

ОСНОВЫ ФУНДАМЕНТАЛЬНОЙ МАТЕМАТИКИ

September 6, 2022

Contents

I	Введение	2
II	Освоение математического языка	4
1	Множества	4
2	Структуры на множествах	5
2.1	Алгебраические операции	6
2.2	Дискретная и локально дискретная математика	7
2.3	Метрика и топология	10
3	Вещественные и комплексные числа	12
III	Начала анализа	15
4	Производные и интегралы	15
4.1	Шкалы поведения функций	15
4.2	Вычисление производных	17
4.3	Суммы, ряды, интегралы	19
4.4	Длина кривой	21
4.5	Кратные интегралы и меры	22
5	Степенные ряды	24
5.1	Основные функции	25
5.2	Аналитические функции	29
6	Асимптотики сумм, произведений и интегралов	32
IV	Линейная алгебра	36
6.1	Линейные операторы и матрицы	36
6.1.1	Комбинаторика матриц	40

6.1.2	Спектр и подобие	41
6.2	Примеры локальной и бесконечномерной линейности	44
6.2.1	Диффеоморфизмы	44
6.3	Преобразование Фурье и обобщенные функции	46
V	Динамические системы	52
7	Детерминированные	52
7.1	На конечном множестве	52
7.2	Линейные дифференциальные уравнения	53
7.2.1	Решение	53
7.2.2	Качественные стороны решения	56
7.3	Нелинейность	60
7.3.1	Существование и единственность	60
7.3.2	Теория возмущений	60
8	Случайные	61
8.1	Классическая вероятность на конечных множествах	61
8.2	Конечные цепи Маркова	62
8.3	Случайные блуждания	64
8.4	Скейлинги и стохастические дифференциальные уравнения	66
8.4.1	От дискретного времени к непрерывному	66
8.4.2	От дискретного пространства к континууму	67
9	Квантовые	70
9.1	На конечном множестве	70
9.2	Квантовые частицы	73
9.3	Квантовые поля	73

Part I

Введение

Математика необъятна, можно развивать ее в глубину, ширину, а также в других, пока неизвестных, направлениях. В основном новые задачи возникают как естественные (к сожалению, часто формальные) обобщения уже известных задач. Но можно, например, искать истоки новых глобальных задач в других науках.

Дело еще в том, что сейчас математика все более расплывается по мелким темам, порождая конкурирующие между собой журналы и кланы. Можно конечно, раскапывая скважину в поисках клада в своем тайном месте, одновременно заглядывать в ямы других копателей. Но как правило, это бывает мало интересно и редко полезно.

И что делать, если кому-то трудно избавиться от желания увидеть картину одним взглядом. Но широта видения невозможна без краткости, без отказа от мелких деталей - такова видимо структура нашего мышления. Особенно важен широкий взгляд, если есть желание влезть со своей математикой в другие науки, чтобы лучше

их понять. Важнейшей движущей силой здесь может быть желание понять математическую структуру других наук, включая физику.

Такая же ситуация и с нашим проектом - есть две возможности работать над ним: заниматься решением конкретных задач (а их очень много), получая все более глубокие результаты, или взглянуть на проект шире и предлагать новые задачи и подходы.

Теперь собственно о содержании текста. В сравнении с толстыми учебниками и монографиями, этот текст выглядит как набросок. В действительности, он делает попытку соединить краткость с доступностью. Я надеюсь что он самодостаточен и может облегчить ряду читателей возможность научиться, не заглядывая в другие книги, видеть большие структуры как целое. Главной трудностью было следующее. Любой математический текст состоит из последовательностей логических скачков (переходов). Обычно серьезный читатель выбирает книгу с уровнем гладкости таких скачков, который требует от него минимальных усилий. В нашем изложении логические переходы менее гладкие и потребуют от читателя больше времени (в зависимости от своего математического уровня) для заполнения их более мелкими переходами. И для студентов также важно постараться заполнять эти переходы самому, редко консультируясь с основными учебниками.

С одной стороны, я старался избегать излишней опоры на школьную программу математики, которая часто мешает логической ясности. Поэтому перед глазами всегда должна быть видна сама логическая структура теории (частичный порядок в множестве теорем). Иначе говоря, для любого утверждения должен быть виден обратный путь к основным аксиомам. С другой стороны, везде явно и неявно должна присутствовать интуиция, которая помогает увидеть те же самые логические переходы.

Понять что необходимо для понимания математической структуры теоретической физики. Помимо множества экспериментальных фактов, множества алгебраических и численных выкладок, должно быть видно само существование строгой логической структуры отдельных теорий, снимающей все существующие противоречия как в отдельных теориях так и между разными физическими теориями. Полезно было бы научить студентов что одна из важнейших целей математической физики - получить максимальное количество качественных физических следствий из минимума аксиом. При этом вывод следствий должен быть наглядным но строгим. Наглядность означает "с доказательством в одну строку".

Этот текст также можно рассматривать также как введение к курсам [9, 10].

Мы почти не даем библиографии - только несколько ссылок. Есть много причин для этого:

1. Молодые математики должны, просматривая много математических книг, сами выбрать те, какие более подходят для их восприятия;
 2. Математические представления, способности, опыт могут быть совсем разные.
- И трудно посоветовать что-то конкретное, не видя индивидуальности студента.

Part II

Освоение математического языка

1 Множества

Основой математического языка (то есть языка на котором общаются математики, и на котором пишутся математические статьи и книги) являются два понятия (помимо натуральных чисел) - множество и элемент множества, которые часто называются не определяемыми (все остальные понятия математики определяются через них):

1. **Элемент** - от слова элементарный. Элементами могут быть любые объекты. Этот термин используется, когда мы на время забываем что элемент из чего то состоит (имеет собственную структуру), но знаем что он принадлежит определенному множеству. Как в физике есть термин “элементарная частица” - и физики могут понимать под этим, что мы еще не знаем из чего он состоит. Или выбирается минимальная возможная для измерений шкала.

Элементы могут называться точками, объектами, и т.д. , и могут обозначаться например латинскими буквами x, y, \dots , или буквами с индексами x_1, x_2, \dots , что подчеркивает принадлежность их одному и тому же множеству. Важно иметь в виду, что равенство $x = y$ не означает, что два элемента подобны, но только, что для элемента x мы используем и обозначение y . Например, в молекуле или в электрическом токе много одинаковых электронов, но они представляют разные элементы (находятся в любой момент времени в разных точках) и должны быть как то перенумерованы чтобы можно было следить за траекторией каждого из них.

2. **Множество** (часто используются другие взаимно заменяемые термины как класс, совокупность и др.). Например, множество всех деревьев в данной местности, всех букв русского алфавита, всех натуральных чисел больших 5, и др.. Если элемент x принадлежит множеству A , то пишут $x \in A$. Если например множество A состоит из трех элементов x_1, x_2, x_3 , то это обозначается как

$$A = \{x_1, x_2, x_3\} = \{x_i : i = 1, 2, 3\}$$

Первые конструкции в теории множеств Подмножество B множества A есть любое множество, элементами которого могут быть только элементы из A , но возможно не все. Точнее, из того что $x \in B$ необходимо следует, что $x \in A$. Пишут $B \subset A$ или $B \subseteq A$.

Пустое множество, обозначаемое \emptyset , есть множество в котором нет элементов. Оно является подмножеством любого множества, включая самого себя.

Объединение $A \cup B$ множеств A и B есть множество состоящее из тех и только тех элементов которые принадлежат хотя бы одному из них. Аналогично определяется объединение большего числа множеств.

Пересечение $A \cap B$ множеств A и B есть множество состоящее из тех и только тех элементов которые принадлежат обоим множествам, то есть и A и B . Аналогично определяется пересечение большего числа множеств.

Разность $A \setminus B$ множеств A и B есть множество состоящее из тех и только тех элементов, которые принадлежат A но не принадлежат B . Если $B \subset A$, то $A \setminus B$ называют также дополнением B в множестве A .

Говорят что совокупность попарно не пересекающихся подмножеств A_1, \dots, A_n (то есть $A_i \cap A_j = \emptyset$ для всех $i \neq j$) множества A определяет **разбиение** A на эти подмножества, если

$$A_1 \cup \dots \cup A_n = A$$

Декартово произведение $X \times Y$ двух множеств X и Y есть множество все пар (x, y) таких, что $x \in X$ и $y \in Y$. В соответствии с обозначением $X \times Y$, пары обозначаются с тем же порядком элементов, то есть на первом (левом) месте элемент X , а на втором (правом) - элемент Y . Аналогично определяется декартово произведение $X_1 \times X_2 \times \dots \times X_N$ любого числа N множеств.

Любое подмножество $A \subset X \times Y$ называется **отношением**. Если $(x, y) \in A$, то можно представлять что между x и y установлено некоторое отношение, обозначаемое A .

Функция f на множестве X со значениями в Y (или **отображение** множества X в Y , которое часто обозначается $f : X \rightarrow Y$), определяется как подмножество $C_f \subset X \times Y$ такое что любой элемент $x \in X$ встречается (на первом месте) в парах из C_f ровно один раз. Более аккуратно было бы называть это **однозначной функцией** (или однозначным отображением), но слово однозначная обычно опускается. Тогда, если пара (x, y) принадлежит C_f , говорят что функция f отображает $x \in X$ в $y = f(x) \in Y$, называемом значением функции f в точке x .

Ниже мы увидим, что в некоторых строго определенных случаях функции могут называться иначе: операторы, функционалы, меры и т.д.

Два множества A и B называются **изоморфными**, если существует подмножество F множества $A \times B$ такое, что в его парах каждый элемент A и каждый элемент B встречаются ровно один раз. Другими словами, подмножество F определяет функцию $f : A \rightarrow B$, для которой существует единственная (называемая **обратной**) функция $f^{(-1)} : B \rightarrow A$ такая что $f^{(-1)}(f(a)) = a$ для всех $a \in A$. Такая функция f называется **взаимно однозначной**. Говорят, что изоморфные множества имеют одинаковую **мощность** (обобщение понятия числа элементов). Множество называется **конечным** с числом элементов N если оно изоморфно множеству $\{1, 2, \dots, N\}$. Множество называется **счетным**, если оно изоморфно множеству всех натуральных чисел $\mathcal{N} = \{1, 2, \dots, N, \dots\}$.

Заметим что расширенное множество натуральных чисел $0, 1, 2, \dots, N, \dots$, получаемое добавлением к \mathcal{N} нуля, часто обозначается \mathcal{N}_0 .

2 Структуры на множествах

Структуры на произвольном множестве A часто определяются как некоторые специальные подмножества декартовых произведений $A \times A, A \times A \times A, \dots$. Пример: элемент N -кратного декартова произведения $A \times A \times \dots \times A$ часто называется **упорядоченным множеством** или **последовательностью** $(a_1, a_2, \dots, a_N) = a_1, a_2, \dots, a_N$ элементов $a_k \in A$, или **словом** $a_1 a_2 \dots a_N$ в **алфавите** A .

2.1 Алгебраические операции

Множество A называется **группой** если определена функция (операция) $f : A \times A \rightarrow A$, обозначаемая, например, как

$$a = f(a_1, a_2) = a_1 a_2$$

При таком обозначении эту операцию естественно называть умножением (a_1 умножается справа на a_2). Требуется, чтобы были выполнены следующие условия (аксиомы):

1. существует элемент (**единица** группы) $1 \in A$ такой что $1a = a1 = a$;
2. для любого $a \in A$ существует **обратный** элемент $a^{-1} \in A$ такой что $aa^{-1} = a^{-1}a = 1$;
3. **ассоциативность** - для всех $a_1, a_2, a_3 \in A$

$$(a_1 a_2) a_3 = a_1 (a_2 a_3)$$

Подмножество группы называется **подгруппой** если оно является группой относительно групповых операций в A . Большинство групп строится как подгруппы множества $\Lambda = \Lambda_X$ всех взаимно однозначных отображений f некоторого множества X в себя, которое является группой относительно умножения $f_1 f_2 = f_1(f_2)$. Единица этой группы - тождественное отображение I множества X в себя, то есть $I(x) = x$ для всех $x \in X$.

Если X конечно, то Λ_X называется **группой перестановок**. Перестановка является биективной (взаимно однозначной) функцией $f : A \rightarrow A$, где $A = \{1, 2, \dots, N\}$. Интуитивно она определяет на какое место $f(k)$ перейдет элемент, находящийся на месте k . Есть $N!$ таких перестановок, и их произведение, как обычно, определяется как $(f_1 f_2)(k) = f_1(f_2(k))$.

Задача 1 *Элементарной перестановкой называется перестановка двух элементов, то есть существуют $a_1, a_2 \in A$, такие что $f(a_1) = a_2, f(a_2) = a_1$, а для остальных a будет $f(a) = a$. Доказать, что любая перестановка f может быть представлена как проведение некоторого числа таких элементарных перестановок. Четность числа таких перестановок зависит только от f , называется **четностью** f и обозначается $\sigma(f)$.*

Группа называется коммутативной, если $f_1 f_2 = f_2 f_1$ для всех f_1, f_2 . Часто операция в коммутативной группе называется **сложением** и обозначается соответственно как $f_1 + f_2$. При этом единица становится нулевым элементом 0 таким что $f+0 = 0+f = f$. Например, на множестве целых чисел $Z = \{\dots, -2, -1, 0, 1, 2, \dots\}$, которое получается добавлением нуля и отрицательных целых чисел к множеству натуральных чисел. Целые числа тогда образуют группу (относительно сложения), где единицей является ноль, а обратный элемент к a будет $-a$.

Другим важным алгебраическим объектом является (коммутативное) **поле**. Так называется коммутативная группа по сложения (с операцией сложения) и нулевым элементом 0 . Кроме того и $A \setminus \{0\}$ является коммутативной группой с операцией умножения и единичным элементом 1 . При этом предполагается аксиома **дистрибутивности**:

$$a_1(a_2 + a_3) = a_1 a_2 + a_1 a_3$$

Первым примером поля является поле **рациональных чисел**, обозначаемая Q . Если сложение и умножение существуют уже для целых чисел, то рациональные числа возникают при попытке определить что такое **деление**. Интуитивно сначала мы вводим половину, треть и т.д. целого, то есть числа вида $\frac{1}{n}$ (где n целое) как числа такие что $n \cdot \frac{1}{n} = \frac{1}{n} \cdot n = 1$. **Формально**, мы множество формальных символов вида $\frac{p}{q}$, где p и $q \neq 0$ целые, и определяем операции на них так как мы учили в школе

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 q_2 + q_1 p_2}{q_1 q_2}, \quad \frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1 p_2}{q_1 q_2}, \quad -\frac{p}{q} = \frac{-p}{q} = \frac{p}{-q}, \quad \frac{p_1 q}{q p_2} = \frac{p_1}{p_2}$$

Задача 2 Определить когда $\frac{p_1}{q_1} = \frac{p_2}{q_2}$ и доказать, что множество рациональных чисел счетно.

Замечание 1 Было изобретено множество алгебраических операций. Мы не будем их четко определять, а только кратко упомянем главные (в дополнение к группам и полям) и приведем операции в них:

линейное пространство (коммутативное) - сложение и умножение на число;

полугруппа - только умножение (без обратного элемента);

кольцо - сложение и умножение;

алгебра - сложение, умножение и умножение на число,

модуль - абелева (то есть коммутативная) группа с умножением на элементы некоторого кольца.

2.2 Дискретная и локально дискретная математика

Под дискретным множеством понимается обычно либо конечное либо счетное множество. Его элементы могут называться **состояниями**, **точками**, **координатами** и др. Определим первичные понятия дискретной математики: граф, автомат, схема, алгоритм и др.

Граф G прежде всего состоит из двух подмножеств - множества V , элементы которого называются **вершинами**, и множества E , элементы которого называются **ребрами**. Они связаны заданным отображением $E \rightarrow V \times V$, то есть каждое ребро отождествляется с парой вершин $\{v_1, v_2\}$. Говорят, что ребро соединяет эти две вершины. Если пара (v_1, v_2) упорядочена, то говорят, что ребро **ориентировано** (от v_1 к v_2). Если все ребра графа ориентированы, то граф называется **ориентированным**.

Пока будем рассматривать неориентированные графы. **Путь** в графе, от вершины v_1 к вершине v_n , есть последовательность ребер $(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)$. Замкнутый путь называется **циклом**, то есть если $v_1 = v_n$. Граф называется **связным**, если между любыми двумя вершинами существует путь.

Первые примеры графов: 1) $V = \{0, 1, \dots, N-1, N\}$ и N ребер $(0, 1), (1, 2), \dots, (N-1, N)$. В частности, это может быть Z_2 , где $N = 1$. Или вся одномерная решетка Z , где вершины - все целые числа, а ребра - все отрезки между соседями; 2) двумерная решетка Z^2 на плоскости, где вершины - все пары (m, n) двух целых чисел, а ребра - все горизонтальные и вертикальные отрезки, соединяющие вершины (m, n) и $(m+1, n)$, а также (m, n) и $(m, n+1)$; 3) связный граф без циклов называется **деревом**.

Выделим в дереве одну вершину и назовем ее **корнем**. Дерево с фиксированным корнем будем называть **корневым деревом**. Из корня в любую другую вершину

существует единственный путь. Припишем каждой вершине порядок - число ребер в этом пути. Порядок корня будет 0. Тогда для каждой вершины порядка $k > 0$ существует ровно одно ребро, связывающее его с ребром порядка $k - 1$.

Гиперграф определяется аналогично, но отображение E будет не парой вершин, а подмножеством множества вершин. Наглядно можно понимать, что вершины этого «гиперребра» связаны некоторой общностью. Например, в физике это может быть несколько частиц, связанных некоторым физическим (в том числе и не двухчастичным) взаимодействием, а в социальных науках это может быть клан индивидуумов, связанных общностью психики или общностью целей.

Частичный порядок на множестве A есть подмножество $T \subset A \times A$, множества упорядоченных пар, такое что выполнены следующие условия:

а) пары с одинаковыми элементами, то есть вида (a, a) не могут принадлежать T ;

б) если $(a, b) \in T$ то (b, a) не принадлежит T . Мы говорим тогда, что a больше (выше, важнее, ранее) b (или b меньше a) и пишем $a > b$ (или $b < a$).

с) если $a > b$ и $b > c$, то $a > c$.

Если для любых двух элементов мы можем сказать, что один из них больше другого, то множество называется **вполне упорядоченным**. Примером могут быть множества \mathcal{N} и \mathcal{N}_0 .

Если множество A конечно или счетно, то с частичным порядком на нем можно связать граф G следующим образом. Множество (граф) G получается из множества T удалением всех пар (a_1, a_2) таких, для которых есть «промежуточный» элемент a , то есть такой что $a_1 > a > a_2$. Если каждый a , кроме одного, называемого **корнем**, может встретиться в парах (из G) справа ровно один раз, а корень встречается только слева, то такой граф является (корневым) **деревом**. Наглядно - каждый элемент (кроме корня) имеет ровно одного непосредственного начальника. Или, в каждую вершину (кроме корня-источника) втекает только один поток.

Булева функция определяется как отображение $f = f(x_1, \dots, x_n) : Z_2^n \rightarrow Z_2$. Примером могут быть полиномы Жегалкина

$$\sum_A C_A \prod_{i \in A} x_i,$$

где суммирование по всем подмножествам A множества $\{1, 2, \dots, n\}$. Кроме того, $C_A \in Z_2$ а операции (сложение и умножение) берутся по модулю 2. Нетрудно доказать, что любая булева функция от n переменных единственным образом представляется в виде полинома Жегалкина. Иначе говоря, все полиномы Жегалкина различны, а их число совпадает с числом булевых функций.

Более общие представления - **схемы из функциональных элементов**. Например, можно ли и как получить произвольную булеву функцию $f(x_1, \dots, x_n)$ от $n > 2$ переменных, используя только некоторое множество \mathbf{F} функций двух переменных (назовем \mathbf{F} базисом) и операцию подстановки. Для построения самой схемы рассмотрим корневое дерево такое что в каждую вершину v порядка k которого либо входит ровно 2 вершины v_1, v_2 большего порядка (то есть порядка $k + 1$) либо не входит ни одной. В последнем случае вершина будет называться начальной. Вершину v сделаем функциональным элементом, приписав ей некоторую функцию $f_v \in \mathbf{F}$.

Фиксируем значения всех переменных x_1, \dots, x_n . Тогда приобретут значения все функции, приписанные начальным вершинам. Если в вершину v входят два ребра

из вершин v_1, v_2 начального порядка, то функция $f_v = f_v(f_{v_1}, f_{v_2})$ также получит значение 0 или 1. Далее по индукции получит некоторое значение и функция приписанная корню дерева. Это значение и будем считать значением функции $f(x_1, \dots, x_n)$. Обозначим описанную схему из функциональных элементов через S . Ее сложностью $C_f(S)$ назовем число вершин этого дерева. Сложностью функции f (при заданном базисе \mathbf{F}) назовем $C(f) = \min_S C_f(S)$, Много исследований было посвящено оценкам максимальной сложности $\max_f C(f)$, где максимум берется по всем булевым функциям от n переменных. К сожалению, пока неизвестно какую роль это понятие сложности играет в математической физике.

Конечный **автомат** (без входов) есть конечное множество A , элементы $\{x\}$ которого называются **состояниями** автомата, вместе с отображением $f : A \rightarrow A$. Работа автомата (в дискретном времени $t = 0, 1, 2, \dots$) состоит в итерации отображения f . Можно представлять это так. Пусть автомат в начальный момент времени $t = 0$ находится в состоянии $x = x(0)$. Тогда в следующие моменты автомат будет переходить в состояния

$$x(1) = f(x(0)), x(2) = f(x(1)), \dots, x(t+1) = f(x(t)), \dots \quad (1)$$

Отображение f порождает полугруппу отображений $f^n : A \rightarrow A$, где $f^0 = I$ есть тождественное отображение, $f^1 = f$ и $f^m f^n = f^{m+n}$ для всех $m, n \in \mathbb{Z}_+$. Если отображение взаимно-однозначно, то эта полугруппа становится группой с элементами $f^n, n \in \mathbb{Z}$, где f^{-1} - отображение обратное к f .

По отображению f можно построить ориентированный **граф переходов**. Его ребрами будут все пары вида $(x, f(x))$. Нетрудно доказать, что граф любого взаимно-однозначного отображения разбивается на связные компоненты, каждое из которых является ориентированным циклом. Более того, каждая связная компонента графа любого отображения содержит ровно один цикл.

Здесь возникает и другая картина, связанная с динамикой частиц в физике. Можно представить множество A как некоторое «пространство», а некоторый объект (например, частица) в момент $t = 0$ был в точке $x = x(0)$ этого «пространства». В следующие моменты наша частица будет проходить траекторию (1).

Автоматы являются примером реализации разных **алгоритмов**, которые определяют как по известной в данный момент информации (набору символов) предпринять необходимые действия (другой набор символов).

