$, $B_X(x_0(\cdot), \delta) = \{x(\cdot) \in X \mid \|x - x_0\|_C \leq \delta\}$ — шар в X с центром в $x_0(\cdot)$ радиуса δ, F — отображение из $B_X(x_0(\cdot), \delta)$ в Y, и при этом существует число θ, $0 < \theta < 1$ такое, что*

$$\|F(\xi(\cdot)) - F(x(\cdot)) - (\xi(\cdot) - x(\cdot))\|_C \leq \frac{\theta}{\gamma}\|\xi(\cdot) - x(\cdot)\|_C,\ \ \forall \xi, x \in B_X(x_0(\cdot), \delta). \qquad (2)$$

Тогда, если $\|F(x_0(\cdot))\|_Y \leq \frac{\delta(1-\theta)}{\gamma}$,последовательность $x_n(\cdot)$, определенная соотношением

$$x_n(\cdot) = x_{n-1}(\cdot) - (F(x_{n-1}(\cdot)),\ n \in \mathbb{N} \qquad (1)$$

сходится к элементу $\widehat{x}(\cdot)$, являющемуся нулем F и при этом $\|\widehat{x}(\cdot) - x_0(\cdot)\|_C \leq K\|F(x_0(\cdot))\|_C, K = \gamma/(1 - \theta)$.

Доказательство. Очевидно, что $x_0(\cdot) \in B_X(x_0(\cdot), \delta)$. Допустим, что удалось построить по методу (1) элементы $x_k(\cdot) \in B_X(x_0(\cdot), \delta)$, $0 \leq k \leq N$. Соотношение (1) можно переписать для $n = N$ как $F(x_{N-1}(\cdot) +$

[6]где $x(\cdot)$ — функция.

$(x_N(\cdot) - x_{N-1}(\cdot)) = 0$, откуда:

$$\|x_{N+1}(\cdot) - x_N(\cdot)\|_X \overset{(1)}{=} \|(F(x_N)(\cdot))\|_X \overset{\mathrm{Id},(1),}{\leq}$$

$$\gamma\|F(x_N)(\cdot)) - F(x_{N-1}(\cdot)) - (x_N(\cdot) - x_{N-1}(\cdot))\|_Y \overset{(2)}{\leq}$$

$$\theta\|x_N(\cdot) - x_{N-1}(\cdot)\|_X \leq \ldots \leq \theta^N\|x_1(\cdot) - x_0(\cdot)\|_X. \qquad (i)$$

Отсюда получаем

$$\|x_{N+1}(\cdot) - x_0(\cdot)\|_X \leq \sum_{k=0}^{N}\|x_{k+1}(\cdot) - x_k(\cdot)\|_X \overset{(i)}{\leq} (\sum_{k=0}^{N}\theta^k)\|x_1(\cdot) - x_0(\cdot)\|_X \overset{(1)}{\leq}$$

$$\frac{\gamma\|F(x_0(\cdot))\|_Y}{(1-\theta)} \leq \delta, \qquad (ii)$$

т. е. $x_{N+1}(\cdot) \in B_X(x_0(\cdot), \delta)$ и значит, $x_n(\cdot) \in B_X(x_0(\cdot), \delta)$ для всех $n \in \mathbb{N}$. Тогда, проведя аналогичную выкладку, получим, что $\|x_{n+k}(\cdot) - x_n(\cdot)\|_X \leq \delta\theta^n$, т. е. последовательность $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ фундаментальна и, следовательно, сходится к функции $\widehat{x}(\cdot)$. Переход к пределу в (1) показывает, что $\widehat{x}(\cdot)$ есть нуль F, а переход к пределу в (ii) приводит к неравенству $\|\widehat{x}(\cdot) - x_0(\cdot)\|_X \leq K\|F(x_0(\cdot))\|_Y, K = \gamma/(1-\theta)$. □

Теорема $3.2'$ хорошо приспособлена к доказательству теорем существования в теории дифференциальных уравнений.

Пусть $D = [t_0 - a, t_0 + a] \times B(x_0, b) \to \mathbb{R},\ f : D \to \mathbb{R}$. Рассмотрим задачу Коши:

$$\dot{x} = f(t, x),\ \ x(t_0) = x_0. \qquad (*)$$

Локальная теорема существования решения задачи Коши. *Если f непрерывна в D и липшицева с константой L по x (т. е. $f(t,\xi) - f(t,x)| \leq L|\xi - x|$), то на отрезке $[t_0 - \delta_0, t_0 + \delta_0]$ ($\delta_0 = \min(a, 1/2L, b/2M)$, где $M = \max_{t\in[t_0-\delta_0,t_0+\delta_0]}|f(t,x_0|)$ существует и единственно решение $(*)$.*

Для доказательства следует применить теорему $3.2'$, в которой $X = Y = C([t_0 - \delta_0, t_1 + \delta_0])$, $x_0(\cdot) \equiv x_0$, $F(x(\cdot))(t) = x(t) - x_0 - \int_{t_0}^t f(s, x(s))ds$, $\theta = 1/2$, $\delta = b$.

Непрерывность отображения F очевидна. Далее

$$\|F(x_0(\cdot))(\cdot)\|_{C([t_0-\delta_0,t_0+\delta_0],\mathbb{R}^n)} \leq \delta M \leq \frac{b}{2},$$

и поскольку для любых $\xi(\cdot),\ x(\cdot)$ выполняется неравенство

$$\|F(\xi(\cdot))(\cdot) - F(x(\cdot))(\cdot) - (\xi(\cdot) - x(\cdot))\|_{C([t_0-\delta_0,t_0+\delta_0],\mathbb{R}^n)} =$$

$$\max_{t\in[t_0-\delta,t_0+\delta]}|\int_{t_0}^{t}(f(s,\xi(s)) - f(s,x(s))ds| \leq$$

$$\delta_0 L\|\xi(\cdot) - x(\cdot)\|_{C([t_0-\delta_0,t_0+\delta_0],\mathbb{R}^n)} \leq \frac{1}{2}\|\xi(\cdot) - x(\cdot)\|_{C([t_0-\delta_0,t_0+\delta_0],\mathbb{R}^n)},$$

то справедливо b). Применение теоремы 1 немедленно приводит к локальной теореме существования. □

Несколько примеров. Простейшее неоднородное дифференциальное уравнение $\dot{x} = v,\ x(t_0) = x_0$ обсуждалось в предыдущем разделе. Его решение дается интегралом $x(t) = x_0 + \int_{t_0}^t v(s)ds$.

Рассмотрим простейшее *однородное* уравнения: $\dot{x} = x$. Для решения задачи Коши $\dot{x} = x,\ x(0) = 1$ этого уравнения, рассмотрим отображение $F : C([-a, a]) \to C([-a, a]),\ F(x(\cdot))(t) = 1 + \int_0^t x(s)ds$.

Если начать итеративную процедуру с функции $x_0(t) = 1$, мы получим $x_1(t) = 1 + t, x_2(t) = 1 + t + \frac{t^2}{2}, \ldots x_n(t) = \sum_{k=0}^n \frac{t^k}{k!}$. Эти суммы равномерно сходятся к функции $t \to e^t$ на некотором отрезке. На самом деле сходимость будет на всей прямой. Из теоремы об умножении рядов получим, что функция $t \to e^t$ — показательная.

Решение задачи Коши $\ddot{x} + x = 0,\ x(0) = 0, \dot{x}(0) = 1$ для уравнения гармонического осциллятора (выражающего второй закон Ньютона для частицы, притягивающейся к началу согласно закону Гука) сводится к нахождению неподвижной точки оператора $F(x(\cdot))(t) = t - \int_0^t(t - s)x(s)ds$. Стартуя от $x_0(t) = t$, приходим к ряду для синуса (убедитесь в этом самостоятельно)

Остановимся здесь, чтобы подвести некоторые итоги. Как связана доказанная нами теорема с окружающим нас миром?

Ньютон научил нас тому, что эволюционные процессы и явления описываются (обыкновенными) дифференциальными уравнениями. Второй закон Ньютона описывает движения взаимодействующих частиц. Теорема существования и единственности решения задачи Коши выражает идею лапласовского детерминизма: если бы Некто мог в одно мгновение узнать все скорости и положения частиц Мира, он знал бы о том, что было в прошлом и узнал бы, все, что случится в будущем. Это вольный пересказ того, о чем не раз говорил Лаплас. Два столетия идея предопределенности была одной из философских доминант (совсем недавно, каких-то полстолетия тому назад, все дороги вели к коммунизму). И лишь недавно в период *современной математики* обнаружилось, что чаще все процессы в итоге запутываются и скорее напоминают хаотическое движение. Но это уже другая история.