Существуют и так называемые **клеточные автоматы**, где автомат состоит из конечного множества X клеток $x = 1, 2, \dots, N$, Каждая клетка x может быть в одном из M состояний $u = u(x) = u_1, \dots, u_M$. Состоянием клеточного автомата в момент t называется набор $U(t) = \{u(x, t) : x \in X\}$. Это понятие соответствует физическому понятию поля, как некоторой функции $u(x, t)$ на физическом пространстве X . Одним из обобщений является конечно детерминированный **процесс с локальным взаимодействием**, где X будет счетным множеством. Пусть $X = \mathbb{Z}$ и рассмотрим множество $\mathbb{Z} \times W$, где W - произвольное конечное множество. Каждая точка $x \in \mathbb{Z}$ в любой момент $t = 0, 1, 2, 3, \dots$ может находиться в одном из состояний $w = w(x, t) \in W$. Состоянием всего процесса в момент t называется набор $w(t) = \{w(x, t), x \in \mathbb{Z}\}$ состояний всех точек x . Кроме того, заданы функции $f_x : W^3 \rightarrow W$, которые и задают динамику всего процесса как

$$w(x, t+1) = f_x(w(x-1, t), w(x, t), w(x+1, t))$$

Формально, это понятие выходит за рамки дискретной математики, так как множество несчетно. Но так как процесс в каждой клетке x зависит лишь от конечного числа клеток, то можно говорить о локально дискретной динамике. Более общо, можно вместо Z взять не более чем счетный неориентированный граф G с множеством вершин $V = X$. Тогда переходные функции f_x могут зависеть от значений $w(y, t)$, где y соединено с x ребром.

Следующее понятие - **автомат со входом и(или) с выходом**, где в каждый момент t времени «извне» приходит некоторый символ $i = i(t)$ - элемент заданного конечного множества I . При этом состояние $x(t+1)$ автомата в момент $t+1$ зависит не только от $x(t)$, но и от пришедшего символа $i(t)$, то есть $x(t+1) = f(x(t), i(t))$, то есть f становится функцией двух переменных - $f(x, i), x \in A, i \in I$. Более того,, в момент $t+1$ автомат производит некий «выход», определяемый другой заданной функцией $g(x, t)$. **Машина Тьюринга** есть такой автомат со входом, но более того - передвигающийся еще по ленте - клеточному автомату, где автомат может менять состояние клетки, на которой он в данный момент находится.

Все введенные здесь понятия имеют аналоги в математической физике. Но для четкого понимания этих аналогий нужны разного рода обобщения - уходы от дискретности:

1) первое из них конечно введение несчетных множеств, и в частности переход к непрерывному времени. Здесь вместо переходных функций возникают дифференциальные уравнения. Обыкновенные, если одна или несколько частиц движутся по пространству. И в частных производных, если речь идет о полях. Здесь уже возникают задачи совместимости разных теорий;

2) но еще более важное направление развития - это скейлинги (типа микро-макро), о которых мы поговорим ниже.

Обратно, в вычислительной математике идет обратный процесс - из дифференциальных уравнений с континуальными множествами возникают разные дискретные вычислительные схемы и автоматы, которые дают приближенные решения соответствующих уравнений.

2.3 Метрика и топология

Здесь все числа (и в частности ϵ) рациональны.

Метрика (или **расстояние**) на множестве X есть симметрическая функция $\rho : X \times X \rightarrow Q$, то есть функция $\rho(x_1, x_2) = \rho(x_2, x_1)$ на $X \times X$ со значениями в поле рациональных чисел, положительная при $x_1 \neq x_2$ и равная нулю при $x_1 = x_2$. Более того, требуется выполнение так называемого неравенства треугольника: для всех $x_1, x_2, x_3 \in X$

$$\rho(x_1, x_3) \leq \rho(x_1, x_2) + \rho(x_2, x_3)$$

Множество с такой метрикой называется **метрическим пространством**.

Замечание 2 *Может показаться странным, что метрика вводится с использованием поля рациональных чисел. Есть две причины. Первая состоит в том, что мы еще не ввели поля вещественных чисел. Вторая, более важная, - что все существующие физически теории - только аппроксимации, а значит для практики достаточно конечного числа цифр в десятичном представлении вещественного числа достаточно. Хотя конечно поле вещественных чисел исключительно удобный объект, но сомнительно его рассматривать как физическую реальность.*

Окрестностью (точнее, ϵ -окрестностью, где $\epsilon > 0$) точки x называется множество $O_\epsilon(x) = \{y : \rho(x, y) < \epsilon\}$. Это частое обозначение которое читается так - множество всех $y \in X$, находящихся на расстоянии меньше ϵ от x .

Теперь мы подошли к двум важнейшим понятиям, пронизывающим всю математику. Последовательность элементов (точек)

$$x_1, x_2, \dots, x_k, \dots \quad (2)$$

метрического пространства X имеет **предел** (предельную точку) $x \in X$, если для любой окрестности O_ϵ точки x существует натуральное число $N = N(\epsilon)$ такое что для всех $k > N$ элементы x_k принадлежат этой окрестности. Говорят также что последовательность (2) **сходится** к x (или является сходящейся).

Последовательность точек (2) метрического пространства X удовлетворяет **условию Коши** (или называется **последовательностью Коши**) если для любого $\epsilon > 0$ существует натуральное число $N = N(\epsilon)$ такое что $\rho(x_k, x_m) < \epsilon$ для всех $k, m > N$. Образно говоря, точки группируются в тесный кластер.

Оба эти определения автоматически обобщаются от счетной последовательности на любое вполне упорядоченное множество. Например, предположим что мы ввели отношение полного порядка на некотором подмножестве $X' \subset X$. Тогда X' имеет предел $x \in X$ если для любой окрестности O_ϵ точки x существует элемент $x_\epsilon \in X'$ такой, что любая точка $x' > x_\epsilon, x' \in X'$, принадлежит O_ϵ .

Метрическое пространство X называется **полным**, если любая последовательность Коши сходится к некоторому элементу из X .

Открытое множество метрического пространства X определяется как любая окрестность или объединение окрестностей в любом числе.

Задача 3 Доказать что пересечение конечного числа открытых множеств открыто.

Границей ∂A открытого множества $A \subset X$ называется множество точек $x \in X$, не принадлежащих A но таких, что каждая окрестность x имеет непустое пересечение с A . Подмножество A метрического пространства X называется **замкнутым** если его дополнение $X \setminus A$ открыто.

Подмножество A полного метрического пространства называется **компактным** (или просто **компактом**) если любая последовательность его точек имеет сходящуюся подпоследовательность. Эквивалентное определение (доказать !!!): из любого покрытия A открытыми множествами можно извлечь конечное (под)покрытие. При этом система множеств называется **покрытием** множества A , если их объединение содержит A .

Говорят, что система открытых множеств на X определяет **топологию** на X . Топология (то есть система открытых множеств) может быть введена без помощи метрики, но здесь нам это не понадобится.

Замечание 3 (Это общее замечание) *Есть два источника развития (идей) математики:*

1) *математический язык со всей структурой его понятий. Число людей свободно владеющих этим языком очень мало, в отличие например от русского или английского языков. Однако, математический язык далеко не основной источник развития науки.*

2) *Интуиция в математике основана в основном на не математических дисциплинах (в основном физических), но и на всей жизни. Совсем новые понятия в математике*

могут возникать из такой интуиции. Однако, возникали многие глубокие идеи вроде бы не несвязанные с окружающей жизнью, и потом постепенно затухали. Примерами могут быть - знаменитая гипотеза континуума, аналитическая теория множеств. Однако, затухание не означает, что эти теории будут навсегда забыты или что они не имеют ничего общего с нашим проектом. Не исключено, что однажды они станут базисом совсем новых теорий. Более того, такие книги как например [1, 2, 3] помогают развить ощущение логики в жизни, а также структуру и глубину математики.

3 Вещественные и комплексные числа

Дадим сначала три определения вещественных чисел. В каждом из них самое важное - установить условия равенства двух объектов, которые мы и будем называть вещественными числами. Так, вещественным числом называется:

1. последовательность Коши $x_1, x_2, \dots, x_n, \dots$ рациональных чисел. Две таких последовательности x_n и y_n определяют одно и то же число тогда и только тогда когда для любого (рационального) $\epsilon > 0$ существует $N = N(\epsilon)$ такое что для всех $n > N, m > N$

$$|x_n - y_m| < \epsilon$$

Если эта последовательность Коши имеет предел из Q , то так определенное вещественное число отождествляется с этим рациональным числом. Таким образом, любое рациональное число вещественно.

2. сечение множества рациональных чисел, то есть разбиение множества рациональных чисел Q на два подмножества Q_1 и Q_2 таких что для всех $x_1 \in Q_1, x_2 \in Q_2$ будет $x_1 < x_2$. Все так определенные вещественные числа различны и, более того, не рациональны за исключением следующих случаев: 1) существует максимум множества Q_1 , который рационален, то есть рациональное число $r \in Q_1$ такое что $x < r$ для всех других $x \in Q_1$; 2) существует минимум множества Q_2 , который рационален, то есть рациональное число $r \in Q_2$ такое что $x > r$ для всех других $x \in Q_2$. в этих и только в этих случаях сечение определяет рациональное число. Значит снова все рациональные числа вещественны.

3. двусторонняя бесконечная последовательность десяти цифр, $a_n = 0, 1, \dots, 9$, вида

$$\alpha = \dots a_{-m} \dots a_{-2} a_{-1}, a_0, a_1 a_2 \dots a_n \dots, \quad (3)$$

где $m, n > 0$, причем только конечное число цифр a_{-k} , с $k > 0$, отлично от нуля. При этом нули $a_{-k}, k > 1$, слева от всех ненулевых a_k обычно не пишутся, то есть десятичное представление имеет вид

$$\alpha, a_1 a_2, \dots,$$

где $\alpha \geq 0$ целое. Все числа, определенные подобными последовательностями считаются различными, с одним исключением: любые два числа α и β вида

$$\alpha = \dots a_{n-1} a_n a_{n+1} \dots, \beta = \dots b_{n-1} b_n b_{n+1} \dots$$

считаются равными если для некоторого n выполнены следующие 5 условий: 1) $a_k = b_k$ для всех $k < n$; 2) $b_n \neq 9$; 3) $a_n = b_n + 1$; 4) $a_k = 0$ для всех $k > n$; 5) $b_k = 9$

для всех $k > n$. Короче, они различны всегда кроме пар чисел вида

$$\alpha, \beta 0000\dots = \alpha, (\beta - 1)9999\dots$$

где α, β натуральные числа в десятичном представлении (в β последняя цифра не ноль). В частности,

$$1 = 0,999\dots$$

И вещественные числа образуют вполне упорядоченное множество: $\alpha > \beta$ если существует (минимальное) n такое что $a_n > b_n$ и $a_k = b_k$ для всех $k < n$.

Задача 4 Доказать эквивалентность всех трех определений. Иначе говоря, для любого вещественного числа в смысле одного определения указать соответствующее число в смысле другого определения.

Задача 5 Определим метрику $\rho(x,y) = |x - y|$ на множестве вещественных чисел. Используя каждое из трех определений, доказать что единичный отрезок $[0,1]$ компактен в смысле каждого из двух определений компактности.

Вещественные числа образуют поле. Чтобы увидеть это, для любого вещественного числа α , см. (3), пусть $\alpha(N) = \dots b_n b_{n+1} \dots$ - рациональное число такое что $b_n = a_n$ для $n \leq N$, и $b_n = 0$ для $n > N$. Тогда сложение, умножение и деление определяются как пределы соответствующих операций на рациональных числах вида $\alpha(N)$ при $N \rightarrow \infty$.

Рассмотрим например сложение. Определим сумму двух вещественных чисел α и β :

$$\alpha + \beta = \lim_{N \rightarrow \infty} (\alpha_N + \beta_N),$$

где символ $\lim_{N \rightarrow \infty}$ означает предел последовательности $\alpha(N)$ при $N \rightarrow \infty$. в следующем смысле. Для любого числа γ и любого $n \leq N$, обозначим $\gamma_n(N)$ n -тый член в последовательности $\gamma(N)$. Пусть теперь α и β заданы. Тогда для любого n и для всех $N > n + 1$ цифры $\gamma_n(N)$ одинаковы.

Множество рациональных чисел \mathbb{Q} удобно рассматривать как подмножество множества вещественных чисел \mathbb{R} , используя десятичное представление. При выполнении стандартной процедуры деления для рационального числа $\frac{p}{q}$, легко видеть что получается бесконечная последовательность вида

$$\alpha, \beta \gamma \gamma \gamma \dots$$

где $\alpha, \beta > 0, \gamma > 0$ - целые в десятичном представлении. Наоборот, докажем что любая такая последовательность определяет рациональное число. Пусть целое число γ имеет ровно k цифр:

$$0, \gamma \gamma \gamma \dots = \lim_{N \rightarrow \infty} \sum_{n=1}^N \frac{\gamma}{10^{kn}} = \frac{1}{1 - 10^{-k}} \frac{\gamma}{10^k}.$$

вспоминая алгебраическую формулу из школы

$$\frac{1 - x^n}{1 - x} = 1 + x + \dots + x^{n-1} \tag{4}$$

что в пределе дает

$$\lim_{n \rightarrow \infty} \frac{1 - x^n}{1 - x} = \frac{1}{1 - x} \tag{5}$$

Комплексные числа Комплексное число формально есть выражение вида $x + iy$, где x, y вещественны, а i есть формальный символ называемый **мнимой единицей**.

Сложение и умножение комплексных чисел определяется как:

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

$$(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

(достаточно помнить что $i^2 = -1$). Множество комплексных чисел обозначается C и чтобы определить на нем структуру (коммутативного) поля, достаточно проверить что любой элемент $a + bi$ имеет обратный

$$\frac{1}{a^2 + b^2}(a - bi)$$

и конечно $0 = 0 + 0i, 1 = 1 + 0i$. Метрически C отождествляется с плоскостью R^2 : комплексное число $x + iy$ отождествляется с точкой $(x, y) \in R^2 = R \times R$ плоскости.

Единичная окружность $S = \{a + bi : a^2 + b^2 = 1\}$, как подмножество множества комплексных чисел, инвариантно относительно умножения и деления но не деления. Иначе говоря, это коммутативная группа относительно умножения. Например, возведение в квадрат дает

$$(a + bi)^2 = a^2 - b^2 + 2abi \implies (a^2 - b^2)^2 + 4a^2b^2 = (a^2 + b^2)^2 = 1,$$

и аналогично для умножения и деления

$$(a + bi)(c + di) = ac - bd + (ad + bc)i \implies a^2c^2 + b^2d^2 - 2abcd + a^2d^2 + b^2c^2 + 2abcd =$$

$$= a^2c^2 + (1 - a^2)d^2 + a^2d^2 + (1 - a^2)c^2 = 1,$$

$$\frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2} = a - bi$$

Абсолютное значение (**модуль**) комплексного числа $z = x + iy$ определяется как, см. определение квадратного корня в (6),

$$r = |z| = \sqrt{x^2 + y^2}$$

А метрика на C , или на двумерной плоскости, определяется как

$$\rho(z, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2},$$

где $x = x_1 + ix_2, y = y_1 + iy_2$.

Part III

Начала анализа

4 Производные и интегралы

Одномерный вещественный анализ имеет дело с функциями $f : A \rightarrow R$, а комплексный анализ - с функциями $f : A \rightarrow C$, где A - некоторое подмножество R или C соответственно. Чаще всего, эти подмножества открыты либо замкнуты.

Напомним что в метрическом пространстве открытое множество есть объединение окрестностей своих точек. Соответственно любое открытое множество на вещественной прямой есть объединение открытых интервалов вида $(a, b) = \{x : a < x < b\}$ для $a < b$.

Первые представления, касающиеся рассматриваемых функций, появляются обычно при виде их **графика**. Если рассматривается функции на R , то график есть кривая на плоскости вдоль вещественной оси. Самые простые функции - константы и кусочно постоянные функции. Последнее означает, что область определения функции разделена на конечное или счетное число подмножеств, на каждом из которых функция постоянна. Более сложные кривая могут иметь разные степени гладкости, к точному определению гладкости мы сейчас и переходим.

4.1 Шкалы поведения функций

Часто необходимо сравнивать функции по степени их роста или убывания. Например, ясно что функция x^{n+1} растет быстрее чем x^n при $x \rightarrow \infty$, то есть при неограниченном росте x . Но попробуйте, используя только определение сложения и умножения, доказать что $n^{1000}2^{-n} \rightarrow 0$ при $n \rightarrow \infty$. Или представьте как сравнивать скорости роста при $N \rightarrow \infty$ таких функций как (с константой $a > 0$) :

$$a, \log_a N, N^a, 2^{aN}, N^{aN}, ..$$

Но конечно надо дать сначала четкое определение этих функций (см. ниже в этой главе). И математический анализ как раз разрабатывает методы решения гораздо более сложных подобных задач.

Есть три часто используемых обозначения. Например, пусть заданы две функции f, g в некоторой окрестности точки a :

1. Говорят что $f(x)$ есть **O -большое** от $g(x)$ при $x \rightarrow a$, и пишут $f(x) = O(g(x))$ если существует константа $C > 0$ такая, что для всех x в некоторой окрестности точки a

$$|f(x)| \leq C|g(x)|$$

2. Рассмотрим поведение двух функций f и $g \neq 0$ при $x \rightarrow a$. Говорят, что f есть **o -малое** от g при $x \rightarrow a$, и пишут $f = o(g)$, если

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0,$$

3. Говорят, что f и $g \neq 0$ **асимптотически эквивалентны** при $x \rightarrow a$ и пишут $f(x) \sim g(x)$, если

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$$

Но часто нужно более точное сравнение, называемое асимптотическим разложением. Например, формула

$$f(N) = N^2 + N + 1 + o(1)$$

есть **асимптотическое разложение** функции $f(N)$ с точностью до $o(1)$ при $N \rightarrow \infty$. Это означает, что $f(N) - (N^2 + N + 1) \rightarrow 0$. Отсюда также следует, что $f(N) \sim N^2$.

Непрерывность и пределы Функция f , определенная в некоторой окрестности точки x , называется **непрерывной** в точке x , если для любого (иногда говорят "достаточно малого") $\delta > 0$ существует $\epsilon = \epsilon(\delta, x) > 0$ такое, что для всех $y \in O_{\epsilon(\delta, x)}(x) = (x - \epsilon, x + \epsilon)$

$$|f(x) - f(y)| < \delta$$

Короче можно сказать: если

$$f(y) - f(x) \rightarrow_{y \rightarrow x} 0$$

Функция, определенная на открытом или замкнутом множестве A , называется непрерывной на A , если она непрерывна в каждой точке $x \in A$. Если для всех $\delta > 0$ существует число $\epsilon = \epsilon(\delta)$, независимое от x , то говорят, что f **равномерно непрерывна** на A .

Нетрудно доказать, используя второе определение компактности, что непрерывная на сегменте $[a, b]$ функция является равномерно непрерывной на нем.

Более общо, правый (левый) пределы на вещественной оси определяются соответственно как

$$\lim_{y \rightarrow x+0} f(y) = f(x), \quad \lim_{y \rightarrow x-0} f(y) = f(x),$$

где $y \rightarrow x+0$ ($y \rightarrow x-0$) означает что рассматриваются только последовательности y_k такие что $y_k > x$ ($y_k < x$). Если эти пределы различны, то их разность называется скачком в этой точке. Есть более сложные типы разрывности функций:

1) в точке и правый и левый пределы могут не существовать. Например, в точке ноль у функции $\sin \frac{1}{x}$ (см. ниже определение тригонометрических функций);

2) пределы могут не существовать во всех точках. Пример - функция равная 1 в рациональных и 0 в иррациональных точках.

Множество всех непрерывных функций на множестве A обозначается $C_0(A)$.

Функция f называется **монотонной** на A , если она либо **возрастающая** либо **убывающая**, то есть для всех $x < y$ из A либо $f(x) < f(y)$ либо $f(x) > f(y)$.

Задача 6 Доказать, что функция x^2 отображает сегмент $[0, \infty)$ в себя взаимно однозначно, как и обратная к ней функция, обозначаемая \sqrt{x} . Более того, обе функции непрерывны и возрастают на этом сегменте.

Гладкость Существенно более сильное (чем непрерывность) ограничение - условие Липшица на функцию f на некотором подмножестве Λ , на котором она определена: существует $C > 0$ такое, что для всех $x, y \in \Lambda$ будет $|f(x) - f(y)| < C|x - y|$.

Еще более сильное условие - существование следующего предела (**дифференцируемость** в точке x)

$$\lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon} = \lim_{\epsilon = \Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta f(x)}{\Delta x}, \quad (6)$$

где $\Delta x = \epsilon$ называется **приращением** аргумента, а $\Delta f = f(x + \Delta x) - f(x)$ - приращением функции (при заданном приращении аргумента). Этот предел называется **производной** функции в точке x , и обозначается

$$\frac{df}{dx}(x) = f'(x) = f_x(x) = f^{(1)}(x),$$

где обозначения dx, df часто называют бесконечно малыми приращениями (как бы в пределе). Функция f при этом называется **дифференцируемой** в точке x . Условие дифференцируемости может быть переформулировано в терминах **разложения Тейлора**

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + o(\Delta x)$$

Действительно, если обозначить разность

$$g(\Delta x) = \frac{f(x + \Delta x) - f(x)}{\Delta x} - f'(x),$$

то $g(\Delta x) = o(1)$, и отсюда $g(\Delta x)\Delta x = o(\Delta x)$.

Функция называется дифференцируемой (или **гладкой**) на некотором интервале, если она дифференцируема в каждой его точке.

Если функция $f'(x)$ также имеет производную, называемую второй производной, обозначим

$$\frac{d^2 f}{dx^2}(x) = f''(x) = f_{xx}(x) = f^{(2)}(x),$$

и говорим, что f дважды дифференцируема. По индукции, производная любого порядка $n = 3, 4, \dots$, может быть аналогично определена, если она существует.

4.2 Вычисление производных

Полиномы

$$\frac{d(x^n)}{dx} = \lim_{\epsilon} \frac{(x + \epsilon)^n - x^n}{\epsilon} = \lim_{\epsilon} \frac{1}{\epsilon} (\epsilon n x^{n-1} + O(\epsilon^2)) = n x^{n-1}$$

и затем индукцией по $k \leq n$

$$\frac{d^k x^n}{dx^k} = n(n-1)\dots(n-k+1)x^{n-k}$$

Благодаря линейности операции дифференцирования (производная суммы дифференцируемых функций равна сумме их производных), многочлены степени n дифференцируемы любое число раз k , причем их производные равны нулю, если $k > n$.

Говорят о разложении Тейлора порядка n в точке x_0 , если существуют константы d_0, d_1, \dots, d_n такие что при $x - x_0 \rightarrow 0$

$$f(x) - \sum_{k=0}^n d_k (x - x_0)^k + o((x - x_0)^n) \quad (7)$$

Тогда производные $f^{(k)}(x_0)$ существуют при $k \leq n$, причем они равны $k!d_k$. С другой стороны обозначим разность

$$\xi_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k,$$

Дифференцируя, получаем, что производные $\frac{d\xi_n(x_0)}{dx^k}$ равны нулю для $k = 0, 1, \dots, n$.