Далее предполагался небольшой набросок теории дифференциальных уравнений с разложением в ряды Маклорена экспоненты и синуса и формклой Тейлора, затем многомерный параграф с понятиями производной функции, отображения, дифференциальной формы, интеграла, интеграла от дифференциальной формы, с формулами Остроградского, Стокса и Пуанкаре, с теоремой 3.3 о корне отображений и правиле множителей Лагранжа для ограничений в виде нескольких равенств и, наконец,— обобщение теоремы 3.2', общее правило множителей Лагранжа для бесконечномерного случая и, как следствие, — доказательство теоремы Гильберта. Но это (вместе со всем предыдущим) заведомо невозможно изложить на двух лекциях, хотя все сюжеты развиваются по уже проложенному пути, стандартно и естественно, без привлечения каких бы то ни было новых идей.

Дополнение: числа, векторные пространства, общая топология

§1. Числа. Совокупность натуральных чисел $\{1, 2, 3\ldots\}$ обозначается $\mathbb{N}$, множество вещественных чисел — через $\mathbb{R}$. Вещественные числа можно определять конструктивно, как совокупность всевозможных десятичных дробей, не оканчивающихся на одни девятки, и аксиоматически, как *полное упорядоченное поле.* Десятичная дробь x — это последовательность чисел вида: $m_1 m_2 \cdots m_N . n_1 n_2 \cdots$, где m_i, $1 \le i \le N$ и n_i, $i \in \mathbb{N}$ — целые числа от нуля до девяти. Не допускается, чтобы все n_i, начиная с некоторого номера, были бы все девятки; дробь называется конечной, если, начиная с некоторого номера, все n_i нули; $m_1 m_2 \cdots m_N$ называется целой частью дроби, остальная часть называется дробной. Естественно определяется сравнение дробей (когда $x = y$ и $x < y$); с помощью действий над конечными дробями и понятия предела определяются сложение, вычитание, умножение и деление десятичных дробей, что превращает их совокупность (т. е. $\mathbb{R}$) в поле (т. е. множество, в котором определены операции сложения, умножения и деления с привычными аксиомами). Вещественные числа обладают свойствами а) полноты; согласно ему фундаментальная последовательность $\{x_n\}_{n\in\mathbb{N}}$ (т. е. такая последовательность $\{x_n\}_{n\in\mathbb{N}}$, что $\forall\varepsilon > 0 \exists N \in \mathbb{N} : |x_{n+k} - x_n| < \varepsilon \ \forall n \ge N, k \in \mathbb{N}$) имеет предел; и существования верхней (нижней) граней для ограниченного сверху (снизу) числового множества (для верхней грани это означает, что если $a < C$ для любого числа a из числового множества A, то $\exists\alpha \in \mathbb{R} : \alpha \ge a \ \forall a \in A \,\&\, \forall\varepsilon > 0 \, \exists a(\varepsilon) : a(\varepsilon) > \alpha - \varepsilon$. Таким образом, описанная нами совокупность десятичных дробей — это модель аксиоматически заданного объекта — полного упорядоченного поля.

§2. Векторные пространства. Векторное пространство (над $\mathbb{R}$) — это совокупность X элементов $\bar{x}$, в которой определены операции сложения элементов и умножения элемента на вещественно число, являющаяся абелевой группой по сложению и удовлетворяющих таким аксиомам умножения на числа $1 \cdot \bar{x} = \bar{x}$, $\alpha(\beta\bar{x}) = (\alpha\beta)\bar{x}$, $(\alpha + \beta)\bar{x} = \alpha(\bar{x} + \bar{x}'0 = \alpha\bar{x} + \beta\bar{x}'$.

§3. Общая топология: метрические пространства, компактность, непрерывность. Общая топология — ветвь математики, изучающая понятия предела и непрерывности. Опишем эти понятия в рамках метрических пространств. Пара (X, d), где X — множество, а $d = d(x, y)$ — вещественная неотрицательная функция, удовлетворяющая аксиомам: a) $d(x_1, x_2) = 0 \Leftrightarrow x_1 = x_2$, b) $d(x_1, x_2) = d(x_2, x_1) \ \forall x_1, x_2 \in X$, c) $d(x_1, x_3) \le d(x_1, x_2) + d(x_2, x_3) \, \forall x_1, x_2, x_3 \in X$, называется *метрическим пространством.* Последовательность $\{x_n\}_{n\in\mathbb{N}}$ называется *схо-*

дящейся, если существует такой элемент $\xi \in X$, что $\forall \varepsilon > 0\, \exists N \in \mathbb{N} : d(x_n, \xi) < \varepsilon\ \forall n \geq N$; последовательность $\{x_n\}_{n\in\mathbb{N}}$ называется *фундаметнальной* , если $\forall \varepsilon > 0\, \exists N \in \mathbb{N} : d(x_n, x_m) < \varepsilon\ \forall n, m \geq N$; вещественная функция f на (X, d) называется непрерывной в точке $x_0 \in X$, если $\forall \varepsilon > 0\, \exists \delta > 0 : d(x, x_0) < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$; функцию f называют непрерывной на (X, d), если она непрерывна в любой точке X метрическое пространство; (X, d) называют компактом, если из каждой последовательности элементов из X можно выбрать сходящуюся подпоследовательность.