Производная произведения Предположим что f и g дифференцируемы в точке x . Тогда

$$\begin{aligned} \frac{1}{\epsilon}(f(x+\epsilon)g(x+\epsilon) - f(x)g(x)) &= \frac{1}{\epsilon}(f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon)) + \frac{1}{\epsilon}(f(x)g(x+\epsilon) - f(x)g(x)) = \\ &= \frac{1}{\epsilon}g(x+\epsilon)(f(x+\epsilon) - f(x)) + \frac{1}{\epsilon}f(x)(g(x+\epsilon) - g(x)) \rightarrow f'(x)g(x) + f(x)g'(x) \end{aligned}$$

В качестве примера вычислим производную функции $\frac{1}{q(x)}$ в точке x , где $q(x) \neq 0$:

$$0 = (q \frac{1}{q})' = q' \frac{1}{q} + q (\frac{1}{q})' \implies (\frac{1}{q})' = -\frac{q'}{q^2}$$

Производная сложной функции Пусть f и g дифференцируемы соответственно в точках $g(x)$ и x . Тогда

$$\begin{aligned} \frac{f(g(x+\epsilon)) - f(g(x))}{\epsilon} &= \frac{f(g(x) + g'(x)\epsilon + o(\epsilon)) - f(g(x))}{\epsilon} = \\ &= \frac{f(g(x)) + f'(g(x))(g'(x)\epsilon + o(\epsilon)) + o(g'(x)\epsilon + o(\epsilon)) - f(g(x))}{\epsilon} \xrightarrow{\epsilon \rightarrow 0} f'(g(x))g'(x) \end{aligned} \quad (8)$$

Производная обратной функции Пусть $y(x)$ отображает интервал (a, b) взаимно однозначно на некоторый другой интервал (a_1, b_1) , и пусть $x(y)$ - обратная к ней функция на (a_1, b_1) . Обе функции предполагаются **непрерывно дифференцируемыми**, то есть имеющими непрерывные производные. Тогда $x(y(x)) = x$ и, как в (8),

$$\frac{d(x(y(x)))}{dx} = \frac{dx}{dy} \frac{dy}{dx} = 1 \implies x'(y) = \frac{1}{y'(x)}$$

4.3 Суммы, ряды, интегралы

Ряд есть формальное выражение вида $\sum_{k=0}^{\infty} a_k$, где a_k могут быть вещественными (или комплексными) числами или функциями. В последнем случае говорят о **функциональном ряде**. Ряд называется **сходящимся**, если последовательность **частичных сумм**

$$S_N = \sum_{k=0}^N a_k$$

сходится. Этот предел $S = \lim_{N \rightarrow \infty} S_N$ называется суммой ряда.

Замечание 4 Если все a_k имеют одинаковые знаки, то сумма ряда не зависит от их порядка.

Если же члены ряда имеют разные знаки, то важную роль играет порядок членов ряда. Однако иногда требуется говорить о сумме счетного числа членов независимо от их нумерации. Тогда надо просуммировать отдельно все неотрицательные a_k и затем все отрицательные. Если по крайней мере одна из сумм конечна, то сумма всего ряда определяется как сумма этих двух членов. Если же обе суммы бесконечны, то надо придумывать другие способы суммирования. Ряд называется **абсолютно сходящимся**, если возрастающая последовательность положительных чисел $\sum_{k=0}^N |a_k|$ ограничена. В этом случае порядок членов ряда также не играет роли.

В теории **кластерных разложений** рассматриваются гораздо более сложные ряды - их члены нумеруются конечными подмножествами счетных графов.

Интеграл кусочно постоянной функции Рассмотрим возрастающую последовательность вещественных чисел $x_0 < x_1 < x_2 < \dots$ и (кусочно постоянную) функцию $f(x)$ на $[x_0, \infty)$, равную a_k на полуинтервалах $[x_k, x_{k+1})$, где $k = 0, 1, \dots$. Интегралы от подобных функций (соответственно на интервалах $[x_0, x_N]$ или $[x_0, \infty)$) определяются соответственно как суммы или ряд

$$\int_{x_0}^{x_N} f(x) dx = \sum_{k=0}^{N-1} a_k (x_{k+1} - x_k), \quad \int_{x_0}^{\infty} f(x) dx = \sum_{k=0}^{\infty} a_k (x_{k+1} - x_k) \quad (9)$$

Последний интеграл называется сходящимся, если ряд справа сходится.

Интеграл от непрерывной функции Теперь мы хотим расширить интегрирование на более широкий класс функций. Рассмотрим например функцию $f(x) = x$ на интервале $[0, y]$. Естественное желание - найти последовательность кусочно-постоянных функций $f_N(x)$ на $[0, y]$, $y \geq 0$, **равномерно сходящуюся** к $f(x)$ при $N \rightarrow \infty$, то есть такую что существует последовательность положительных чисел $\delta_N \rightarrow 0$ такую, что для всех $x \in [0, y]$

$$|f(x) - f_N(x)| < \delta_N$$

И определить интеграл от $f(x)$ как предел интегралов от этих кусочно-постоянных функций

$$\int_0^y f(x) dx = \lim_{N \rightarrow \infty} \int_0^y f_N(x) dx$$

В данном случае это сделать легко, и даже предел вычислить. Введем кусочно-постоянную функцию $f_N(x) = x_k = \frac{k}{N}$, если $x_k = \frac{k}{N} \leq x < x_{k+1} = \frac{k+1}{N}$. Тогда, для любого $0 \leq y$ существует m такое что $0 \leq y - \frac{m}{N} \leq \frac{1}{N}$, или $yN - 1 \leq m \leq yN$. Тогда при $N \rightarrow \infty$

$$S_N(y) = \int_0^y f_N(x) dx = \sum_{k=0}^m \frac{k}{N} \frac{1}{N} + O\left(\frac{1}{N}\right) = \frac{1}{N^2} \frac{m^2}{2} + O\left(\frac{1}{N}\right) = \frac{y^2}{2} + O\left(\frac{1}{N}\right) \rightarrow \frac{y^2}{2} \quad (10)$$

Аналогично, для любой непрерывной (скажем на интервале $[0, 1]$) функции $f(x)$ можно доказать сходимость интегралов от кусочно-постоянных функций $f_N(x) = f(x_k)$, где x_k - точка вида $\frac{k}{N}$. ближайшая к x слева.

Задача 7 Рассмотрим снова непрерывную функцию $f(x)$ на конечном интервале $[a, b]$. Назовем $R(f, B) = \sup_{x \in B} f(x) - \inf_{x \in B} f(x)$ **размахом** функции f на произвольном подмножестве $B \subset [a, b]$. Здесь **супремум** $\sup_{x \in B} f(x)$ определяется как минимальное число не меньшее любого числа $f(x)$, $x \in B$. Аналогично определяется **инфимум** $\inf_{x \in B} f(x)$ как максимальное число, не меньшее любого числа $f(x)$, $x \in B$.

Доказать, что для любого $\epsilon > 0$, интервал $[0, 1]$ может быть разделен на конечное число полуинтервалов $A_k(\epsilon)$, $k = 1, \dots, N = N(\epsilon)$ так, чтобы $R(f, A_k(\epsilon)) < \epsilon$ для всех k . Воспользоваться свойством равномерной непрерывности $f(x)$ на этом сегменте.

Однако, мы не сможем вычислить этот интеграл для любой f как для $f(x) = x$. Вернемся назад к формуле (10) и удивимся, что функция x является производной своего интеграла $\frac{x^2}{2}$. Оказывается, это чудо есть очевидное правило и в общем случае. Для этого отметим два свойства аддитивности интеграла, которые следуют из его определения:

$$\int_0^y (f_1(x) + f_2(x)) dx = \int_0^y f_1(x) dx + \int_0^y f_2(x) dx,$$

$$\int_0^y f(x) dx = \int_0^{y_1} f(x) dx + \int_{y_1}^y f(x) dx, \quad 0 \leq y_1 \leq y$$

В одномерном случае, **анти-производной** (или **неопределенным интегралом**) функции $f(x)$, $x \in A = [a, b] \subset R$, называется любая функция (если такие существуют) на A , чья производная равна $f(x)$.

Задача 8 Доказать единственность неопределенного интеграла (с точностью до прибавления константы). Иначе говоря, доказать, что если функция дифференцируема и ее производная тождественно равна нулю на A , то она константа.

Чтобы понять связь интеграла с анти-производной всегда надо помнить что производная соответствует нормированной разности, а интеграл - нормированной сумме значений функции. Поэтому, разобьем интервал $[a, b]$ на N частей

$$a = x_0 < x_1 < \dots < x_N = b,$$

так что существуют константы $0 < C_1 < C_2$ такие что для всех k

$$\frac{C_1}{N} < x_{k+1} - x_k < \frac{C_2}{N}$$

Возьмем функцию f_N равную $f(x_k)$ на интервале $[x_k, x_{k+1}), k = 0, 1, \dots, N - 1$. Тогда ее интеграл записывается как

$$\int_a^b f(x)dx = \lim \sum_{k=0}^{N-1} f(x_k)(x_{k+1} - x_k)$$

Если $f(x)$ непрерывна на $[a, b]$, то этот предел существует. Положим $b = x$ и найдем производную интеграла в точке x как предел

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_x^{x+\epsilon} f(t)dt = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_x^{x+\epsilon} (f(x) + O(\epsilon))dt = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\epsilon f(x) + o(\epsilon)) = f(x) \quad (11)$$

Значит, интеграл и является анти-производной функции $f(x)$. Более легкий способ - вспомнить формулу для суммы разностей

$$f(x_N) - f(x_0) = \sum_{k=0}^{N-1} (f(x_{k+1}) - f(x_k)) = \sum_{k=0}^{N-1} \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k} (x_{k+1} - x_k)$$

и ее предел

$$f(b) - f(a) = \int_a^b f'(x)dx = \int_a^b f'(x)dx$$

Благодаря этому, часто вычисление интеграла сводится к методам вычисления производных. Например, метод **интегрирования по частям** позволяет свести вычисление одного из интегралов в левой части к вычислению другого

$$\int_a^b f'gdx + \int_a^b fg'dx = \int_a^b (fg)'dx = f(b)g(b) - f(a)g(a)$$

4.4 Длина кривой

Определим d -мерное пространство $R^d = \{(x_1, \dots, x_n)\}$ как множество всех последовательностей (x_1, \dots, x_n) вещественных чисел $x_k \in R$. Оно становится полным метрическим пространством, если ввести на нем евклидово расстояние $\rho(x, y) = \sqrt{\sum_{k=1}^d (y_k - x_k)^2}$.

Определим **кривую** в $R^d = \{(x_1, \dots, x_n)\}$ как отображение интервала $[a, b]$ в R^d . Это отображение может быть задано непрерывной вектор-функцией $x(t) = (x_1(t), \dots, x_n(t))$, то есть все функции $x_k(t)$ непрерывны, где $t \in [a, b]$. Кривая называется **прямой**, если все функции $x_k(t)$ линейны. Кривая называется **ломаной** (или кусочно линейной), если она состоит из конечного числа прямых, то есть если для некоторых точек $a = t_0^{(N)} < t_1^{(N)} < \dots < t_N^{(N)} = b$ она является прямой на каждом из сегментов $[t_{k-1}, t_k]$ для $k = 1, \dots, N$. Длина прямой есть расстояние между ее крайними точками. Длина кусочно-линейной есть сумма длин своих прямых кусков. Кривая называется **гладкой**, если все функции $x_k(t)$ гладкие. Длина такой кривой определяется как предел длин ломаных, определенных функциями $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_d^{(N)}(t))$, если при $N \rightarrow \infty$ и $\max_k |t_k^{(N)} - t_{k-1}^{(N)}| \rightarrow 0$ равномерно по t и k будет

$$|x_k(t) - x_k^{(N)}(t)| \rightarrow 0$$

Рассмотрим сначала длину L кривой в R^2 , определенной функцией $f(x)$ вдоль вещественной оси x . Пусть $f(x)$ определена на интервале $[a, b]$ и имеет непрерывную производную $f'(x)$. В нашем случае разделим сегмент $[a, b]$ на интервалы длины $\Delta x = \frac{b-a}{N}$. Тогда длина L_N ломаной, определенной точками $(a + \frac{k}{N}(b-a), f(a + \frac{k}{N}(b-a)))$, $k = 0, 1, \dots, N$ будет

$$L_N = \sum_k \sqrt{(\Delta x)^2 + (\Delta f)^2} = \sum \sqrt{(\Delta x)^2 + (\Delta f)^2} \frac{\Delta x}{\Delta x} = \sum \sqrt{1 + \frac{(\Delta f)^2}{(\Delta x)^2}} \Delta x \rightarrow$$

$$\rightarrow L = \int_a^b \sqrt{1 + \left(\frac{df}{dx}\right)^2} dx$$

Пусть теперь кривая задана на $R^2 = (x_1, x_2)$ двумя функциями $x_1(t), x_2(t)$, где параметр t меняется на интервале $0 \leq t \leq s$. Тогда длина определяется как предел длин ломаных, аппроксимирующих нашу кривую. И тогда аналогично

$$L = \int_0^s \sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2} dt$$

Обобщение на размерности $d > 2$ очевидно.

4.5 Кратные интегралы и меры

Алгебра множеств и меры Рассмотрим произвольное множество Ω с элементами $\omega \in \Omega$ и некоторое множество Σ его подмножеств. Последнее называется **алгеброй множеств**, если $\Omega \in \Sigma$ и для любых подмножеств $A, B \in \Sigma$ их объединение, пересечение и разность также принадлежат Σ . **Мера** обычно определяется, для заданных (Ω, Σ) , как вещественная (возможно равная $+\infty$), функция $\mu(A)$, $A \in \Sigma$, на Σ . такая что выполняется свойство **конечной аддитивности**: $\mu(\cup A_k) = \sum \mu(A_k)$ для конечного набора не пересекающихся множеств $A_k \in \Sigma$. Приведем примеры:

1. Пусть Ω конечно. В качестве Σ возьмем множество всех подмножеств. Рассмотрим любую функцию $\mu(\omega)$ на Ω и определим меру множества $A \subset \Omega$ как $\mu(A) = \sum_{\omega \in A} \mu(\omega)$.

2. Пусть $\Omega = R$. В качестве Σ возьмем минимальную алгебру, содержащую все точки и все интервалы $(a, b) \subset \Omega$. Мера их будет соответственно 0 и $b - a$. В Σ войдут также все сегменты, полуинтервалы и их конечные объединения. Нетрудно доказать, что эта мера единственным образом продолжается на все Σ . Эта мера называется **мерой Лебега** на Ω .

3. Пусть $\Omega = R^2$. Открытое множество O в R^d называется **связным**, если для любых двух его точек x, y существует непрерывная кривая (например, ломаная), лежащая в O и соединяющая эти точки. В качестве Σ возьмем минимальную алгебру, содержащую все принадлежащие Ω точки, отрезки и все связные открытые множества, граница которых является замкнутой ломаной линией. Мера их определим как соответственно 0 и площадь в последнем случае. Называется также мерой Лебега на R^2 .

Определим интеграл от непрерывной функции $f(x, y)$ по ограниченному открытому множеству $D \in \Sigma$. Рассмотрим разбиение Q_ϵ области D на конечное число множеств

$A_k \in \Sigma$ с диаметром меньшим $\epsilon > 0$ (доказать что такое разбиение существует) и мерой $\mu_k = \mu(A_k)$. Определим **интегральную сумму**

$$S(\mathbb{Q}_\epsilon) = \sum_k f(x_k, y_k)\mu(A_k),$$

где $(x_k, y_k) \in A_k$ для всех k . Тогда для любой последовательности $\epsilon_n \rightarrow 0$, любых определенных выше разбиений \mathbb{Q}_{ϵ_n} и любого выбора точек $(x_k, y_k) \in A_k$, последовательность $S(\mathbb{Q}_{\epsilon_n})$ является последовательностью Коши (доказательство существования предела такое же как для одномерного интеграла), а значит имеет предел

$$\int \int_D f(x, y) dx dy = \lim_{n \rightarrow \infty} S(\mathbb{Q}_{\epsilon_n}),$$

который и называется интегралом функции f по мере μ по области D .

Теперь важнейший вопрос - как такие двумерные интегралы оценивать или вычислять. Рассмотрим например прямоугольник $D = \{a_1 \leq x \leq b_1, a_2 \leq y \leq b_2\}$ на плоскости и непрерывную функцию $f(x, y)$ на нем. В качестве A_k возьмем прямоугольники со сторонами параллельными координатным осям и имеющими длины $\frac{L_1}{N_1}, \frac{L_2}{N_2}$, где $L_i = b_i - a_i$, а N_1, N_2 - большие целые числа. Тогда интегралом будет предел при $N_1, N_2 \rightarrow \infty$ суммы

$$\sum_{k=0}^{N_1-1} \sum_{m=0}^{N_2-1} f\left(a_1 + \frac{kL_1}{N_1}, a_2 + \frac{mL_2}{N_2}\right) \frac{1}{N_1} \frac{1}{N_2}$$

Можно доказать, что двойной предел $N_1, N_2 \rightarrow \infty$ можно понимать как угодно. Например, как $N_1 = N_2 \rightarrow \infty$, или можно фиксировать большое N_2 и взять намного большее N_1 , то есть разбить D на очень узкие прямоугольники. Более того, для фиксированных N_2, m можно перейти к пределу $N_1 \rightarrow \infty$. Можно сказать, что мы вычисляем одномерные интегралы по x от функции $f(x, a_2 + \frac{mL_2}{N_2})$, или даже интеграл от $f(x, y)$ для фиксированного $y \in [a_2, b_2]$. Пусть результатом будет функция $F(a_2 + \frac{mL_2}{N_2})$ или $F(y)$, а потом мы вычислим интеграл по y . Получаем

$$\int \int_D f(x, y) dx dy = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x, y) dx \right) dy$$

Этот метод называется **повторным интегрированием**. Он исключительно важен, так как для вычисления интеграла позволяет использовать (часто более простое) вычисление производных.

Мы показали, как простые предельные переходы позволяют ориентироваться в основных определениях анализа. Везде выше множества были просты, а наши разбиения и суммы были конечны. Однако, математики углубились дальше. Алгебра множеств называется **σ -алгеброй**, если Σ замкнуто относительно счетных объединений и пересечений. В последнем случае пара (Ω, Σ) называется **измеримым пространством**.

Структура σ -алгебр довольно сложна - ее элементы (называемые **измеримыми множествами**) образуют довольно любопытную иерархию, которая изучалась в науке под названием "Аналитическая теория множеств". Определяются они обычно так: сначала определяется некоторый набор подмножеств $\Sigma_0 \subset \Omega$, называемый **базой**. В качестве базы обычно выбираются самые простые множества. Конечно,

всегда есть минимальная σ -алгебра, обозначаемая Σ , содержащая Σ_0 . Она конечно единственна - пересечение всех σ -алгебр, содержащих Σ_0 . Если Σ_0 - класс всех открытых множеств, то Σ называется **Борелевской σ -алгеброй**.

Мера обычно определяется, для заданного измеримого пространства (Ω, Σ) , как вещественная функция $\mu(A)$, $A \in \Sigma$, на Σ . если выполняется свойство σ -аддитивности (**счетной аддитивности**): $\mu(\cup A_k) = \sum \mu(A_k)$ для конечного или счетного набора не пересекающихся множеств $A_k \in \Sigma$. Здесь конечно надо избежать случая, когда существуют два подмножества с бесконечными мерами разных знаков, иначе мы не знаем чему равна мера их объединения.

Перечислим основные утверждения:

1. Мера Лебега μ на R^d продолжается единственным образом до σ -аддитивной меры на Σ со следующим свойством:

$$\mu(\cap A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$$

для любой убывающей последовательности $A_1 \supset A_2 \supset \dots$ множеств из Σ .

??

Но вся σ -алгебра совсем не первостепенна для нас, так как в математической физике как правило рассматриваются только достаточно простые множества. Более того, не исключено, что само евклидово пространство есть только хорошее приближение для чего-то более фундаментального.

5 Степенные ряды

Формальный степенной ряд есть следующее формальное выражение

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{12}$$

Простейший пример такого ряда - геометрическая прогрессия с $a_n \equiv 1$. Ниже, мы в основном рассматриваем случаи, когда a_n и z комплексные числа.

Говорят, что ряд (12) сходится в некоторой окрестности нуля, если для некоторого $z_0 > 0$ абсолютно сходится ряд (12) с $z = z_0$. Тогда ряд сходится для всех z с $|z| < z_0$. Например, геометрическая прогрессия сходится для всех z с $|z| < 1$ к $\frac{1}{1-z}$.

Заметим, что формальное дифференцирование n раз ряда (12) в точке $z = 0$ даст

$$a_n = \frac{f^{(n)}(0)}{n!}$$

Обратно, если мы сможем вычислить или оценить модули производных функции $f(z)$ в точке $z = 0$, можно будет проверить сходимость ее **ряда Тейлора**

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n,$$

и, в случае его сходимости в некоторой окрестности, мы получим то, что называют **разложением функции f в ряд в точке 0**.

Рассмотрим теперь комплекснозначную функцию $f(z)$, определенную на некотором открытом множестве A комплексной плоскости. Если для точки $z_0 \in A$ существует предел отношения

$$\frac{f(z) - f(z_0)}{z - z_0}$$

при $z \rightarrow z_0$ (точнее, при $|z - z_0| \rightarrow 0$), то он называется комплексной производной и обозначается $f'(z_0) = \frac{df}{dz}(z_0)$.

Замечание 5 Функцию $f(z)$ комплексного переменного z можно рассматривать как функцию $f(x+iy)$ двух вещественных переменных x, y с комплексными значениями. Частные производные по x и y определяются как

$$\frac{\partial f}{\partial x} = \lim_{x \rightarrow 0} \frac{f(z_0 + x) - f(z_0)}{x}, \quad \frac{\partial f}{\partial y} = \lim_{y \rightarrow 0} \frac{f(z_0 + iy) - f(z_0)}{y}$$

Если комплексная производная $f(z)$ существует, то

$$\frac{\partial f}{\partial x} = \frac{df}{dz}, \quad \frac{\partial f}{\partial y} = i \frac{df}{dz},$$

откуда

$$\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0$$

Комплексная функция $f(z)$, определенная в некоторой окрестности точки z_0 , называется **аналитической** в точке z_0 , если она совпадает с некоторым степенным рядом по $z - z_0$, сходящимся при $|z - z_0| < \epsilon$ для некоторого $\epsilon > 0$. Оказывается, что тогда $f(z)$ аналитична в любой точке z этой окрестности. Но есть и более сильное (общее) утверждение.

Теорема 1 Если $f(z)$ комплексно дифференцируема в каждой точке некоторого открытого множества A , то она аналитична в каждой точке z множества A , и более того, соответствующий степенной ряд сходится в любой окрестности $O(z)$ такой что $O(z) \subset A$.

Это утверждение будет доказано позже.

5.1 Основные функции

Рациональные функции Это функции вида $\frac{P(z)}{Q(z)}$, где $P = p_0 + p_1z + \dots + p_mz^m$ и $Q = q_0 + q_1z + \dots + q_nz^n$ многочлены. Если $Q(0) \neq 0$, то есть $q_0 \neq 0$, то Q^{-1} (а значит и $\frac{P}{Q}$) разлагаются в степенной ряд по z в некоторой окрестности точки 0

$$Q^{-1} = \frac{1}{q_0} \frac{1}{1 - y} = \frac{1}{q_0} \sum_{k=0}^{\infty} y^k, \quad y = -\left(\frac{q_1}{q_0}z + \dots + \frac{q_n}{q_0}z^n\right),$$

который сходится при $|y| < 1$, а значит и для всех достаточно малых $|z|$. Аналогично, если $Q(z_0) \neq 0$ для некоторой точки z_0 , то Q^{-1} может быть разложено в степенной ряд по $z - z_0$ в некоторой окрестности z_0 .

Замечание 6 Заметим, что любой многочлен Q может быть представлен (как при $z_1 = 0$) в виде $Q = (z - z_1)Q_{n-1} + C$, где Q_{n-1} - некоторый многочлен степени $n - 1$, а C - константа. Если же $Q(z_1) = 0$, то Q делится на $z - z_1$, то есть $C = 0$. Тогда по индукции следует (предполагая что каждый многочлен имеет хотя бы один корень), что каждый Q может быть представлен в виде

$$Q = C(z - q_1)\dots(z - q_n) \quad (13)$$

для некоторых комплексных чисел C, q_1, \dots, q_n , где q_k не обязательно различны. Также следует, что Q не может иметь более чем n корней. Ниже мы докажем, что Q имеет ровно n (возможно кратных) корней, то есть представим в виде (13) (основная теорема алгебры).

Число e Есть два эквивалентных определения числа e :

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \iff e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \quad (14)$$

Сходимость ряда доказывается мажорированием геометрической прогрессией

$$1 + \sum_{n=0}^{\infty} \frac{1}{2^n} = 3$$

Для доказательства (14) надо оценить сверху модуль разности

$$\sum_{n=0}^N \frac{1}{n!} - \left(1 + \frac{1}{N}\right)^N$$

для чего надо выполнить некоторые (\pm) -сокращения.

Функция e^z Полностью аналогично можно определить функцию $z(t) = e^t$

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} \iff e^t = \lim_{n \rightarrow \infty} \left(1 + \frac{t}{n}\right)^n$$

или как функцию $z = z(t)$, удовлетворяющую уравнению

$$\frac{dz}{dt} = z,$$

то есть как (единственную с точностью до умножения на константу) функцию, которая совпадает со своей производной, а значит и со всеми своими производными.

Основное свойство

$$e^{x+y} = e^x e^y \quad (15)$$

может быть доказано перемножением рядов справа и использованием очевидных комбинаторных формул

$$x^m y^n \frac{1}{(m+n)!} C_{m+n}^m = x^m y^n \frac{1}{m!n!}$$

В частности, $e^z e^{-z} = 1$, откуда следует, что e^z нигде не равна нулю, а также $e^{-z} = \frac{1}{e^z}$. Кроме того, e^t возрастает на $[0, \infty)$ от 1 до ∞ , а на $[0, -\infty)$ убывает от 1 до 0. Таким образом, она принимает все положительные значения. Ниже мы покажем, что она принимает все комплексные значения, кроме нуля..

Тригонометрические функции Функции $\cos x, \sin x$ можно ввести как, соответственно, вещественную и мнимую части ряда для e^{ix}

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!},$$

Отсюда следует, что они вещественны для вещественных x , а дифференцирование рядов дает

$$(\sin x)' = \cos x, (\cos x)' = -\sin x \quad (16)$$

Надо теперь проверить формулу

$$\sin^2 x + \cos^2 x = 1, \quad (17)$$

из которой следует ограниченность этих функций для вещественных x . Для доказательства надо возвести оба ряда в квадрат, заметить что в полученном ряде не будет членов с нечетными степенями и написать формулы коэффициентов перед x^{2n} и показать что они равны нулю для всех $n > 0$.

Формула (17) показывает, что любое вещественное φ определяет точку $(x, y) = (\cos \varphi, \sin \varphi)$ на единичной окружности S на плоскости R^2 , и формально можно считать это число φ углом. Например, угол для точки $(1, 0)$ равен нулю - обычно угол измеряется от этой точки против часовой стрелки. Покажем, что угол равен длине дуги окружности от точки $(1, 0)$ до точки (x, y) . Действительно, длина кривой $y = f(x)$ (то есть длина $\varphi(x)$ части единичной окружности над интервалом $[0, x] \subset R$, в положительной четверти плоскости) равна

$$\varphi(x) = \int_0^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_0^x \frac{1}{\sqrt{1-x^2}} dx = \arcsin x, |x| < 1,$$

так как можно вычислить производную арксинуса (функции обратной к синусу на интервале $(-\frac{\pi}{2}, \frac{\pi}{2})$) как производную обратной функции (см. выше)

$$\frac{d(\arcsin y)}{dy} = \frac{1}{\cos(\arcsin y)} = \frac{1}{\sqrt{1-\sin^2(\arcsin y)}} = \frac{1}{\sqrt{1-y^2}} \quad (18)$$

Тогда первая производная

$$\frac{d\varphi}{dx} = \frac{1}{\sqrt{1-x^2}}, |x| < 1,$$

а отсюда и все остальные производные выглядят просто и можно написать ряд для $\varphi(x)$, то есть ряд для решений уравнения $\cos \varphi = x$ или ряд для $\arccos \varphi$.

Заметим инвариантность длин относительно группы вращений круга.

Число π Есть разные определения:

1) длина окружности единичного радиуса

$$x^2 + y^2 = 1 \iff y = \sqrt{1-x^2}$$

равна 2π . Или можно определить число π как интеграл

$$\frac{\pi}{2} = \int_0^1 \frac{1}{\sqrt{1-x^2}} dx \quad (19)$$

2) четверть площади единичного круга

$$\frac{\pi}{4} = S = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \sqrt{1 - \frac{n^2}{N^2}} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{n=1}^N \sqrt{N^2 - n^2},$$

откуда получаем ряд Грегори

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

3) другое определение, см. (38), известное любому вероятностнику,

$$\sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$$

Периодичность Докажем периодичность тригонометрических функций. Заметим, что из вида их рядов следует

$$\sin 0 = 0, \cos 0 = 1$$

а из (18) и (19) следует что

$$\sin \frac{\pi}{2} = 1, \cos \frac{\pi}{2} = 0, \dots, \cos 2\pi = 1, \sin 2\pi = 0$$

Тогда

$$e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1,$$

и мы получаем периодичность экспоненты

$$e^{2\pi i + x} = e^{2\pi i} e^x = e^x$$

Так же получаем, что функция $f(x) = e^{x+i\pi}$, $x \in R$ принимает все отрицательные значения. Аналогично, рассматривая все прямые на плоскости, проходящие через ноль, получаем, что e^t принимает все значения кроме нуля..

Логарифм и степень Функция $\ln x$ определяется сначала для $x > 0$ как функция обратная к e^x . Тогда, из (15) имеем

$$\ln(xy) = \ln x + \ln y$$

и

$$x = e^{\ln x} \implies 1 = (e^{\ln x})' = e^{\ln x} (\ln x)' \implies (\ln x)' = \frac{1}{x}$$

Теперь легко найти все производные и ряд для логарифма в точке 1

$$\ln(1+x) = x + \sum_{k=2}^{\infty} \frac{x^k}{k} (-1)^{k-1}$$

В силу периодичности экспоненты, в любой точке $z \neq 0$ логарифм - многозначная функция. В частности на окружности $|z| = 1$ он принимает мнимые значения $\ln e^{i\phi} = i(\phi + 2\pi k)$ с целыми k .

Для любого комплексного $x \neq 0$, его степень определяется как $x^a = e^{a \ln x}$. Отсюда легко вывести все свойства степени, и в частности,

$$(x^a)' = (e^{a \ln x})' = x^a \frac{a}{x} = ax^{a-1}$$

5.2 Аналитические функции

Комплексный интеграл вдоль кривой Пусть даны связное открытое множество $\Lambda \subset C$, комплексно дифференцируемая функция $f(z)$ на Λ и кривая $\Gamma : z(t), t \in [0, T]$, в Λ , где $z(t)$ дифференцируема по t . Выберем $N + 1$ точек $z_k = z(k\frac{T}{N})$ на этой кривой, обозначим $z_N = z(T) = z$ и

$$F_N(z) = \sum_{k=1}^N f(z_k)(z_k - z_{k-1}) \quad (20)$$

Предел $F(z) = \lim_{N \rightarrow \infty} F_N(z)$ существует, что доказывается так же как для вещественного интеграла по прямой. Например, если $f(z) = C = const$, то уже до взятия предела видно, что

$$F(z) = F_N(z) = C(z_N - z_0)$$

Также и факт, что $F(z)$ есть анти-производная $f(z)$ "в комплексном смысле", доказывается в точности, как в (11), так как производная "в комплексном смысле" существует и одинакова в любом направлении (включая направление кривой Γ в точке z). Но в то же время анти-производная определена с точностью до константы C . Так что для контура Γ (из точки z_0 до точки z) для любой однозначной анти-производной $F(z)$ функции $f(z)$

$$\int_{\Gamma} f(z) dz = F(z) - F(z_0),$$

то есть $C = -F(z_0)$, так как при $z \rightarrow z_0$ интеграл стремится к нулю.

Например, при $n \neq -1$, анти-производная функции z^n есть функция $\frac{1}{n+1}z^{n+1} + const$. Тогда для любой кривой (любого контура) Γ с начальной и конечной точками z_0 и z соответственно

$$\int_{\Gamma} z^n dz = \frac{1}{n+1}z^{n+1} - \frac{1}{n+1}z_0^{n+1}$$

Если контур замкнут, то есть $z = z_0$, то интеграл равен нулю. Вместо полинома мы могли бы взять любую функцию $f(z)$, однозначную и дифференцируемую в любой точке односвязной области, и тогда для любой замкнутой кривой Γ

$$\int_{\Gamma} f(z) dz = 0,$$

где **односвязность** области Λ означает что Λ само и его дополнение связны.

В то же время, если $n = -1$, и Γ - любая окружность с центром в точке 0, то

$$\int_{\Gamma} \frac{1}{z} dz = \ln(ze^{2\pi i}) - \ln z = 2\pi i,$$

и анти-производной для $\frac{1}{z}$ будет $\ln z$. Мы договариваемся, что далее такие интегралы берутся против часовой стрелки.

Пусть $f(z)$ - аналитична в точке ноль. Обозначим $\Gamma(r) = \{z : |z| = r\}$ окружность радиуса r . Тогда для достаточно малых r

$$f(0) = \frac{1}{2\pi i} \int_{\Gamma(r)} \frac{f(z)}{z} dz, \quad (21)$$

что следует из предыдущих формул и почленного интегрирования ряда. Последняя формула называется **формулой Коши**.

Другой пример. Пусть в малой окрестности точки z_0 функция представима в виде ($c \ m \leq -1$)

$$f(z) = \sum_{k=m}^{\infty} a_k (z - z_0)^k \quad (22)$$

Тогда a_{-1} называется **вычетом** в точке z_0 , и если $\Gamma(r)$ - окружность достаточно малого радиуса r вокруг точки z_0 , то

$$\frac{1}{2\pi i} \int_{\Gamma(r)} f(z) dz = a_{-1}$$

Аналитичность и дифференцируемость

Утверждение 1 *Предположим что функция f комплексно дифференцируема в окрестности нуля некоторого радиуса $0 < R < \infty$. Тогда:*

- 1) *формула Коши в точке $z = 0$ имеет место для всех $r < R$,*
- 2) *функция f имеет все производные в точке 0 , и ее ряд Тейлора сходится для всех $z : |z| < R$.*

Для доказательства нам понадобится следующая часто используемая конструкция. Фиксируем два числа $0 < r_1 < r_2 < R$. Определим два замкнутых контура на C : 1) $\Gamma_+ = \Gamma_+(r_1, r_2)$ - начинается в точке $(r_1, 0)$, идет вдоль вещественной оси к точке $(r_2, 0)$, потом по окружности радиуса r_2 против часовой стрелки к точке $(-r_2, 0)$, затем по вещественной оси к точке $(-r_1, 0)$, и наконец по окружности радиуса r_1 по часовой стрелке возвращается в точку $(r_1, 0)$. 2) симметричный (относительно вещественной оси) контур $\Gamma_- = \{(r_2, 0) \rightarrow (r_1, 0) \rightarrow (-r_1, 0) \rightarrow (-r_2, 0) \rightarrow (r_2, 0)\}$. Мы хотим доказать что

$$\int_{\Gamma(r_1)} = \int_{\Gamma(r_2)}$$

Действительно, интегралы по контурам Γ_+ и Γ_- равны нулю (внутри нет сингулярных точек), и значит $0 = \int_{\Gamma_+} + \int_{\Gamma_-} = -\int_{\Gamma(r_1)} + \int_{\Gamma(r_2)}$, так как в левой сумме интегралы вдоль обеих интервалов вещественной оси взаимно сокращаются.

Чтобы доказать 1) заметим, что когда $|z| = r_1 \rightarrow 0$ можно написать

$$f(z) = f(0) + f'(0)z + o(r_1)$$

Тогда интеграл Коши от второго члена правой части даст 0, а третий будет стремиться к нулю. Так что в пределе останется только $f(0)$. Так как r_1 произвольно, то интеграл по окружности радиуса r_2 равен этому пределу.

2) Предположим теперь что z_0 произвольная точка, f дифференцируема в этой точке и Γ - окружность достаточно малого радиуса $r < r$ вокруг z_0 . Тогда из 1) и (21) следует, что

$$f(z_0) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{f(z)}{z - z_0} dz$$

Дифференцируя под знаком интеграла, получим

$$f'(z_0) = \lim_{\epsilon \rightarrow 0} \frac{f'(z_0 + \epsilon) - f'(z_0)}{\epsilon} = \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} \int_{\Gamma} \frac{1}{\epsilon} \left(\frac{1}{z - z_0 - \epsilon} - \frac{1}{z - z_0} \right) f(z) dz =$$

$$= \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} \int_{\Gamma} \frac{1}{(z - z_0 - \epsilon)(z - z_0)} f(z) dz = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^2} dz$$

и аналогично

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_{\Gamma(r)} \frac{f(z)}{(z - z_0)^{n+1}} dz$$

Так, если например $z_0 = 0$, мы имеем оценку

$$|f^{(n)}(0)| \leq \frac{n!}{r^{n+1}} \max_{|z|=r} |f(z)|$$

Теперь, для произвольной z с $|z| < R$ выберем r так чтобы $|z| < r < R$. Таким образом, ряд Тейлора для $f(z)$ в точке 0 абсолютно сходится при $|z| < r$.

Теорема о вычетах: предположим что в односвязной области Λ функция аналитична везде кроме конечного числа сингулярных точек, в которых она имеет форму (22). Тогда интеграл вдоль любого замкнутого контура без самопересечений равен сумме вычетов в сингулярных точках внутри этого контура.

Она доказывается аналогично.

Заметим, что такие функции называются **мероморфными** в Λ , а функции (как например e^z), аналитические в каждой точке комплексной плоскости называются **целыми** функциями.

Основная теорема алгебры Рассмотрим многочлен $P(z) = z^n + p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$, и функцию $f(z) = \frac{P'(z)}{P(z)}$. Например, если $P = z^n, n = 0, 1, \dots$, то для любого $r > 0$

$$\frac{1}{2\pi i} \int_{\Gamma(r)} f(z) dz = n, \quad (23)$$

то есть равен кратности корня $z = 0$. Пусть z_0 будет одним из корней P . Тогда такой же интеграл по окружности малого радиуса вокруг z_0 даст кратность этого корня.

Из за доминантности члена z^n для достаточно больших R , все корни будут внутри $\Gamma(R)$ для некоторого R . Но

$$\begin{aligned} f(z) &= \frac{P'(z)}{P(z)} = \frac{nz^{n-1} + p'(z)}{z^n + p(z)} = \frac{nz^{-1} + \frac{p'(z)}{z^n}}{1 + \frac{p(z)}{z^n}} = nz^{-1} \frac{1 + \frac{p'(z)}{nz^{n-1}}}{1 + \frac{p(z)}{z^n}} = \\ &= nz^{-1} \frac{1 + \sum_{k=1}^{\infty} a_k z^{-k}}{1 + \sum_{k=1}^{\infty} b_k z^{-k}} = \frac{n}{z} + \sum_{k=2}^{\infty} c_k z^{-k} \end{aligned}$$

Последний ряд сходится для достаточно больших R . Значит интеграл от него равен нулю. А значит и формула (23) справедлива.

Но кроме того, интеграл от $f(z)$ по $\Gamma(R)$ равен сумме интегралов от $f(z)$ вдоль маленьких окружностей вокруг всех нулей P . Это может быть доказано, как и выше, построением малых окружностей вокруг каждого корня и контуров, связывающих их вместе и с $\Gamma(R)$. Так что сумма кратностей всех корней $P(z)$ равна n .

6 Асимптотики сумм, произведений и интегралов

Каждый из этих методов получения асимптотики основан на простых очевидных идеях: 1) аппроксимация сумм интегралами, 2) сокращения при частых сменах знака, 3) доминантность окрестности экспоненциально большого максимума, позволяющая пренебречь вкладом остальной части, 4) выбор пути интегрирования.

Суммы

Задача 9 Какой самый простой способ доказательства расходимости ряда $\sum_{n=1}^{\infty} \frac{1}{n}$?

Для получения асимптотики докажем сначала, что

$$\sum_{k=1}^n \frac{1}{k} = \ln n + C + o(1), \quad (24)$$

где $C > 0$ некоторая константа. Растущие члены в асимптотике (в отличие от констант) в таких задачах можно найти легче, так как они не зависят от конечного числа членов ряда. Рассмотрим кусочно-постоянную функцию $f(x)$, равную $\frac{1}{k}$ на интервале $[k, k+1)$. Тогда наша сумма равна площади под ней на интервале $[1, n+1)$. Так как разность $f(x) - \frac{1}{x} \geq 0$, то

$$\sum_{k=1}^n \frac{1}{k} > \int_1^{n+1} \frac{dx}{x} = \ln(n+1) = \ln n + \ln\left(1 + \frac{1}{n}\right) = \ln n + O\left(\frac{1}{n}\right),$$

а площадь между этими двумя функциями равна

$$\sum_{k=1}^n \left(\frac{1}{k} - \int_k^{k+1} \frac{dx}{x} \right) = C > 0,$$

что дает сходящийся ряд из положительных членов.

Произведения - формула Стирлинга Аналогично, в асимптотике произведений найти мультипликативную константу существенно сложнее чем возрастающие множители. Примером может быть знаменитая формула Стирлинга

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \quad (25)$$

Все сводится к нахождению асимптотического разложения суммы

$$\ln n! = \ln 1 + \dots + \ln n = n \ln n - n + \frac{1}{2} \ln n + \frac{1}{2} \ln(2\pi) + C + o(1), \quad (26)$$

где все просто, кроме нахождения константы. Существует много связей между двумя знаменитыми числами: π и e . Например, даваемая формулой $e^{2\pi i} = 1$, а также сама формула Стирлинга (25). Повидимому трудно все эти связи объединить в одну простую структуру.

Рассмотрим сумму (26) как площадь S_n под графиком кусочно-постоянной функции $f(x)$ на $(1, n]$, равной $\ln k$ на полуинтервале $(k-1, k]$. Площадь под функцией $\ln x$ на этом интервале

$$\int_1^n \ln x dx = - \int_1^n x d(\ln x) + (x \ln x)|_1^n = -n + 1 + n \ln n$$

дает оценку снизу и основной множитель $(\frac{n}{e})^n$. Множитель \sqrt{n} получить немного сложнее из-за асимптотики разности двух площадей

$$\begin{aligned} S_n - \int_1^n \ln x dx &= \sum_{k=2}^n [\ln k - \int_{k-1}^k \ln x dx] = \\ &= \sum_{k=2}^n [\ln k - (k \ln k - (k-1) \ln(k-1) - 1)] = \frac{1}{2} \ln n + C + o(1), \end{aligned} \quad (27)$$

которая подсчитывается после разложения (для больших k)

$$\begin{aligned} \ln(k-1) &= \ln k + \ln(1 - \frac{1}{k}) = \ln k - \frac{1}{k} - \frac{1}{2k^2} + O(\frac{1}{k^3}) \ln(k-1) = \\ &= \ln k + \ln(1 - \frac{1}{k}) = \ln k - \frac{1}{k} - \frac{1}{2k^2} + O(\frac{1}{k^3}), \end{aligned}$$

и использования последнего равенства в (27) и асимптотики сумм (24).

Интегрирование по частям Типичный пример - следующая оценка интеграла функции f , дифференцируемой на интервале $[a, b]$: существует константа $A > 0$ такая что

$$| \int_a^b f(x) \cos(Nx) dx | \leq \frac{A}{N} \quad (28)$$

Интуитивно это получается благодаря быстрым флуктуациям косинуса на коротких интервалах, на которых f почти постоянна, что дает плюс-минус сокращения. Формально это выглядит так:

$$\begin{aligned} f(x) \cos(Nx) &= \frac{1}{N} f(x) (\sin(Nx))' = \frac{1}{N} [(f(x) \sin(Nx))' - f'(x) \sin(Nx)] \implies \\ \implies \int_a^b f(x) \cos Nx dx &= \frac{1}{N} [f(b) \sin(Nb) - f(a) \sin(Na) - \int_a^b f'(x) \sin(Nx) dx] \end{aligned}$$

Если $f(b) \sin(Nb) = f(a) \sin(Na)$, то предположив дважды дифференцируемость f , мы можем применить похожую процедуру к последнему интегралу чтобы получить оценку порядка N^{-2} , и т.д.

Метод Лапласа Рассмотрим интеграл

$$\int_a^b f(x) e^{NS(x)} dx,$$

где S имеет единственный невырожденный максимум, то есть $S(x_0) > 0, S'(x_0) = 0, S''(x_0) < 0$, во внутренней точке $x_0 \in [a, b]$. Сначала пусть $f = 1$, тогда вынесем $e^{NS(x_0)}$ за знак интеграла

$$\int_a^b e^{NS(x)} dx = e^{NS(x_0)} \int_a^b e^{N(S(x)-S(x_0))} dx$$

Мы покажем, что в интеграле справа в асимптотике можно пренебречь интегралом вне интервала

$$I_0 = (x_0 - cN^{-\frac{1}{2}}, x_0 + cN^{-\frac{1}{2}})$$

для любой константы c . Действительно, в окрестности x_0 , имеем

$$S(x) - S(x_0) = \frac{1}{2}S''(x_0)(x - x_0)^2 + O((x - x_0)^3)$$

Для достаточно малых $\delta > 0$ введем еще меньший интервал $I_\delta = (x_0 - cN^{-\frac{1}{2}-\delta}, x_0 + cN^{-\frac{1}{2}-\delta})$, где

$$\int_{I_0} > \int_{I_\delta} \geq c_1 N^{-\frac{1}{2}-\delta} e^{N^\delta}$$

В то же время, подинтегральное выражение вне I_0 не превышает константы. Поэтому

$$\int_{I_0} e^{N(S(x)-S(x_0))} dx \sim \int_{I_0} e^{N\frac{1}{2}S''(x_0)(x-x_0)^2} dx \sim \int_{-\infty}^{\infty} e^{N\frac{1}{2}S''(x_0)(x-x_0)^2} dx = \sqrt{\frac{2\pi}{NS''(x_0)}}$$

Последний интеграл называется гауссовым и вычисляется в (38).