Лемма (Гейне) *Для того, чтобы функция f, определенная на метрическом пространства (X, d) была непрерывной в точке x_0 необходимо и достаточно, чтобы она была непрерывной по Гейне, т. е. для любой последовательности $\{x_n\}_{n\in\mathbb{N}}$, сходящейся к x_0, последовательность $\{f(x_n)_{n\in\mathbb{N}}\}$ сходилась бы к $f(x_0)$.*

Доказательство. Из определения непрерывности по определению следует непрерывность по Гейне. Пусть непрерывность по Гейне имеется, но непрерывности нет. Отсутствие непрерывности означает, что имеется такое число $\varepsilon > 0$, что $\forall n \in \mathbb{N}\, \exists x_n : d(x_n, x_0) < \frac{1}{n}$, но $|f(x_n) - f(x_0)| > \varepsilon$. Это противоречит непрерывности по Гейне. □

Теорема 1. (Вейерштрасса о непрерывных функциях на компакте) *Функция, непрерывная на компакте, достигает на нем свое минимальное и максимальное значение.*

Доказательство. Пусть (X, d) — метрический компакт и $f \in C(X)$. a) Если допустить, что f неограничена, то $\forall n \in \mathbb{N}\, \exists x_n : |f(x_n)| > n$. По определению компактности $\exists \{n_k\}_{k\in\mathbb{N}} : x_{n_k} \to x_0$, что противоречит непрерывности по Гейне. b) Пусть f ограничена сверху и M — верхняя грань значений f. По определению верхней грани $\forall\, n \in \mathbb{N}\, \exists x_n : f(x_n) > M - \frac{1}{n}$. По определению компактности $\exists \{n_k\}_{k\in\mathbb{N}} : x_{n_k} \to x_0$. По лемме Гейне $f(x_0) = M$. □

Вещественная функция f на (X, d) называется равномерно непрерывной на X, если $\forall \varepsilon > 0\, \exists \delta > 0 : d(x_1, x_2) < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$.

Теорема 2 (Кантора). *Функция, непрерывная на компакте, равномерно непрерывна.*

Доказательство. Если допустить обратное, то найдется такое число $\varepsilon > 0$, что $\forall n \in \mathbb{N}\, \exists x_n, x - n' : d(x_n, x'_n) < \frac{1}{n}$, но $|f(x_n) - f(x'_n)| > \varepsilon$. По определению компактности $\exists \{n_k\}_{k\in\mathbb{N}} : x_{n_k} \to x_0$, при этом $f(x'_n) \notin f(x_0)$, что противоречит лемме Гейне. □

Введем метрику в совокупность непрерывных функций на компакте X: $d_C(f_1, f_2) = \max_{x\in X} |f_1(x) - f_2(x)|$. Пространство непрерывных функций на компакте X с метрикой d_C обозначают $C(X)$.

Теорема 3 (о полноте $C(X)$). *Пространство $(C(X), d_C)$ полно.*

Доказательство. Пусть $\{f_n\}_{n\in\mathbb{N}}$ — фундаментальная (в смысле метрики d_C) последовательность, т. е. $\forall \varepsilon > 0\, \exists N \in \mathbb{N} : d_C(f_n, f_m) < \varepsilon\ (i)\ \forall n, m \geq N$. Отсюда следует, что для каждого $x \in X$ $\lim_{n\to\infty} f_n(x) = f(x)$. Переходя к пределу в (i), получим, что при $n \geq N$ $|f(x) - f_n(x)| \leq \varepsilon\ (ii)$. Пусть $x_0 \in X$. Для некоторого $n_0 \geq N$ найдем $\delta > 0 : d_C(x, x_0) < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon/3\ (iii)$. Из (ii) и (iii) следует, что $|f(x) - f(x_0)| < \varepsilon$, т.е. f непрерывна. □