При вычислении различных криволинейных интегралов, часто можно деформировать путь (не изменяя значения интеграла) так чтобы он проходил через точку с явно выраженным максимумом. В этом случае говорят о **методе перевала**.

Метод стационарной фазы Обобщением интегралов типа (28) являются интегралы вида

$$\int f(x) \exp(iNS(x)) dx$$

Однако здесь есть важный нюанс - концы интервала также могут давать существенный вклад в асимптотику. Поэтому предположим, что $f(x)$ гладкая и финитная (то есть равна нулю вне некоторого конечного интервала). Если $S(x)$ не имеет стационарных точек, то есть точек где $S'(x) = 0$, то интеграл оценивается интегрированием по частям, как выше. Существование таких точек препятствует сокращениям благодаря переменам знака.

Тогда в окрестности стационарной точки интеграл трактуется как в методе Лапласа, а вне ее - интегрированием по частям. Предположим что в некоторой точке a будет $S'(a) = 0$, а также, например, $S''(a) > 0, f(a) \neq 0$. Тогда, как в методе Лапласа, надо выбрать достаточно малое $\epsilon = \epsilon(N)$ так чтобы

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \exp(iNS(x)) dx &\sim \int_{a-\epsilon}^{a+\epsilon} f(x) \exp(iNS(x)) dx \sim \\ &\sim f(a) \exp(iNS(a)) \int_{a-\epsilon}^{a+\epsilon} \exp(iN(S(x) - S(a))) dx \sim \end{aligned}$$

$$\begin{aligned}
&\sim f(a) \exp(iNS(a)) \int_{a-\epsilon}^{a+\epsilon} \exp\left(\frac{iN}{2} S''(a)(x-a)^2\right) dx \sim \\
&\sim f(a) \exp(iNS(a)) \int_{-\infty}^{\infty} \exp\left(\frac{iN}{2} S''(a)(x-a)^2\right) dx \sim \\
&\sim C \frac{\exp(iNS(a))}{\sqrt{N}}, \quad C = f(a) \exp\left(\frac{i\pi}{4}\right) \sqrt{\frac{2\pi}{S''(a)}}
\end{aligned}$$

Таким образом, становится очевидной связь с методом Лапласа, а также что асимптотические разложения всегда требуют больших вычислений, хотя методы (алгоритмы) часто предсказуемы.

Part IV

Линейная алгебра

6.1 Линейные операторы и матрицы

Линейное пространство L (над полем C или R) есть множество элементов, называемых **векторами**, которые образуют коммутативную группу (по сложению) с нулевым элементом 0 , и для любых двух элементов x, y определена их сумма $x + y$. Более того, для любого вектора x определено умножение λx на любое (комплексное или соответственно вещественное) число λ , удовлетворяющее следующим аксиомам:

$$0x = 0, \lambda(x + y) = \lambda x + \lambda y, (\lambda_1 \lambda_2)x = \lambda_1(\lambda_2 x), (\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$$

Далее, в качестве основного поля у нас будет C , но многие определения могут быть использованы и для других полей.

Вектора x_1, \dots, x_N называются **линейно независимыми**, если не существует таких чисел λ_k (не всех равных нулю) что

$$\lambda_1 x_1 + \dots + \lambda_N x_N = 0$$

Максимальное число линейно независимых векторов в L называется **размерностью** L и обозначается $\dim L$. Тогда (если размерность равна N) любые N таких векторов могут быть выбраны в качестве **базиса**. Это означает, что любой вектор $y \in L$ может быть записан в виде линейной комбинации элементов этого базиса. Другими словами, существуют числа $a \neq 0, a_1, \dots, a_N$, такие что

$$ay + a_1 x_1 + \dots + a_N x_N = 0,$$

то есть вектора x_1, \dots, x_N, y линейно зависимы. И конечно, можно взять $a = 1$.

Два линейных пространства L_1, L_2 одной размерности (линейно) изоморфны, то есть существует взаимно однозначное отображение $\phi : L_1 \rightarrow L_2$, “уважающее” сложение и умножение на числа, то есть

$$\phi(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \phi(x_1) + \lambda_2 \phi(x_2)$$

Для доказательства надо только линейно продолжить произвольное взаимно-однозначное соответствие между любыми их базисами.

Простое представление (иногда говорят о **стандартном представлении**) линейного пространства размерности N есть множество векторов (последовательностей) $a = (a_1, \dots, a_N)$, длины N , где $a_k \in C$ называются **координатами** (или **компонентами**) вектора a . В этом представлении операции над двумя векторами a и $b = (b_1, \dots, b_N)$ выглядят так:

$$a + b = (a_1 + b_1, \dots, a_N + b_N), \lambda a = (\lambda a_1, \dots, \lambda a_N)$$

Удобный базис здесь $e_k, k = 1, \dots, N$, где координаты вектора e_k равны 0, кроме k -ой координаты, равной 1.

Тогда утверждение о линейной независимости векторов $x_k = (x_{k1}, \dots, x_{kN}), k = 1, \dots, N$, может быть переформулировано так: одно векторное уравнение, или эквивалентная система N линейных уравнений

$$\sum_k x_k \lambda_k = y \iff \sum_k x_{kl} \lambda_k = y_l, l = 1, 2, \dots, N,$$

с N неизвестными λ_k , имеет единственное решение $\{\lambda_k\}$ для любого вектора $y = (y_1, \dots, y_N)$.

Прямая сумма N линейных пространств $L_k, k = 1, \dots, N$, определяется как линейное пространство такое что: 1) как множество оно есть (декартово) прямое произведение множеств L_k , с элементами (x_1, \dots, x_N) , в котором $x_k \in L_k$, 2) операции сложения и умножения на числа определяются как

$$(x_1, \dots, x_N) + (y_1, \dots, y_N) = (x_1 + y_1, \dots, x_N + y_N), \quad \lambda(x_1, \dots, x_N) = (\lambda x_1, \dots, \lambda x_N)$$

В частности, любое линейное пространство размерности N есть прямая сумма N одномерных пространств.

Линейный функционал на L есть функция $f : L \rightarrow C$, удовлетворяющая свойствам линейности

$$f(x + y) = f(x) + f(y), \quad f(\lambda x) = \lambda f(x) \quad (29)$$

Очевидно, что в стандартном представлении любой линейный функционал может быть представлен в виде

$$f(x) = x_1 f(e_1) + \dots + x_N f(e_N) = f_1 x_1 + \dots + f_N x_N$$

для некоторых чисел $f_k = f(e_k)$.

Линейный оператор в C^N есть функция $A : L \rightarrow L$ имеющая свойства линейности в точности подобные (29). Множества линейных функционалов, линейных операторов (и матриц, см. ниже) в C^N также представляют собой линейные пространства размерностей N и N^2 соответственно. Множество линейных операторов образуют также алгебру с умножением $(L_2 L_1)x = L_2(L_1)x$. **Ядро** $\text{Ker} A$ оператора A , то есть множество векторов $x \in L$ таких что $Ax = 0$, является линейным подпространством L . Аналогично, **образ** $\text{Im} A$ также является линейным пространством, размерность которого называется **рангом** оператора A .

Любой линейный оператор может быть записан (в произвольном базисе e_1, \dots, e_N), с некоторыми числами $a_{kl}, k, l = 1, \dots, N$,

$$Ae_k = \sum_l a_{lk} e_l$$

Тогда для любого вектора $x = (x_1, \dots, x_N)$ в этом же базисе будет

$$Ax = \sum_k x_k Ae_k = \sum_k x_k \sum_l a_{lk} e_l = (\sum_{k=1}^N a_{1k} x_k, \dots, \sum_{k=1}^N a_{Nk} x_k)$$

Эти числа a_{ij} удобно представлять в матричном виде (как квадратная $(N \times N)$ -таблицу с N строками и N столбцами) $A = (a_{ij}), i, j = 1, \dots, N$, где i нумерует строки, а j - столбцы. Произведение двух операторов $AB = D$ соответствует (в этом базисе) умножению их матриц, определяемому как

$$d_{ij} = \sum_{k=1}^N a_{ik} b_{kj}$$

Более того, и умножение вектора на матрицу можно понимать как умножение двух матриц, если вектор записать в виде $(1 \times N)$ -матрицы, или как $(N \times 1)$ -матрицы, то есть

$$xA = (x_1, \dots, x_N)A = y = (y_1, \dots, y_N) = \left(\sum_i x_i a_{i1}, \dots, \sum_i x_i a_{iN} \right),$$

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \sum_j a_{1j} x_j \\ \vdots \\ \sum_j a_{Nj} x_j \end{pmatrix}$$

Отметим, что в первом случае (в xA) суммирование ведется по первым индексам a_{ij} и по вторым индексам в Ax .

Если в каждом из L_k в прямой сумме $L = L_1 \times \dots \times L_N$ задан некоторый линейный оператор $A_k : L_k \rightarrow L_k$, то прямой суммой этих операторов называется линейный оператор $A : L \rightarrow L$ действующий покомпонентно

$$Ax = A(x_1, \dots, x_N) = y = (y_1, \dots, y_N) = (A_1 x_1, \dots, A_N x_N)$$

Задача 10 Аналогично, определить линейные операторы из C^N в C^M и соответствующие им матрицы. Для $M = 1$ очевидно это будут линейные функционалы.

Мультилинейные формы Пусть заданы линейные пространства L_1, \dots, L_n размерностей N_k . Векторы этих пространств будем обозначать $x_k = (a_{k1}, \dots, a_{kN_k})$.

Функция $f : L_1 \times \dots \times L_n \rightarrow C$ называется мультилинейной формой, если для любого индекса k , как например для $k = 1$, имеет место:

1) при прибавлении вектора $y \in L_1$ к x_1 будет

$$f(x_1 + y, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) + f(y, x_2, \dots, x_n),$$

2) умножение x_1 на число λ умножает на него всю f :

$$f(\lambda x_1, x_2, \dots, x_n) = \lambda f(x_1, x_2, \dots, x_n)$$

Далее, предполагаем что все $L_k = L$ имеют размерность n , и будем записывать вектора x_i из L в стандартном базисе e_1, \dots, e_n :

$$x_i = \sum_{j=1}^n a_{ij} e_j$$

Тогда форму можно рассматривать как функцию от n^2 чисел a_{ij} .

Форма называется **кососимметрической**, если любая перестановка двух векторов изменяет знак формы. Например,

$$f(x_2, x_1, x_3, \dots, x_n) = -f(x_1, x_2, x_3, \dots, x_n)$$

Мы покажем теперь, что условие нормировки

$$f(e_1, e_2, \dots, e_n) = 1 \tag{30}$$

кососимметрической формы, рассматриваемой как функции от a_{ij} , делает ее единственной и дает ее представление

$$f(\{a_{ij}\}) = \sum_{\sigma} a_{1j_1} a_{2j_2} \dots a_{nj_n} (-1)^{|\sigma|}, \quad (31)$$

где сумма берется по всем перестановкам

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix},$$

где $|\sigma|$ четность σ . Это число (31) называется **детерминантом** (или **определителем**) $\det A$ матрицы $A = (a_{ij})$.

Для доказательства (31), достаточно сначала раскрыть все скобки, используя свойство 1), и вынести все a_{ij} из-под f , используя 2) :

$$f(x_1, \dots, x_n) = f\left(\sum_{j=1}^n a_{1j} e_j, \dots, \sum_{j=1}^n a_{nj} e_j\right) = \sum_{j_1, \dots, j_n=1}^n a_{1j_1} \dots a_{nj_n} f(e_{j_1}, \dots, e_{j_n})$$

После чего надо использовать условие нормировки и то что $f(e_{j_1}, \dots, e_{j_n}) = (-1)^{|\sigma|}$, если все e_{j_1}, \dots, e_{j_n} различны, и нулю в остальных случаях. Заметим, что (31) может быть переписано в виде

$$\sum_{\sigma} a_{j_1 1} a_{j_2 2} \dots a_{j_n n} (-1)^{|\sigma|},$$

откуда следует, что определители матрицы $A = (a_{ij})$ и ее **транспонированной** $A' = (a'_{ij} = a_{ji})$ равны.

Задача 11 Доказать прямым перемножением, что $\det(AB) = \det(A) \det(B)$.

1.1 Геометрия определителя

Скалярное произведение (a, b) двух векторов $a = (a_1, \dots, a_N), b = (b_1, \dots, b_N) \in R^N$ и **длина** $|a|$ вектора a определяются как

$$(a, b) = a_1 b_1 + \dots + a_N b_N, \dots, |a| = \sqrt{(a, a)} \geq 0$$

Косинус угла ϕ между векторами a и b определяется (как на плоскости) формулой

$$\cos \phi = \frac{(a, b)}{|a||b|}$$

Легко доказать, что это определение угла и косинуса не противоречит определениям в разделе о степенных рядах.

Определим n -мерные параллелепипеды $P_n \subset R^n$, порожденные n векторами a_1, \dots, a_n как

$$P_n = P(a_1, \dots, a_n) = \{s_1 a_1 + \dots + s_n a_n : 0 \leq s_1, \dots, s_n \leq 1\}$$

Мы хотим также определить их (ориентированные) объемы $v(P(a_1, \dots, a_n))$ индукцией по n . Пусть M_{n-1} - подпространство, порожденное векторами a_1, \dots, a_{n-1} . Тогда вектор a_n может быть записан в виде $a_n = h + a$, где h перпендикулярен к M_{n-1} , то

есть имеет нулевое скалярное произведение с каждым из векторов a_1, \dots, a_{n-1} , а a - линейная комбинация векторов a_1, \dots, a_{n-1} .

Это означает, что P_n является объединением прямых сегментов $\cup_{x \in P_{n-1}} [x, x + a_n]$. Тогда $|v(P_n)| = |v(P_{n-1})||h|$, как и в двумерном случае. Нам надо показать, что если $|v(P_{n-1})| = |\det A_{n-1}|$, то $|v(P_n)| = |\det A_n|$. Для этого определим матрицу $A = (a_{ij})$ с вектор-столбцами $a_i = (a_{i1}, \dots, a_{iN})$, Тогда матрица AA' может быть записана в виде **матрицы Грама** $G = G(a_1, \dots, a_N)$ системы векторов a_1, \dots, a_N с элементами

$$g_{ij} = (a_i, a_j)$$

Тогда

$$\det G = \det(AA') = (\det A)^2$$

Так как $(a_k, h) = 0$ при $k = 1, \dots, n - 1$, то

$$\det G(a_1, \dots, a_{n-1}, a_n) = \det G(a_1, \dots, a_{n-1}, h) = \det G(a_1, \dots, a_{n-1})(h, h)$$

Замечание 7 В этих рассуждениях мы опустили важный нюанс: объем может иметь знак, зависящий от ориентации - знак определителя определяет смену ориентации. Простой пример в двумерном случае: разность между $P_2(e_1, e_2)$ и $P_2(e_2, e_1)$ объясняет что такое ориентация.

Теперь можно пояснить геометрический смысл формулы

$$\det(A_1 A_2) = \det A_1 \det A_2$$

Действительно, определитель можно рассматривать не только как объем но также как число, показывающее во сколько раз меняется объем единичного куба, порожденного базисными векторами e_1, e_2, \dots, e_n , (а также и любой объем) под действием матрицы $A : R^n \rightarrow R^n$. Поэтому, при последовательном умножении на матрицы объем соответственно меняется.

6.1.1 Комбинаторика матриц

Решение линейных уравнений Нахождение вектора $x = (x_1, \dots, x_n)$ из системы уравнений $Ax = y = (y_1, \dots, y_n)$ с $(n \times n)$ -матрицей $A = (a_{ij})$ есть хорошо известный естественный трюк. Например, если $a_{11} \neq 0$, то в системе линейных уравнений

$$\sum_{j=1}^n a_{ij} x_j = y_i \tag{32}$$

можно исключить неизвестное x_1 (если оно есть) из уравнений с $i = 2, \dots, n$, прибавляя к уравнению i уравнение 1, умноженное на соответствующее число. Если $a_{11} = 0$. то найдем строку i с $a_{i1} \neq 0$, переставим ее с первой строкой, а далее сделаем то же самое. После этого мы будем иметь уравнение 1 и $n - 1$ уравнений с матрицей A_{n-1} , где не будет неизвестного x_1 . Если строки матрицы A линейно независимы, то и строки A_{n-1} будут также линейно независимы. Следующие шаги - исключаем по индукции x_2, x_3, \dots . На последнем шаге n мы получаем одну строку с неизвестной x_n , которая очевидно находится. Далее обратная процедура дает нам все неизвестные.

Эта процедура также показывает, что следующие 3 элементарные операции оставляют то же множество решений: 1) перестановка уравнений, 2) умножение левой и правой

частей одного из уравнений на одно и то же число, 3) прибавление одного из уравнений к другому. Заметим также, что эти операции не меняют в матрице ни числа линейно независимых строк ни числа линейно независимых столбцов.

Вместо того, чтобы делать эти операции “руками”, можно использовать умножение (как например $A \rightarrow E(i, j)A$) на соответствующие матрицы (называемые **элементарными матрицами**): 1) матрица $E(i, j) = E^{-1}(i, j)$, полученная из единичной матрицы $E = (\delta_{ij})$ (где δ_{ij} называется символом Кронекера, причем $\delta_{ii} = 1$, а остальные равны нулю) перестановкой строк i и j , 2) $E(bi) = E^{-1}(b^{-1}i)$ получается из E умножением строки i на число b , 3) $E(j \rightarrow i) = E^{-1}(-j \rightarrow i)$, получается из E прибавлением строки j к строке i . Конечно, одновременно та же матрица должна быть применена к правой части системы уравнений. Этот подход позволяет лучше понять хорошо известные факты и формулы:

1) Любая обратимая матрица является произведением элементарных.

2) У квадратной матрицы A максимальное число линейно независимых строк равно максимальному числу линейно независимых столбцов. Это число совпадает с ее рангом $r(A)$. Действительно, используя элементарные операции, можно сделать нулевыми $n - r(A)$ строк и столбцов.

3) Также на этом пути может быть понята знаменитая формула Крамера

$$x_k = \frac{\det A_k}{\det A}, \quad (33)$$

для решения линейной системы уравнений (32) с $\det A \neq 0$, где A_k есть A , в которой k -ый столбец заменен столбцом y правой части в (32). Например, пусть мы приводим A к треугольному виду с x_1, \dots, x_{n-1}, x_n на диагонали, и одновременно A_n (теми же операциями), получая треугольную матрицу (определение см. ниже) с диагональю x_1, \dots, x_{n-1}, y_n . Тогда $\det A = x_1 \dots x_{n-1} x_n$, $\det A_n = x_1 \dots x_{n-1} y_n$, что дает $x_n = y_n$.

4) Формула для обратной (с ненулевым определителем)

$$A^{-1} = \frac{1}{\det A} ((-1)^{i+j} M_{ij}), \quad (34)$$

где (M_{ij}) называется минором A , и является матрицей A без i -ой строки и j -того столбца. Действительно, вычисление обратной матрицы может быть сведено к последовательному решению систем уравнений. Поэтому достаточно доказать эту формулу для простейших (см. ниже) матриц.

6.1.2 Спектр и подобие

Спектр линейного оператора A есть множество всех λ таких что оператор $A - \lambda E$ не является обратимым (не имеет обратного). То есть существует $\lambda \in C$, называемое **собственным значением** и хотя бы один вектор ψ , называемый **собственным вектором** (с собственным значением λ), такие что

$$A\psi = \lambda\psi \quad (35)$$

Число точек спектра не превосходит N и не меньше чем 1, то есть существует по крайней мере одна пара λ и ψ , удовлетворяющие (35). Для доказательства определим **резольвенту** оператора A в C^N как следующую рациональную матриц-функцию λ

$$R(\lambda) = (A - \lambda E)^{-1}, \lambda \in C$$

Резольвента существует для всех $\lambda \in C$, кроме случаев когда $A - \lambda E$ необратим, то есть кроме конечного числа нулей **характеристического полинома** $P_A(\lambda) = \det(A - \lambda E)$, то есть когда

$$P(\lambda) = 0 \quad (36)$$

Важно что могут существовать несколько линейно независимых собственных векторов с одним собственным значением λ . Они порождают подпространство называемое **собственным подпространством** $L_\lambda = L_\lambda(A)$. Пусть d_λ - его размерность. При этом любое число n собственных векторов ψ_k с разными собственными значениями λ_k линейно независимы. Для $n = 2$ это очевидно, а далее может доказано индукцией по n . Тогда, все L_λ линейно независимы, но проблема в том, что возможно $\sum_\lambda d_\lambda < N$, то есть они не порождают все пространство. Сейчас мы хотим понять как A действует на остальные векторы.

Простые типы матриц Матрица называется **блочной**, если множество индексов $\{1, \dots, n\}$ разбивается на m последовательных блоков

$$\{1, \dots, k_1\}, \{k_1 + 1, \dots, k_2\}, \dots, \{k_{m-1}, \dots, k_m = n\}$$

таких что $a_{ij} = 0$, если i, j принадлежат различным блокам. Блочная матрица определяет базис, разбитый на несколько групп, что в свою очередь, определяет разбиение C^n в прямую сумму линейных подпространств, в каждом из которых матричный блок определяет линейный оператор. А вся матрица является прямой суммой соответствующих операторов. Также легко видеть, что определитель блочной матрицы есть произведение определителей блоков. Специальный случай блочной матрицы есть **диагональная** матрица, где только диагональные элементы a_{ii} могут быть отличны от нуля.

Матрица называется **верхне (нижне) треугольной**, если $a_{ij} = 0$ при $i > j$ ($i < j$). Очевидно, что определитель треугольной матрицы равен произведению диагональных элементов. Матрица называется **блочно-треугольной** если каждый из ее блоков является треугольной матрицей.

Матрица A порядка n называется **Жордановым блоком** (порядка n), если для всех $i = 1, \dots, n-1$ будет $a_{i,i+1} = 1$, и $a_{ii} = \lambda$ для всех i и некоторого λ , а все остальные равны нулю. Будем обозначать Жорданов блок через $J(N, \lambda)$. Блочная матрица называется **Жордановой** матрицей (или имеющей Жорданову форму) если каждый из ее блоков является Жордановым блоком.

Подобие матриц Две матрицы A и B называются **подобными** (иногда их называют также сопряженными), если существует обратимая матрица C такая что $B = CAC^{-1}$. Мы покажем, что матрицы подобны тогда и только тогда когда они представляют один и тот же линейный оператор в разных базисах. Предположим, что есть два базиса, представленные вектор-столбцами e_1, \dots, e_n и g_1, \dots, g_n . Пусть эти базисы связаны линейным преобразованием C как $Ce_i = g_i$, и более того, в этих базисах операторы A, C представлены матрицами

$$g_i = Ce_i = \sum_j c_{ij}e_j, Ae_i = \sum_j a_{ij}e_j, Ag_i = \sum_j b_{ij}g_j,$$

которые мы будем обозначать теми же буквами C, A, B соответственно. Тогда

$$\begin{aligned} C^{-1}Ag_i &= \sum_j b_{ij}C^{-1}g_j = \sum_j b_{ij}e_j \implies \\ \implies C^{-1}ACC^{-1}g_j &= C^{-1}ACe_j = \sum_j b_{ij}e_j \implies B = CAC^{-1} \end{aligned}$$

Легко видеть, что отношение подобия является симметричным, рефлексивным и транзитивным. Эти три свойства определяют **отношение эквивалентности**, то есть разбиение всего множества матриц на классы "подобия". Иначе говоря, матрицы из разных классов не могут быть подобными.

Если A и B подобны, то у них одинаковые определители, ранги, характеристические полиномы, спектр, а также размерности $\dim L_\lambda(A) = \dim L_\lambda(B)$ для любых λ . Это легко доказать. Однако, наличия всех этих свойств недостаточно для того чтобы A и B были подобными. См. пример в [5].

Естественный вопрос - найти в каждом классе эквивалентности матрицу простейшего вида. Но сначала полезно подумать, как могли бы выглядеть простейшие матрицы, не имеющие диагонального представления. Например, чтобы показать, что любая матрица подобна треугольной матрице, рассмотрим произвольную матрицу A , и выберем, как мы делали выше, элементарные матрицы D_k так, чтобы матрица $D_m \dots D_1 A$ была треугольной. Так, если матрица имеет N линейно независимых собственных векторов, то она подобна диагональной матрице с соответствующими собственными значениями на диагонали.

Теорема 2 (Основное утверждение)

Любая матрица подобна Жордановой матрице.

Рассмотрим сначала линейный оператор A в $L = C^N$ с единственным собственным вектором g_1 с собственным значением 0. Покажем, что тогда A подобен Жорданову блоку $J(N, 0)$. Точнее, покажем что A есть **нильпотентный оператор**, то есть что $L^n = 0$ для некоторого n (в нашем случае $n = N$) и существует вектор g_N такой что следующие векторы

$$g_N, g_{N-1} = Ag_N, \dots, g_1 = Ag_2$$

образуют базис L .

Например, пусть $N = 2$, тогда для любого вектора x , линейно независимого с g_1 , будет $Ax = ag_1 + bx, a \neq 0$. Возможны 4 случая: 1) $a = b = 0$, тогда существует второй собственный вектор с нулевым собственным значением, 2) $a = 0, b \neq 0$, тогда x является вторым собственным вектором с нулевым собственным значением, 3) $a \neq 0, b = 0$, это именно то что нам нужно, 4) $a \neq 0, b \neq 0$, тогда вектор был бы $y = \frac{a}{b}g_1 + x$ был бы вторым собственным вектором с собственным значением b . Значит, $b = 0$ и $A(a^{-1}x) = g_1$.

Для любого N , предположим, что для некоторого $n < N$ существует последовательность

$$g_n, g_{n-1} = Ag_n, \dots, g_1 = Ag_2$$

Пусть L_n - подпространство, в котором эти векторы образуют базис. Выберем другие вектора x_1, \dots, x_{N-n} такие, что они вместе с g_1, \dots, g_n образуют базис L . Пусть L_x - подпространство, порожденное x_1, \dots, x_{N-n} . Тогда могут быть две возможности:

1) не существует вектора $x \in L_x$ такого что $Ax \in L_n$, тогда L_n инвариантно по отношению к A и существует по крайней мере один собственный вектор внутри него, что противоречит нашим предположениям;

2) существует вектор x такой, что $Ax = g = c_1g_1 + \dots + c_n g_n \in L_n$. Если $c_n = 0$. Тогда существуют два вектора $x_1 = x$ и $x_2 = c_1g_2 + \dots + c_{n-1}g_n$ такие что $x_1 - x_2$ есть второй собственный вектор. Если $c_n \neq 0$ то $Ay = g_n$ для $y = c_n^{-1}(x - c_1g_2 - \dots - c_{n-1}g_n)$. Отсюда $n = N$.

В случае нескольких собственных векторов ξ_1, \dots, ξ_m с собственным значением 0 будет m инвариантных подпространств, где A будет нильпотентным с показателями n_1, \dots, n_m так что для всех $k = 1, \dots, m$ и некоторых векторов $\xi_{k1} = \xi_k, \xi_{k2}, \dots, \xi_{kn_k}$ (базис подпространства L_k)

$$A\xi_{k1} = 0, A\xi_{k2} = \xi_{k1}, \dots, A\xi_{kn_k} = \xi_{k,n_k-1},$$

и $L_1 \oplus \dots \oplus L_m = L$. Для нескольких ξ_1, \dots, ξ_m с произвольными собственными значениями λ_k ситуация аналогична. Единственная разность состоит в том, что

$$A\xi_{k1} = \lambda_k \xi_{k1}, A\xi_{k2} = \xi_{k1} + \lambda_k \xi_{k2}, \dots, A\xi_{kn_k} = \xi_{k,n_k-1} + \lambda_k \xi_{kn_k}$$

Доказательство - сведение к предыдущему случаю. Например, пусть $N = 2$ и единственный собственный вектор g_1 имеет собственное значение $\lambda \neq 0$. Рассмотрим оператор $B = A - \lambda E$, единственный собственный вектор g_1 имеет собственное значение 0. Тогда B подобен Жорданову блоку $J(2, 0)$, а A подобен $J(2, \lambda)$, так что $Ax = ag_1 + bx$, $a \neq 0$. Если $b \neq 0$ то вектор $y = \frac{a}{b}g_1 + x$ есть собственный вектор с собственным значением b . Отсюда $b = 0$ и $A(a^{-1}x) = g_1$. Если для $N = 2$ есть только один собственный вектор g_1 с собственным значением λ , рассмотрим оператор $B = A - \lambda E$, у которого только один собственный вектор g_1 с собственным значением 0. Тогда B подобен Жорданову блоку $J(2, 0)$, а A подобен $J(2, \lambda)$.

6.2 Примеры локальной и бесконечномерной линейности

6.2.1 Диффеоморфизмы

Локальный диффеоморфизм

Любое гладкое отображение локально линейно в некотором смысле - чем меньше окрестность, тем лучше отображение на этой окрестности приближается линейным.

Отображение открытой области $\Lambda_1 \subset R^d$ на другую открытую область $\Lambda_2 \subset R^n$ называется **диффеоморфизмом** если оно взаимно однозначно и непрерывно дифференцируемо.

Сначала пусть $d = 1$ а функция $f(x)$ определена в окрестности нуля. Когда можно утверждать что она отображает некоторую окрестность нуля взаимно-однозначно на некоторую окрестность точки $f(0)$? Для этого достаточно чтобы $f(x)$ была непрерывно дифференцируема на некотором интервале, содержащем точку 0, а также чтобы $f'(0) \neq 0$. Тогда на некотором интервале $[x_1, x_2]$, содержащем 0, ее производная имеет один и тот же знак, и ответ очевиден из монотонности функции на этом интервале, так как

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} f'(x) dx \neq 0 \quad (37)$$

Аналогично для любой размерности d . Пусть отображение открытого множества $O_1 \subset R^d$ в R^d задано системой функций $f_i(x_1, \dots, x_d), i = 1, \dots, d$, которые имеют непрерывные частные производные $\frac{\partial f_i}{\partial x_j}, i, j = 1, \dots, d$. Тогда, если матрица Якоби $(a_{ij} = \frac{\partial f_i}{\partial x_j})$, в каждой точке O_1 , имеет ненулевой определитель (**Якобиан**), то некоторая окрестность каждой точки (x_1^0, \dots, x_d^0) взаимно однозначно отображается на некоторую окрестность точки $(f_1(x_1^0, \dots, x_d^0), \dots, f_d(x_1^0, \dots, x_d^0))$.

Действительно, рассмотрим две точки (x_1^0, \dots, x_d^0) и $(x_1^0 + \delta_1, \dots, x_d^0 + \delta_d)$, а также прямой сегмент между ними: $x_k^0 + t\delta_k, 0 \leq t \leq 1, k = 1, \dots, d$. Производная вектор функции $F = (f_1, \dots, f_d)$ в направлении этого пути равна

$$\frac{dF}{dt} = \left\{ \sum_j \delta_j \frac{\partial f_i}{\partial x_j}, i = 1, \dots, d \right\}$$

Но ни один из этих векторов не может быть тождественным нулем из-за линейной независимости строк и столбцов матрицы Якоби. Значит, по крайней мере одна из компонент отлична от нуля. Тогда, аналогично (37), вектора $f(x_1^0, \dots, x_d^0)$ и $f(x_1^0 + \delta_1, \dots, x_d^0 + \delta_d)$ различны.

Замена меры Обозначим $\mu = \mu_0$ меру Лебега на R . В математической физике есть популярное понятие «замена меры». Простейшие примеры:

1. Пусть дана функция $f(x)$ на интервале $[a, b]$. Тогда, как мы знаем, мерой Лебега любого подинтервала $I = [c, d]$ будет его длина $d - c$. Новой мерой будет $\mu_1(I) = \int_I f(x)dx$, если конечно все эти интегралы существуют:

2. Пусть функция $f(x)$ возрастающая, тогда еще одной новой мерой будет

$$\mu_2(I) = f(d) - f(c) = \int_I f'(x)dx = \mu(f(I))$$

Замена переменных В одномерном случае формула замены переменных в интеграле

$$\int_a^b f(y)dy = \int_{x(a)}^{x(b)} f(y(x))y'(x)dx$$

проще всего понимается через предел сумм

$$\sum f(y_k)(y_{k+1} - y_k) \sim \sum f(y(x_k))y'_x(x_k)(x_{k+1} - x_k)$$

или

$$\sum f(y_k)\Delta_k(y) \sim \sum f(y(x_k))y'_x(x_k)\Delta_k(x)$$

В многомерном случае, если $\Lambda \subset R^d$ - открытое множество и $y : \Lambda \rightarrow y(\Lambda)$ - ее диффеоморфизм на $y(\Lambda)$, то формула замены переменной имеет соответственно вид

$$\int_{\Lambda} f(y)dy = \int_{x(\Lambda)} f(y(x))J(x)dx,$$

где $J(x)$ - Якобиан отображения $y(x)$.

В качестве примера рассмотрим вычисление гауссова интеграла

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi} \quad (38)$$

Полярные координаты (r, φ) на плоскости с координатами (a, b) определяются как ???

$$(a, b) = (r, \varphi), r = \sqrt{a^2 + b^2}, \varphi = ???$$

Формула (38) тогда выводится переходом к полярным координатам. Действительно, достаточно доказать, что двойной интеграл (квадрат нашего интеграла)

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy$$

равен 2π . При переходе к полярным координатам Якобиан равен r , что дает

$$\int_0^{\infty} \left(\int_0^{2\pi} d\phi \right) e^{-\frac{r^2}{2}} r dr = \int_0^{\infty} \int_0^{2\pi} d\phi e^{-\frac{r^2}{2}} d\left(\frac{r^2}{2}\right) = \int_0^{\infty} (2\pi) e^{-x} dx = 2\pi$$

Фундаментом всей физики является время и динамика - эволюция состояния различных системы со временем. В основном рассматривается три вида динамик, а также их смеси. Мы даем здесь их простейшие примеры.

6.3 Преобразование Фурье и обобщенные функции

Конечные циклические группы Рассмотрим циклическую группу $I_N = \{0, 1, \dots, N-1\}$ со сложением по модулю N , и гильбертово пространство $l_2(I_N)$ комплексных функций $f(n), n \in I_N$, на нем со скалярным произведением и нормой

$$(f, g) = \sum_{n=0}^{N-1} f(n)g^*(n), \quad \|f\|^2 = \sum |f_n|^2 < \infty \quad (39)$$

Неравенство Коши

$$|(f, g)| \leq \|f\| \|g\|, \quad (40)$$

которое легко следует из неравенства $2|ab| \leq a^2 + b^2$, если возвести обе части (40) в квадрат и раскрыть скобки.

В $l_2(I_N)$ есть два примечательных базиса: 1) $e_m(n) = \delta_{m,n}$, 2) $\tilde{e}_m(n) = \frac{1}{\sqrt{N}} \exp(2\pi i \frac{m}{N} n)$, где ортогональность и нормировка легко проверяются с помощью формулы

$$\sum_{m=0}^{N-1} (\exp 2\pi i \frac{n}{N})^m = 0$$

для $n \neq 0$. **Преобразованием Фурье** здесь называется унитарный оператор U в $l_2 - (I_N)$, который отображает e_m на \tilde{e}_m , а обратное преобразование есть обратный оператор $U = U^{-1} = U^*$. Таким образом, явные формулы прямого и обратного преобразования, с функцией $f(m)$, таковы

$$\tilde{f}(n) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} f(m) \exp(2\pi i \frac{m}{N} n), \quad f(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \tilde{f}(n) \exp(-2\pi i \frac{n}{N} m)$$

Гильбертово и Банахово пространства Здесь впервые мы поднимаемся от конечномерной линейной алгебры к бесконечномерной. Гильбертово пространство может быть определено как векторное пространство со скалярным произведением и ортонормированным счетным базисом таким что свойство типа (39) имеет место при $N = \infty$. Однако иногда базис не так легко найти, и необходимо ввести более общие понятия. **Нормированное пространство** - линейное пространство над R или C с нормой $0 \leq \|x\| < \infty$ такой что: 1) $\|x\| = 0 \iff x = 0$, 2) $\|ax\| = |a|\|x\|$ для любого числа a , 3) $\|x+y\| \leq \|x\| + \|y\|$. Метрика в нем определяется как $\rho(f, g) = \|f-g\|$. Пространство называется **полным** если любая последовательность Коши в этой метрике сходится к некоторому элементу пространства. Полное нормированное пространство называется **банаховым пространством**.

Два примера банаховых пространств на конечном или счетном множестве X : $l_1(X) \subset l_\infty(X)$, соответственно с нормами

$$\|f\|_1 = \sum_{x \in X} |f(x)|, \quad \|f\|_\infty = \sup_{x \in X} |f(x)|$$

Банахово пространство, в котором норма определена скалярным произведением, называется **гильбертовым пространством**. Легко показать, что любое гильбертово пространство имеет ортонормированный базис, и если оно счетный, то пространство называется **сепарабельным**.

Проекция вектора x на вектор y определяется как $(x, y) \frac{y}{\|y\|}$, а на подпространство с ортонормированным базисом $\{g_k\}$ как $\sum_k (x, g_k) g_k$.

Классический пример гильбертова пространства - множество $L_2([0, 1])$ комплексных или вещественных функций на интервале $[0, 1]$ таких что квадрат их модуля интегрируем. Так как мы, специально, не рассматриваем измеримых функций во всей общности, удобно определить $L_2([0, 1])$ и такие же как пополнение любого класса достаточно простых функций в данной норме, не отождествляя эти предельные элементы с конкретными функциями. Например, класса кусочно-постоянных, непрерывных или гладких функций. Легко показать, что в каждом из этих трех случаев мы получим одно и то же пространство.

Группа Z Сначала мы определим два основных гильбертовых пространства. Первое из них $l_2(Z)$, где Z - целочисленная решетка (множество всех целых чисел). Пусть $e_n(k) = \delta_{nk}$ - стандартный ортонормированный базис в нем. Любой элемент (функция f на Z) $l_2(Z)$, по определению, может быть записан в виде

$$f = \sum_{n \in Z} f_n e_n$$

с комплексными коэффициентами $f_n = f(n)$, со скалярным произведением и нормой, определяемыми как выше. Выполнено и неравенство Коши.

Второе пространство обозначается $L_2(S)$. Это множество функций на окружности S длины 1 с конечной нормой и скалярным произведением

$$\|f\|^2 = \int_0^1 |f(x)|^2 dx < \infty, \quad (f, g) = \int f(x) g^*(x) dx,$$

где конечно неравенство Коши также имеет место. Ортонормированный базис $g_n = e^{2\pi i n x}$, $x \in [0, 1)$ определяет изоморфизм между l_2 и $L_2(S)$:

$$f = f(n) = \sum_n f(n)e_n \iff \phi(x) = \sum_n f(n)e^{2\pi i n x} \quad (41)$$

Мы покажем унитарность (то есть сохранение скалярных произведений) преобразования Фурье:

$$\begin{aligned} (\tilde{f}, \tilde{g}) &= \int \tilde{f}(\tilde{g})^* du = \int_0^1 du \sum_n f(n)e^{2\pi i n u} \sum_m g^*(m)e^{-2\pi i m u} = \\ &= \sum_n f(n) \sum_m g^*(m) \int_0^1 e^{2\pi i (n-m)u} du = \sum_n f(n)g^*(n) = (f, g) \end{aligned} \quad (42)$$

В данном случае, обратное преобразование имеет вид:

$$f(n) = \int \tilde{f}(u)e^{-2\pi i n u} du$$

Остается показать, что образом преобразования Фурье является все L_2 , или иначе говоря, для любой функции f из L_2 по крайней мере один коэффициент (f, g_n) отличен от нуля. Согласно нашему отказу от общепринятого определения L_2 , докажем это например для любой непрерывной функции. Ограничимся вещественными функциями, и пусть для некоторой f , все коэффициенты Фурье равны нулю. Заметим тогда, что любая сумма

$$\sum_k (a_k \cos 2\pi k + b_k \sin 2\pi k) \quad (43)$$

будет ортогональной к нашей функции. Предположим теперь, что f положительна в некоторой точке. Например, пусть $f(0) = 1$, $f(x) < C$ для всех x и некоторой $C > 0$. Докажем, что для достаточно большого N функция

$$\left(\frac{1 + \cos 2\pi x}{2}\right)^N,$$

которая конечно представима в виде (43), не ортогональна к f . Разделим интервал $[-\frac{1}{2}, \frac{1}{2}]$ на 3 части:

$$I_1 = \{x : |x| < \epsilon_1\}, I_2 = \{x : \epsilon_1 \leq |x| < \epsilon_2\}, I_3 = \{x : \epsilon_2 \leq |x|\}$$

Для достаточно малых (но не зависящих от N) $0 < \epsilon_1 < \epsilon_2$ существуют малые $0 < \delta_1 < \delta_2$ и $C_1 > 0$ (также независящие от N). такие что $f(x) > C_1$, $x \in I_1 \cup I_2$, и

$$1 - \delta_1 < \frac{1 + \cos 2\pi x}{2} \leq 1, x \in I_1,$$

$$\frac{1 + \cos 2\pi x}{2} \leq 1 - \delta_2, x \in I_2$$

Разобьем интеграл

$$\int_0^1 \left(\frac{1 + \cos 2\pi x}{2}\right)^N f(x) dx = \int_{I_1} + \int_{I_2} + \int_{I_3},$$

так чтобы первый член был больше чем $2\epsilon_1 C_1 (1-\delta_1)^N$, второй член был бы положителен, а третий - меньше чем $C(1-\delta_2)^N$. Отсюда следует результат.

Так же получаем, что ряд Фурье (41) сходится в метрике L_2 . Ряды Фурье ранее были чрезвычайно популярной областью математики, и многие новые техники были развиты.

Замечание 8 Для некоммутативных групп есть похожая наука - теория характеров и представлений групп, но там алгебра более сложна и во многих областях математической физики не необходима.

Интегралы Фурье Обозначим S класс (пространство Шварца) функций на R , которые бесконечно дифференцируемы и все их производные убывают в бесконечности быстрее любой степени. Преобразование Фурье \tilde{f} функции f (прямое и обратное) обычно определяется как

$$\tilde{f}(u) = C_1 \int_R f(x) e^{-C_2 i u x} dx, \quad (44)$$

и отображает S в себя, что доказывается интегрированием по частям, см. ниже. Исторически несколько вещественных констант использовались. Например: 1) $C_1 = 1 = -C_2$ для прямого и $C_1 = C_2 = 1$ для обратного, 2) $C_1 = \frac{1}{\sqrt{2\pi}}$, $C_2 = -1$ и $C_1 = \frac{1}{\sqrt{2\pi}}$, $C_2 = 1$. Ниже везде мы используем способ 1).

Основные результаты:

1. все эти преобразования взаимно однозначны на S - достаточно доказать, что существует константа $C \neq 0$ такая что для любой $f \in S$

$$\int_R \tilde{f}(u) e^{i u x} du = C f(x) \quad (45)$$

2. для функций из S они сохраняют (с точностью до множителя зависящего только от C_1 и C_2) скалярное произведение в $L_2(R)$.

Докажем сначала (45) для гауссовой плотности, то есть что преобразование Фурье функции $\exp(-\frac{x^2}{2})$ совпадает с ней самой, умноженной на 2π :

$$\begin{aligned} \int_{-\infty}^{\infty} \exp(-\frac{x^2}{2} - itx) dx &= \int_{-\infty}^{\infty} \exp(-(\frac{x}{\sqrt{2}} + \frac{it}{\sqrt{2}})^2 + (\frac{it}{\sqrt{2}})^2) dx = \\ &= \exp(-\frac{t^2}{2}) \int_{-\infty}^{\infty} \exp(-(\frac{x}{\sqrt{2}} + \frac{it}{\sqrt{2}})^2) dx = \exp(-\frac{t^2}{2}) \int_{-\infty+it}^{\infty+it} \exp(-\frac{z^2}{2}) dz, \end{aligned}$$

где мы произвели замену переменных $z = x + it$. Но

$$\int_{-\infty+it}^{\infty+it} \exp(-\frac{z^2}{2}) dz = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$$

Действительно, рассмотрим комплексный интеграл от $\exp(-\frac{z^2}{2})$ вдоль замкнутого контура, состоящего из 4 прямых сегментов на комплексной плоскости:

$$[(-N, 0), (N, 0)], [(N, 0), (N, N+it)], [(N, N+it), (-N, -N+it)], [(-N, -N+it), (-N, 0)]$$

Он равен нулю. Но при $N \rightarrow \infty$ интегралы по второму и четвертому сегментам стремятся к нулю. Тогда первый и третий интегралы имеют разные знаки. Остается воспользоваться формулой (38) ниже.

Для любой $f \in S$:

$$\begin{aligned} \int_R \tilde{f}(u) e^{iux} du &= \int_R \left(\int_R f(y) e^{-iuy} dy \right) e^{iux} du = \int_R \int_R f(y) e^{iu(x-y)} du dy = \\ &= \lim_{N \rightarrow \infty} \int_R f(y) dy \left(\int_{-N}^N e^{iu(x-y)} du \right) = \lim_{N \rightarrow \infty} \int_R f(y) \frac{2 \sin N(x-y)}{x-y} dy = \\ &= (y = \frac{z}{N} + x) = \lim_{N \rightarrow \infty} 2 \int_R f\left(\frac{z}{N} + x\right) \frac{\sin z}{z} dz = f(x) \end{aligned} \quad (46)$$

Последний интеграл, при $N \rightarrow \infty$, стремится к $f(x) \int_{-\infty}^{\infty} \frac{\sin z}{z} dz$, где последний интеграл называется интегральным синусом (или интегралом Дирихле), и хорошо известно, что

$$\int_{-\infty}^{\infty} \frac{\sin z}{z} dz = \pi$$

Значит желаемая константа равна 2π . Что можно проверить, если в (46) взять гауссову плотность в качестве $f(x)$, где мы уже это знаем.

То, что скалярные произведения сохраняются, проверяется аналогичными вычислениями.

Замечание 9 В случае двойных пределов, как например в $\int_{-\infty}^{\infty}$, есть нюанс - в каком смысле его понимать. Во многих случаях, для простоты, будем понимать его как $\lim_{N \rightarrow \infty} \int_{-N}^N$.

При решении многих уравнений очень полезно знать, как изменяется преобразование Фурье после простейших преобразований функции f :

- 1) $f(x) \rightarrow g(x) = f(x + s) \implies \tilde{f}(u) \rightarrow \tilde{g}(u) = e^{ius} \tilde{f}$;
- 2) $f(x) \rightarrow g(x) = \frac{df}{dx} \implies \tilde{f}(u) \rightarrow \tilde{g}(u) = iu \tilde{f}$;
- 3) Свертка двух функций определяется как функция

$$(f * g)(x) = \int f(x-y)g(y)dy,$$

чье преобразование Фурье равно произведению их преобразований Фурье

$$\int e^{-iux} f(x-y)g(y)dydx = \int [f(x-y)e^{-iu(x-y)}d(x-y)]e^{-iuy}g(y)dy = \tilde{f}\tilde{g}$$

Обобщенные функции Часто полезно расширить пространство C^∞ , добавляя к нему “худшие” функции (или даже не функции) без уменьшения свободы алгебраических манипуляций с ними. Например, функция равная 1 в одной точке и нулю во всех остальных, очень важна на Z , но на вещественной оси неотличима от нуля. А вот функция, равная ∞ в одной точке и нулю во всех остальных - неизвестно что, но оказалась очень иллюстративной и важной.

Расширение делается так: каждая функция $f \in C^\infty$ определяет также линейный функционал L_f на C^∞ :

$$L_f(g) = \int fg dx, \quad g \in C^\infty$$

И мы хотим рассмотреть более широкий класс линейных функционалов.

Объясним это на примере знаменитой δ -функции. Рассмотрим функцию $D_N(x) = N$ на интервале $(-\frac{1}{2N}, \frac{1}{2N})$ и 0 в других точках. В пределе $N \rightarrow \infty$, она как раз равна нулю везде кроме точки 0, где она бесконечна. более того, ее интеграл с произвольной функцией из C^∞

$$\int f(x)D_N(x)dx \rightarrow_{N \rightarrow \infty} f(0) = \int f(x)\delta(x)dx$$

Это линейный функционал на C^∞ , определяемый так называемой δ -функцией $\delta(x) = \lim D_N(x)$. Как говорят физики “функцией равной нулю везде, кроме точки 0, в которой она бесконечна, а интеграл от нее равен 1. Вместо простейшей функции D_N можно использовать много других, например гауссову плотность $\frac{1}{\sqrt{2\pi N}} \exp(-\frac{x^2}{N})$.

Другой пример

$$L_1(f) = \int_R f(x)dx$$

линейного функционала на S порождается функцией, тождественно равной 1.

В общем случае, преобразование Фурье \tilde{L} линейного функционала L определяется следующим образом: $\tilde{L}\tilde{f} = Lf$. Мы покажем двумя способами, что преобразование Фурье δ -функции есть функция, тождественно равная 1. Первый способ:

$$\begin{aligned} \tilde{D}_N(u) &= N \int_{-\frac{1}{2N}}^{\frac{1}{2N}} e^{-iux} dx = N \int_{-\frac{1}{2N}}^{\frac{1}{2N}} \cos uxdx = \frac{N}{u} (\sin ux) \Big|_{-\frac{1}{2N}}^{\frac{1}{2N}} = \\ &= \frac{2N}{u} (\sin \frac{u}{2N}) \rightarrow 1 \end{aligned}$$

Второй способ:

$$\begin{aligned} \int_{-\infty}^{\infty} \tilde{f} du &= \lim_{N \rightarrow \infty} \int_{-N}^N \tilde{f} du = \lim_{N \rightarrow \infty} \int f(x) (\int_{-N}^N e^{-iux} du) dx = \lim_{N \rightarrow \infty} \int f(x) \frac{2}{x} \sin Nxdx = \\ &= \lim_{N \rightarrow \infty} \int (f(0) + f'(0)x + O(x^2)) \frac{2}{x} \sin Nxdx \sim \lim_{N \rightarrow \infty} f(0) \int \frac{2}{x} \sin Nxdx = \\ &= \lim_{N \rightarrow \infty} f(0) \int \frac{2}{Nx} \sin Nxd(Nx) = \frac{1}{2\pi} f(0) \int \frac{2}{y} \sin ydy = f(0) \end{aligned}$$

Дадим еще примеры обобщенных функций:

1) Сдвиги δ -функций $\delta(x - x_0)$ дают линейные функционалы $L(f) = \int_R f(x)\delta(x - x_0)dx = f(x_0)$;

2) Обобщенные функции можно складывать, умножать на число и на C_∞ -функции;

3) Производные обобщенных функций определяются так $L'(f) = -L(f')$. В частности, $L'_\delta(f) = -L_\delta(f') = -f'(0)$;

4) Регуляризовать сингулярную (обычную) функцию x^{-1} - значит преобразовать ее в обобщенную. есть два способа: **интеграл в смысле главного значения**

$$Lf = \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \frac{1}{x} f(x) dx,$$

и

$$Lf = \int_R \frac{1}{x} (f(x) - f(0)) dx,$$

что дает простейший пример **вычитания расходимостей**, часто используемый в квантовой теории поля.

Part V

Динамические системы

Теория динамических систем

м

7 Детерминированные

7.1 На конечном множестве

Итерации преобразований конечного множества

Рассмотрим отображение $F : X \rightarrow X$ конечного множества $X = \{1, \dots, N\}$ (называемого **множеством состояний**) в себя. Графом этого отображения называется ориентированный (направленный) граф с множеством вершин X и N направленными ребрами $l_{ij} = (i, j = F(i))$. Если F взаимно-однозначно, то граф состоит из нескольких циклов, где циклом называется последовательность вершин i_1, \dots, i_n таких, что $F(i_k) = i_{k+1}$ для $k = 1, \dots, n-1$ и $F(i_n) = i_1$. Цикл длины $n = 1$ называется **неподвижной точкой**. Начальное состояние i_0 и итерации $i_0, i_1 = F(i_0), \dots, i_{k+1} = F(i_k), \dots$ отображения F определяют эволюцию состояний в дискретном времени $t = k = 0, 1, 2, \dots$

Отображение F индуцирует отображение $L_F : \Phi(X) \rightarrow \Phi(X)$ множества $\Phi(X)$ комплексных функций $f(x)$ на X себя

$$(L_F f)(i) = f(Fi),$$

а также отображение $M_F : M(X) \rightarrow M(X)$, множества $M(X)$ мер μ на X в себя

$$(M_F \mu)(A) = \mu(F^{-1}A), \quad (47)$$

Мера μ называется **инвариантной**, если $M_F \mu = \mu$.

Самые очевидные утверждения почти очевидны (согласны?). Мера называется постоянной на подмножестве $A \subset X$, если она имеет то же самое значение на каждом одноточечном подмножестве A . Если F взаимно-однозначно, то инвариантная мера постоянна на каждом цикле, и единственна (с точностью до постоянного множителя) если есть только один цикл.

Отображение F называется **эргодическим** если инвариантная мера единственна (с точностью до постоянного множителя). В этом случае, для любой начальной меры μ

$$\frac{1}{T} (M_F(\mu) + M_F^2(\mu) + \dots + M_F^T(\mu)) \quad (48)$$

стремится к единственной инвариантной мере при $T \rightarrow \infty$. В общем случае, для любой начальной меры μ (48) сходится к некоторой (зависящей от μ) инвариантной мере.

Аналогичная сходимость имеет место для функций f на X . При $T \rightarrow \infty$

$$\frac{1}{T}(L_F(f) + L_F^2(f) + \dots + L_F^T(f))$$

сходится к некоторой функции $\phi(x)$ зависящей от f . Это конечно тривиальный случай знаменитой теоремы Биркгофа-Хинчина. Если F эргодично, то функция $\phi(x)$ константа.

7.2 Линейные дифференциальные уравнения

7.2.1 Решение

Естественно сначала понять ситуацию с одним (но общим), то есть с одной неизвестной функцией $x(t)$, линейным уравнением первого порядка

$$\frac{dx}{dt} = a(t)x + f(t) \quad (49)$$

Если $f(t) = 0$, а a не зависит от t , то решением будет Ce^{at} , где C - произвольная константа. Она однозначно определяется из начального условия $x(0) = x_0 = C$. Будет ли это решение единственным зависит от класса функций, для которого мы хотим исследовать единственность. Доказательство единственности легко в классе функций, которые в окрестности начальной точки разлагаются в степенной ряд. Из (49) мы имеем рекуррентное соотношение

$$x(t) = \sum_{k=0}^{\infty} a_k t^k \implies (k+1)a_{k+1} = a_k, k = 0, 1, \dots,$$

и подставляя $t = 0$, получаем $a_0 = x(0)$. Отсюда единственность.

С произвольной функцией $a(t)$ мы также ищем решение в виде $x(t) = Ce^{g(t)}$, где $C = x(0)$. Подставляя, получаем простое уравнение, откуда

$$g(t) = \int_0^t a(t)dt, \quad (50)$$

Частное решение неоднородного (то есть с функцией $f(t)$) уравнения будем искать в виде $x = D(t) \exp(\int_0^t a(t)dt)$, откуда получаем простое уравнение для D , решением которого будет

$$\frac{dD}{dt} = f(t) \exp(-\int_0^t a dt) \implies D = \int_0^t f(t) \exp(-\int_0^t a(s)ds) dt$$

Общим решением будет сумма этого решения (оно равно нулю в точке $t = 0$) и решения однородного уравнения с заданным начальным условием.

Матричные (операторные) ряды В множестве всех матриц данной размерности N можно ввести норму $\|A\|$, и значит расстояние $\|A - B\|$, определив тем самым метрическое пространство всех матриц данной размерности. Например,, нормой может быть:

$$\|A\| = \sum_{i,j} |a_{ij}|$$

Тогда для любой квадратной матрицы A матричный ряд (матричная экспонента)

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = E + A + \frac{A^2}{2} + \dots$$

сходится в смысле этой нормы.

Системы уравнений Рассмотрим теперь систему линейных уравнений вида

$$\frac{d\psi}{dt} = A\psi + f, \quad \psi(0) = \varphi_0 \quad (51)$$

где $\psi = \psi(t) = \{\psi_k(t) : k = 1, \dots, n\}$ - неизвестный вектор-столбец, $f = f(t) = \{f_k(t) : k = 1, \dots, n\}$ - известный вектор-столбец, а A - $(n \times n)$ -матрица с постоянными коэффициентами.

Эта система может быть явно решена заменой неизвестного вектора вектором $C(t)$

$$\psi(t) = e^{At}C(t) \iff C(t) = e^{-At}\psi(t), \psi(0) = C(0),$$

откуда получаем уравнение для нового вектора и его решение

$$\frac{dC}{dt} = e^{-At}f \implies C(t) - C(0) = \int_0^t e^{-As}f(s)ds,$$

откуда получается решение уравнения (51), с начальным условием $\psi(0) = C(0)$:

$$\frac{dC}{dt} = e^{-At}f \implies C(t) - C(0) = \int_0^t e^{-As}f(s)ds$$

В случае однородного уравнения, то есть если $f = 0$, например, если возможно, приведем матрицу A к диагональному виду

$$A = DBD^{-1}.$$

с диагональной B . Тогда для вектора $\xi(t) = C^{-1}\psi(t)$ получим уравнение

$$\frac{d\xi}{dt} = B\xi,$$

и проблема сводится к одномерным уравнениям.

В случае если A зависит от времени, то похожий трюк возможен только в одномерном случае, см. уравнение (51). Вот почему, теория, даже для простейшей системы двух уравнений, к которой сводится например уравнение Хилла

$$\ddot{x} = -k(t)x, \quad (52)$$

намного более сложна, см. ниже. Основная проблема - найти линейно независимые решения однородного уравнения. Однако, если эти решения известны, найти решение соответствующего неоднородного уравнения легко. Так, для уравнения

$$\frac{d^2x}{dt^2} + g_1(t)\frac{dx}{dt} + g_0(t)x = f(t)$$

с произвольными g_1, g_0, f , общее решение таково

$$c_1F_1(t) + c_2F_2(t) + F_2(t) \int_0^t F_1(s)f(s)W^{-1}(s)ds - \\ - F_1(t) \int_0^t F_2(s)f(s)W^{-1}(s)ds,$$

где F_1, F_2 - линейно независимые решения однородного уравнения и

$$W = F_1 \frac{dF_2}{dt} - \frac{dF_1}{dt} F_2$$

Заметим, что даже в случае $g_1 = 0, f = 0$, в матричном случае невозможно найти решение как в (50).

Для уравнения

$$\frac{d^2x}{dt^2} = -\omega_0^2x + f(t)$$

будет $F_1 = \cos \omega t, F_2 = \sin \omega t$, и $W(t) = \omega$. И общее решение имеет вид

$$x(t) = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t + \sin \omega_0 t \int_0^t \cos \omega_0 s f(s) ds - \cos \omega_0 t \int_0^t \sin \omega_0 s f(s) ds$$

Отсюда во многих случаях легко найти условия ограниченности.

Уравнение Хилла и мультипликативный интеграл Мы обсудили выше два простых случая уравнения Хилла с постоянной $k(t) = k$. Если $k < 0$ то, как мы видели, решение ограничено. Если $k > 0$, то решение

$$x(t) = \frac{1}{2}(x(0) + \frac{v(0)}{\sqrt{k}})e^{\sqrt{k}t} + \frac{1}{2}(x(0) - \frac{v(0)}{\sqrt{k}})e^{-\sqrt{k}t}$$

растет экспоненциально за исключением случая когда $x(0) = -\frac{v(0)}{\sqrt{k}}$. В случае иррегулярного или случайного $k(t)$ можно ожидать совершенно другое поведение решений. Существует общий подход, см. [6, 7, 8], основанный на понятии **мультипликативного интеграла**, которое является прямым обобщением обычного (аддитивного) интеграла. Если последний является пределом сумм многих малых членов, каждый из которых стремится к нулю, то первый есть предел произведения многих множителей, где каждый стремится к 1. Возможно, что некоторые математики не слышали о деталях такой науки, но ее идеи конечно известны всем. Например, существует естественный приближенный метод (принадлежащий видимо Эйлеру) решения линейных векторных уравнений $\dot{x} = A(t)x$ на интервале $[0, 1]$ с начальным условием $x(0)$. При этом сегмент подразделяется

точками $t_k = \frac{k}{N}$ на N интервалов длины $\frac{1}{N}$, и рассматривается линейная рекуррентная последовательность:

$$x_{k+1} = x_k + \dot{A}(t_k)(t_{k+1} - t_k)x_k = (E + \dot{A}(t_k)(t_{k+1} - t_k))x_k, k = 0, 1, \dots, N - 1$$

Мы получим тогда системы кусочно-постоянных решений на интервалах $t_k \leq t < t_{k+1}$

$$x^{(N)}(t) = x_k^{(N)} = B(k(N))x_0, \quad B(k(N)) = \prod_{k=0}^{\frac{k(N)}{N}} (E + \dot{A}(t_k))$$

Более того, при определенных условиях можно доказать, что существуют предельные матрицы $B(t)$ такие что $B(t)x_0 = x(t)$. и при $\frac{k(N)}{N} \rightarrow t, N \rightarrow \infty$, будет $B(k(N)) \rightarrow B(t)$,

Однако, это может привести скорее в вычислительную математику и в вычислительные методы физики. Более того, неясно, все ли качественные эффекты могут быть замечены на этом пути.

7.2.2 Качественные стороны решения

Общее -линейное уравнение второго порядка имеет вид

$$m\ddot{x} = F(x, \dot{x}, t) = -k(t)x - \alpha(t)\dot{x} + f(t), \quad (53)$$

с начальными условиями $x(0), v(0)$. Правая его часть включает три вида сил для линейных (по x и \dot{x}) **неоднородных** уравнений. Если $f(t) \equiv 0$, то уравнение называется **однородным**. Сначала предположим, что $k(t) \geq 0$ и будем обозначать $k(t) = \omega_0^2(t)$.

Замечание 10 Заметим, что общее линейное уравнение порядка n имеет вид

$$\sum_{k=0}^n a_k(t) \frac{d^k x(t)}{dt^k} = f(t)$$

и может быть сведена к системе порядка n вида (???) заменой переменных

$$\varphi_k = \frac{d^k x}{dt^k}, k = 0, 1, \dots, n - 1$$

Например уравнение превратится в систему с матрицей

Это простейшее из уравнений Ньютона, на которых основана классическая механика.

Ниже мы рассмотрим частные случаи этого уравнения, где важную роль будет играть физическая интерпретация. Здесь $x(t) \in R$ - координата точечной частицы в момент t , число $m > 0$ называется **массой**, далее для упрощения обозначений мы предполагаем, что $m = 1$. Функция F называется **силой** (действующих на частицу), и в общем случае она зависит от **положения**, **скорости** и **времени**. Мы рассматриваем только случаи когда F есть сумма этих трех сил. По порядку:

1) **потенциальная сила** которая часто держит частицу в потенциальной яме. Этот член может быть записан в виде $-\frac{\partial U}{\partial x}$. где функция $U(x, t) = \omega_0^2(t) \frac{x^2}{2}$ называется

потенциальной энергией. Она имеет минимум (низ “потенциальной ямы”) в точке $x = 0$. Полная энергия частицы (или **Гамильтониан**) $H = T + U$ есть сумма потенциальной и **кинетической энергии** $T = m \frac{v^2}{2}$ частицы, где $v = \frac{dx}{dt}$ называется **скоростью** частицы. Если ω_0^2 не зависит от времени и $\alpha = f = 0$, то закон сохранения энергии ($\frac{dH}{dt} = 0$) доказывается в одну строчку. В этом случае надо использовать тот факт, что наше уравнение (53) может быть переписано в виде **гамильтоновой системы** двух уравнений первого порядка: $\dot{x} = v, \dot{v} = -\omega_0^2 x$;

2) **диссипативная сила** (далее предполагается $\alpha(t) \geq 0$), которая уменьшает скорость, а значит и кинетическую энергию;

3) $f(t)$ часто называется **внешней силой**.

Только одна сила Здесь ω^2 и α не зависят от времени.

Если это только потенциальная сила, то уравнение таково

$$\ddot{x} = -\omega_0^2(t)x$$

Его общее решение зависит от констант C_1, C_2 , определяемых из начальных данных

$$x(t) = C_1 \cos \omega_0 t + C_2 \sin \omega_0 t, C_1 = x(0), C_2 = \frac{v(0)}{\omega_0}$$

Закон сохранения энергии говорит, что **на фазовой плоскости** $R^2 = \{(x, v)\}$ точка совершает периодическое движение по эллипсу

$$\frac{v^2(t)}{2} + \omega_0^2 \frac{x^2(t)}{2} = H(0)$$

Если есть только диссипация, то уравнение

$$\frac{d^2x}{dt^2} = -\alpha \frac{dx}{dt}$$

имеет решение

$$v(t) = \frac{dx}{dt} = C_1 \exp(-\alpha t), C_1 = v(0) \implies$$

$$\implies x(t) = C_2 + C_1 \int_0^t \exp(-\alpha s) ds, C_2(0) = x(0)$$

Отсюда следует, что скорость частицы стремится к нулю, а координата стремится к постоянному значению.

Только зависящая от времени внешняя сила Рассмотрим уравнение

$$\frac{d^2x}{dt^2} = \frac{dv}{dt} = f(t)$$

Когда траектории будут ограниченными? Имеем

$$\frac{dx}{dt} = v(t) = v(0) + \int_0^t f(s) ds$$

Скорость ограничена тогда и только тогда, когда последний интеграл ограничен. Например, если f периодична и имеет нулевой интеграл по периоду. Но для координаты

$$x(t) = x(0) + \int_0^t v(s)ds = x(0) + v(0)t + \int_0^t du \int_0^u f(s)ds$$

ситуация другая. Пусть для примера $v(0) = 0, f = \sin(t + A)$. Тогда

$$\begin{aligned} v(t) &= -\cos(t + A) + \cos A \implies \\ \implies x(t) &= x(0) + \int_0^t v(s)ds = x(0) + t \cos A - \sin(t + A) + \sin A, \end{aligned}$$

и траектория ограничена тогда и только тогда, когда $\cos A = 0$.

Две силы Если внешняя сила и диссипация, то уравнение имеет первый порядок

$$\frac{d^2x}{dt^2} = \frac{dv}{dt} = -\alpha v + f(t),$$

и, как мы видели выше, имеет явное решение

$$x(t) = x(0) + \frac{v(0)}{\alpha} - \frac{v(0)}{\alpha} e^{-\alpha t} - \alpha^{-1} \int_0^t e^{\alpha(s-t)} f(s)ds + \alpha^{-1} \int_0^t f(s)ds,$$

Ограниченность будет для тех f , когда последний интеграл ограничен.

Потенциал и диссипация

$$\ddot{x} = -\omega_0^2 x - \alpha v, \alpha > 0$$

Здесь два линейно независимых решения получаются подстановкой $x = e^{\lambda t}$ в уравнение, что дает квадратичное уравнение для λ с двумя корнями

$$\lambda = -\frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2}{4} - \omega_0^2}$$

Например, если $\frac{\alpha^2}{4} > \omega_0^2$, то оба корня λ_1 и λ_2 отрицательны, и решение имеет вид

$$C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t},$$

где C_1, C_2 определяются из начальных условий. Это определяет экспоненциальную сходимость к нулю для всех ω и всех начальных условий. Случаи $\frac{\alpha^2}{4} < \omega_0^2$ и $\frac{\alpha^2}{4} = \omega_0^2$ читатель с легкостью рассмотрит сам.

Потенциал и постоянная внешняя сила В случае постоянной силы $f = const$ и диссипации, мы имеем уравнение

$$\ddot{x} = -\omega^2 x + f - \alpha v - \ddot{x} = -\omega_0^2 x - \alpha(v - \frac{f}{\alpha}), \quad (54)$$

Без диссипации решение

$$x = \frac{f}{\omega_0^2} + C_1 \cos \omega_0 t + C_2 \sin \omega_0 t$$

просто сдвигает точку, около которой происходят осцилляции. Если $\alpha \neq 0$, то подстановка $x = \frac{\alpha}{\omega_0^2} + y$ показывает что будет экспоненциальная сходимость $x(t)$ к точке $\frac{f}{\omega_0^2}$.

Резонанс Для $f = a \cos \omega t$, $\omega \neq \omega_0$, общее решение имеет вид суммы частного решения неоднородного уравнения и общего решения однородного

$$\frac{a}{\omega_0^2 - \omega^2} \cos \omega t + C_1 \cos \omega_0 t + C_2 \sin \omega_0 t. \quad (55)$$

При этом константы C_1, C_2 выбираются так чтобы решение (55) удовлетворяло начальным условиям. Чем ближе частота ω внешней силы к собственной частоте ω_0 , тем больше становится амплитуда колебаний (явление **резонанса**). При $\omega = \omega_0 > 0$, частное решение неоднородного уравнения имеет вид

$$\frac{a}{2\omega_0} t \sin \omega_0 t, \quad (56)$$

и максимум амплитуды растет линейно со временем. Такие резонансные члены (обычно называемые **секулярными**) как (56) часто появляются и доставляют уйму проблем в много более сложных задачах, например, в случае общей внешней силы $f = f(x, t)$.

Таким образом, траектория будет неограниченной только при одной частоте $\omega = \omega_0$. И может показаться, что неограниченность траектории из-за внешних возмущений - редкое явление. Но если мы сдвинем фазу резонансной силы как $f(t) = \cos(\omega t + \phi)$, то решение снова будет резонансным

$$\frac{a}{2\omega_0} t \sin(\omega_0 t + \phi)$$

Общего понимания ситуации для общих систем пока нет. Для этого надо понять механизм резонанса менее формально.

Непериодическая сила Более сложный пример: рассмотрим произвольную гладкую функцию $a(\omega)$ и

$$f(t) = \int a(\omega) \cos \omega t d\omega$$

Предположим, что $a(\omega)$ гладкая и имеет носитель в интервале $(\omega_0 - \epsilon, \omega_0 + \epsilon)$, где $0 < \epsilon < \omega_0$ (предполагая $\omega_0 > 0$) не обязательно мало.

Утверждение 2 Для любых начальных данных решение равномерно ограничено интервале времени $[0, \infty)$.

Доказательство. Решение таково (где обозначено $\omega = \omega_0 + x$)

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \int \frac{a(\omega) \cos \omega t}{\omega_0^2 - \omega^2} d\omega &= \lim_{\epsilon \rightarrow 0} \int \int_{-\epsilon}^{\epsilon} \frac{a(\omega_0 + x) \cos(\omega_0 + x)t}{\omega_0^2 - (\omega_0 + x)^2} dx = \lim_{\epsilon \rightarrow 0} \int \int_{-\epsilon}^{\epsilon} \frac{a(\omega_0 + x)}{x} \frac{1}{-2\omega_0 - x} \cos(\omega_0 + x)t dx \\ &= \lim_{\epsilon \rightarrow 0} \int \frac{a(\omega_0)}{-2\omega_0} \int_{-\epsilon}^{\epsilon} \frac{\cos(\omega_0 t + xt)}{x} dx = \lim_{\epsilon \rightarrow 0} \int \frac{a(\omega_0)}{-2\omega_0} \int_{-\epsilon}^{\epsilon} \left(\frac{\cos \omega_0 t \cos xt}{x} - \frac{\sin \omega_0 t \sin xt}{x} \right) dx \end{aligned}$$

Достаточно рассмотреть только случай $a(\omega_0) \neq 0$. Видно, что неограниченность по времени может возникнуть при интегрировании по малой окрестности ω_0 Первый член в окрестности точки $x = 0$ нулевой из-за (\pm) -симметрии. В то же время

$$\int_{-\epsilon}^{\epsilon} \frac{\sin xt}{x} dx = \int_{-t\epsilon}^{t\epsilon} \frac{\sin x}{x} dx$$

ограничен равномерно по t . Действительно, на произвольном периоде $(N, N + 2\pi)$ положим $x = N + y$, тогда

$$\frac{1}{x} = \frac{1}{N} \frac{1}{1 + \frac{y}{N}} = \frac{1}{N} - \frac{y}{N^2} + \dots$$

Первый член даст 0 в интегралах по таким периодам, а остальные дадут сходящуюся сумму.

Замечание 11 *Есть много других примеров ограниченных решений для других внешних сил $f(t)$, включая случайные, но интересно было бы иметь полный обзор имеющихся результатов.*

7.3 Нелинейность

7.3.1 Существование и единственность

7.3.2 Теория возмущений

8 Случайные

8.1 Классическая вероятность на конечных множествах

Рассмотрим произвольное конечное или счетное множество Ω . Его элементы ω назовем **элементарными событиями**. Все подмножества $A \subset \Omega$ назовем **событиями**. Пусть на Ω задана функция $P(\omega) \geq 0$, причем $\sum_{\omega \in \Omega} P(\omega) = 1$. Тогда

$$P(A) = \sum_{\omega \in A} P(\omega)$$

называется **вероятностью события** A , а сама функция $P(A)$ на множестве всех подмножеств Ω называется **вероятностной мерой**. Любую функцию $\xi(\omega)$ на Ω назовем **случайной величиной**. С ней можно связывать разные события, например событие $\{\omega : \xi(\omega) < x\}$, что значение ξ не превосходит числа x . **Функцией распределения** ξ называется функция от $x \in R$

$$F_\xi(x) = P(\{\omega : \xi(\omega) < x\}) = \sum_{\omega: \xi(\omega) < x} P(\omega),$$

а ее производная $\rho_\xi(x) = \frac{dF_\xi(x)}{dx}$ называется **плотностью распределения**. Среднее и **дисперсия** определяются соответственно как

$$E\xi = \sum_{\omega} P(\omega)\xi(\omega), D\xi = E(\xi - E\xi)^2$$

Характеристической функцией случайной величины ξ называется функция

$$f_\xi(t) = Ee^{it\xi}, t \in R$$

Одно из самых популярных конечных подмножеств Ω - множество $X_N = \{\omega = (x_1, \dots, x_N) : x_i = a, b\}$ всех 2^N последовательностей (слов) длины N из двух вещественных чисел a, b . Самый популярный способ задать вероятностную меру на нем это задать два положительных числа $p, q : p + q = 1$, и положить

$$P(\omega) = p^{N_a} q^{N - N_a},$$

где N_a - число i таких, что $x_i = a$. Введем случайные величины $\xi_i(\omega) = \xi_i(x_1, \dots, x_N) = x_i$. Легко видеть, что, например, вероятность того, что данного i случайная величина ξ_i равна a

$$P(\xi_i = a) = P\{\{\omega : \xi_i(\omega) = a\}\} = p$$

Более того, для любых $1 \leq i_1 < i_2 < \dots < i_k \leq N$ вероятность того, что $\xi_i(\omega) = a_i$ для всех $i = i_1, i_2, \dots, i_k$ и некоторых $a_i \in \{a, b\}$, равна

$$\prod_{i=i_1, \dots, i_k} P(\xi_i = a_i)$$

Такое свойство называется (взаимной) **независимостью** случайных величин ξ_1, \dots, ξ_N . Пусть $S_N = \xi_1 + \dots + \xi_N$, тогда

$$E\xi_k = pa + qb = d, ES_N = N(pa + qb),$$

$$D\xi_k = c = p(a-d)^2 + q(b-d)^2, DS_N = \sum_{k=1}^N D\xi_k = cN,$$

так как $E(\xi_i - d)(\xi_j - d) = 0$ при $i \neq j$.

И имеет место **закон больших чисел** - с большой вероятностью случайная величина S_N примерно равна Nd . Точнее, для любого $\epsilon > 0$ при $N \rightarrow \infty$

$$P\left(\left|\frac{S_N}{N} - d\right| > \epsilon\right) \rightarrow 0$$

Для доказательства введем случайные величины $\eta_k = \xi_k - E\xi_k \implies E\eta_k = 0$. Тогда

$$S_N = dN + Q_N, Q_N = \sum_{k=1}^N \eta_k,$$

и нам достаточно доказать, что при $N \rightarrow \infty$

$$P\left(\frac{|Q_N|}{N} > \epsilon\right) \rightarrow 0$$

Но это следует из более сильного неравенства - что для любого $\frac{1}{2} < \alpha < 1$

$$DQ_N = D|Q_N| > P(|Q_N| > N^\alpha)N^{2\alpha} \implies P(|Q_N| > N^\alpha) = P\left(\frac{|Q_N|}{N} > N^{\alpha-1}\right) < N^{-2\alpha} D|Q_N|$$

А сейчас более точно, как ведет себя $Q_N = \sum_{k=1}^N \eta_k$ - именно, что с большой вероятностью Q_N имеет порядок \sqrt{N} . Для простоты обозначений пусть теперь $\eta_k = \pm 1$ с вероятностями $\frac{1}{2}$. Тогда характеристическая функция $f_N(t)$ случайной величины $\frac{Q_N}{\sqrt{N}}$ равна

$$E \exp(it \frac{Q_N}{\sqrt{N}}) = \prod E \exp(it \frac{\eta_k}{\sqrt{N}}) = \prod \frac{e^{\frac{it}{\sqrt{N}}} + e^{-\frac{it}{\sqrt{N}}}}{2} = \cos^N \frac{t}{\sqrt{N}}$$

и отсюда

$$f_N(t) = \left(1 - \frac{t^2}{2N} + o\left(\frac{t^2}{N}\right)\right)^N \rightarrow e^{-\frac{t^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx$$

то есть это характеристическая функция случайной величины, имеющей гауссову плотность распределения $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

8.2 Конечные цепи Маркова

Конечные цепи Маркова с дискретным временем Здесь состояниями системы будут не только точки множества X , но и все вероятностные (неотрицательные и нормированные) меры p на X . Меры одноточечных множеств $\{k\}$ (одноточечные меры) будем обозначать p_k . Так, состояние системы в момент $t = 0, 1, \dots$ будем обозначать вектор-строкой $p(t) = (p_1(t), \dots, p_N(t))$.

Общая линейная динамика на множестве вероятностных мер на X определяется $(N \times N)$ -матрицей $P = (p_{ij})$ с неотрицательными элементами, и более того, $\sum_j p_{ij} = 1$ для всех i . Такие матрицы называются **стохастическими**. Динамика определяется начальным вектором-строкой и матрицей P

$$p(t) = p(0)P^t,$$

где компоненты $p_i(t) : i \in X$ последнего вектора называются вероятностями системе быть i в момент t . Если компоненты $p_i(0), i = 1, \dots, N$ начального вектора $p(0)$ неотрицательные и нормированные, то есть $\sum_{i=1}^N p_i(0) = 1$, то легко доказать что все $p(t)$ обладают этими свойствами. Простые примеры: 1) если матрица P имеет одинаковые строки), то все меры $p(t)$ будут одинаковыми для $t > 0$, 2) каждая строка имеет ровно один ненулевой элемент (равный конечно единице), то это соответствует индуцированной детерминированной динамике (47).

Построенная динамика мер называется конечной цепью Маркова с дискретным временем. Вероятностная мера π называется инвариантной, если $\pi P = \pi$. Марковская цепь называется **эргодической** если для любого начального состояния $p(0)$ при $t \rightarrow \infty$

$$p(0)P^t \rightarrow \pi$$

Докажем, что любая матрица P , все элементы которой положительны, определяет эргодическую цепь Маркова. Для этого используем **метод последовательных приближений**. Он состоит в следующем утверждении.

Пусть M - полное метрическое пространство с расстоянием ρ и конечным диаметром $D = D(M) = \sup_{x,y \in M} \rho(x,y)$. Задано отображение $F : M \rightarrow M$ такое что существует число $0 < \alpha < 1$ такое что для любых двух точек $x, y \in M$

$$\rho(Fx, Fy) < \alpha \rho(x, y) \tag{57}$$

Тогда F имеет единственную неподвижную точку $Fx = x$.

Действительно, тогда для любой x_0 последовательность

$$x_0, x_1 = F(x_0), \dots, x_{k+1} = F(x_k), \dots$$

является последовательностью Коши, так как

$$\rho(x_k, x_{k+1}) < \alpha \rho(x_{k-1}, x_k) < \alpha^k D$$

Тогда последовательность x_k имеет предел $x = Fx$. Любая другая последовательность y_k , начинающаяся с произвольной y_0 имеет тот же предел, так как точки x_k и y_k сближаются когда $k \rightarrow \infty$.

В нашем случае определим норму $|\mu|$ произвольной меры μ , и расстояние между двумя мерами μ_1, μ_2 (не обязательно вероятностными) как

$$|\mu| = \sum_{x \in X} |\mu(x)|, \quad \rho(\mu_1, \mu_2) = |\mu_1 - \mu_2| = \sum_{x \in X} |\mu_1(x) - \mu_2(x)|$$

Назовем носителем меры μ множество точек с ненулевой мерой. Тогда, если две меры μ_1, μ_2 имеют непересекающиеся носители, то $\rho(\mu_1, \mu_2) = \sum_{x \in X} |\mu_1(x)| + |\mu_2(x)|$. Меры $\mu_1 - \mu_2$ и $(\mu_1 - \mu_2)P$ могут быть записаны в виде разности двух неотрицательных мер ν_1, ν_2 с одной нормой и непересекающимися носителями. Мы хотим доказать, что для любых таких двух мер норма меры $(\mu_1 - \mu_2)P$ уменьшится в сравнении с нормой меры $(\mu_1 - \mu_2)$

$$\|(\mu_1 - \mu_2)P\| \leq \alpha \|\mu_1 - \mu_2\|$$

где

$$\alpha \leq 1 - \min_{i,j \in X} p_{ij}$$

Действительно, рассмотрим точки x_1, x_2 , где $\nu_m(x_m) \geq \frac{1}{N}|\nu_m|, m = 1, 2$. Тогда для любой $y \in X$ в разности $\nu_1(x_1)p_{x_1y} - \nu_2(x_2)p_{x_2y}$ возникнут плюс-минус сокращения в

$$(\min \nu_1(x_1))(\min p_{xy}) \geq \frac{1}{N}|\nu_1|(\min p_{xy})$$

Суммирование по y даст результат.

Это утверждение обобщается на более широкий класс матриц P - именно ???

Сейчас мы по-другому посмотрим на эту динамику. Рассмотрим множество

$$\Omega = \Omega_\infty = \{\omega = (x_0, x_1, \dots, x_n, \dots)\}$$

бесконечных последовательностей элементов $x_n \in X$. Элементы $\omega \in \Omega$ будем называть **элементарными событиями** и представлять их как возможные **траектории** (пути) частицы, прыгающей по X от точек x_n к x_{n+1} . А индексы $0, 1, 2, \dots$ при этом будем рассматривать последовательность дискретных моментов времени.

Введем основные события $A(i_0, \dots, i_n)$ как множества бесконечных последовательностей ω таких что $x_k = i_k$ для $k = 0, 1, \dots, n$. Обозначим Σ минимальную σ -алгебру, порожденную всеми этими основными событиями. Пусть дана мера $p_0(x)$ на X и стохастическая матрица P . Определим вероятностную меру μ_n на событиях $A(i_0, \dots, i_n)$

$$\mu_n(A(i_0, \dots, i_n)) = p_0(i_0)p_{i_0i_1}p_{i_1i_2}\dots p_{i_{n-1}i_n}$$

Эти меры согласуются для разных n в том смысле, что

$$\mu_n(A(i_0, \dots, i_n)) = \sum_{i_{n+1} \in X} \mu_{n+1}(A(i_0, \dots, i_{n+1}))$$

Тогда теорема Колмогорова утверждает, что на σ -алгебре Σ , порожденной основными событиями, существует единственная мера μ , которая на базовых множествах совпадает с мерами μ_n . Множества $A \in \Sigma$ называются **событиями**, а меры $\mu(A)$ - **вероятностями** этих событий. Вещественные функции $\xi = \xi(\omega)$ на Ω называются случайными величинами, если удовлетворяют следующим условиям измеримости: прообраз $\xi^{(-1)}(I)$ любого множества $I \subset R$ принадлежит Σ . Например, для любого n , все функции зависящие только от первых n координат, измеримы.

8.3 Случайные блуждания

Пусть теперь множество состояний X счетно, например, Z . Простейший пример случайного блуждания на Z задается вероятностями p_{ij} скачков из i в j : 1) $p_{i,i-1} = p > 0, p_{i,i+1} = q = 1 - p > 0$. Можно представлять частицу, которая в момент $t = 0$ находится в точке $x(0) = n$, а траектория частицы $x(0), x(1), \dots, x(t), \dots$ и динамика мер определяется как и для конечных цепей. При этом $x(t+1) - x(t) = \pm 1$. А вероятности отрезков траектории длин N определяется как

$$P_N(x(0) = i_0, x(1) = i_1, \dots, x(N) = i_N) = p_{i_0i_1}p_{i_1i_2}\dots p_{i_{N-1}i_N}$$

и они согласованы в очевидном смысле

$$P_N(x(0) = i_0, x(1) = i_1, \dots, x(N) = i_N) = \sum_i P_{N+1}(x(0) = i_0, x(1) = i_1, \dots, x(N) = i_N, x(N+1) = i)$$

Можно представлять это движение частицы иначе. Пусть $\xi_1, \xi_2, \dots, \xi_t, \dots$ - последовательность случайных величин, каждая из которых равна ± 1 с вероятностью q, p соответственно. При этом они взаимно-независимы, то есть для любого t и любой последовательности u_1, \dots, u_t из ± 1

$$P(\xi_1 = u_1, \dots, \xi_t = u_t) = \prod_{k=1}^t P(\xi_k = u_k)$$

Правая часть совпадает с вероятностью траектории $x(0), x(1), \dots, x(t)$ частицы, если $u_k = x(k) - x(k-1)$ для всех $k = 1, \dots, t$. И при этом сумма $S_n = \sum_{k=1}^n \xi_k$ случайных величин ξ_k совпадает с $x(t) - x(0)$ при $t = n$. И мы имеем как выше закон больших чисел и ЦПТ для движения частицы.

Теперь более тонкие вопросы. Обозначим μ_n среднее время достижения точки 0 частицей с $x(0) = n$, и пусть p_n - вероятность того, что выйдя из n частица достигнет точки 0 когда-нибудь. Точнее

$$p_n = \lim_{N \rightarrow \infty} p_n(N),$$

где $p_n(N)$ - вероятность того, что частица попадет в 0 хотя бы в один из моментов $t \leq N$.

Пусть $n > 0$, и мы изменим вероятности скачков из точки 0 на $p_{00} = p, p_{01} = q$, и конечно тогда $p_{0,-1} = 0$, то получим блуждание на Z_+ .

$$d = -p+q < 0 \implies \mu_n < \infty, \rho_n = 1; \quad d = 0 \implies \mu_n = \infty, \rho_n = 1; \quad d > 0 \implies \mu_n = \infty, \rho_n < 1$$

Классификация общих счетных цепей Маркова конечно более сложна, где например может возникать новое явление - существование стационарного распределения. Но общий характер можно представить уже на примере нашего блуждания на Z_+ , и дается следующим утверждением:

1) если снос $d = q - p < 0$, то $p_n = 1, \mu_n < \infty$, для всех n . В этом случае говорят о положительно возвратном (или эргодическом) процессе. При этом существует стационарная (инвариантная) вероятностная мера $\pi_n > 0$, то есть такая, что $\pi P = \pi$. При этом $\pi_n \rightarrow 0$ экспоненциально быстро при росте n . Интуитивно, частица редко отходит далеко от точки 0;

2) если снос $d = 0$, то $\mu_n = \infty, p_n = 1$. Интуитивно, частица бесконечно часто возвращается в 0, но промежутки между возвращениями становятся все более длинными. Стационарной меры нет;

3) если снос $d = q - p > 0$, то $\mu_n = \infty$ и $p_n < 1$ стремятся к нулю экспоненциально быстро. Стационарной меры нет. Интуитивно, частица уходит в бесконечность согласно закону больших чисел.

Существуют три подхода к решению подобных задач: 1) ЗБЧ и ЦПТ как выше, 2) Аналитические методы - рекуррентные уравнения и их решение методом производящих функций, преобразований Фурье, см. ???, 3) метод функций Ляпунова, см. ???

8.4 Скейлинги и стохастические дифференциальные уравнения

8.4.1 От дискретного времени к непрерывному

Начнем с формального определения, используя язык полугрупп: $(N \times N)$ матрица $H = (\lambda_{ij})$ с элементами

$$\lambda_{ij} \geq 0, i \neq j, \lambda_{ii} = - \sum_{j:j \neq i} \lambda_{ij}$$

называется генератором матричной полугруппы

$$U^t = e^{Ht} = \sum_{n=0}^{\infty} \frac{(Ht)^n}{n!},$$

где ряд сходится для любого t . Очевидно также полугрупповое свойство:

$$U^{t+s} = U^t U^s$$

Докажем сначала, что все матрицы U^t являются стохастическими. Достаточно доказать это для всех достаточно малых t , и воспользоваться полугрупповым свойством чтобы доказать для всех t . Если $\lambda_{ij} > 0$ для всех $i \neq j$, то положительность для малых Δt следует из

$$U_{ij}^{\Delta t} = \lambda_{ij} \Delta t + o(\Delta t), i \neq j, U_{ii}^{\Delta t} = 1 + \lambda_{ii} \Delta t + o(\Delta t) \quad (58)$$

Без этого предположения, неотрицательность элементов U^t следует из их непрерывной зависимости от λ_{ij} . Условие нормировки следует из того, что все степени H сохраняют свойство самого H - нулевую сумму элементов каждого ряда.

Если $p(0) = \{p_x(0), x \in X\}$ - начальная мера, то любая компонента вектора $p(t) - p(0)U^t$ является аналитической функцией t . Теперь надо определять разные события и подсчитывать их вероятности. Например, элементы U_{ij}^t будем понимать как вероятности события - быть в j в момент t , при условии что в момент 0 мы были в i . Более сложные события: фиксируем конечное число моментов времени $0 \leq t_1 < \dots < t_k$ и определим вероятности (конечномерные распределения)

$$P(\omega(t_1) = i_1, \dots, \omega(t_k) = i_k) = \sum_i p_i(0) U_{ii_1}^{t_1} U_{i_1 i_2}^{t_2 - t_1} \dots U_{i_{k-1} i_k}^{t_k - t_{k-1}} \quad (59)$$

события, где мы были в эти дискретные моменты. Другие события можно определить как пределы более простых событий. Например, как определить и найти вероятность события, что в течение времени t мы все время находились в состоянии i , иначе говоря что в течение времени t не было скачков в другие состояния. Определим это как предел распределений (59)

$$\lim_{N \rightarrow \infty} P(\omega(\frac{kt}{N}) = i, k = 1, 2, \dots, N) = \lim_{N \rightarrow \infty} (1 + \lambda_{ii} \frac{t}{N} + o(\frac{t}{N}))^N = e^{\lambda_{ii} t}$$

А вероятность, что до момента t мы были в состоянии i , а в интервале времени $(t, t + \Delta t)$ уже нет, по полугрупповому свойству, будет

$$e^{\lambda_{ii} t} (1 - e^{\lambda_{ii} \Delta t}) = -\lambda_{ii} e^{\lambda_{ii} t} \Delta t + o(\Delta t) = \int_t^{t+\Delta t} f(t) dt + o(\Delta t),$$

где функция $f(t) = -\lambda_{ii}e^{\lambda_{ii}t}$ называется экспоненциальной плотностью распределения момента первого скачка из состояния i . Заметим, что вероятность того, что такой скачок когда либо случится, равна 1.

В качестве пример рассмотрим случай (процесс Пуассона), когда $X = \{0, 1\}$ и

$$H = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix},$$

который описывает детерминистическую эволюцию в “случайные моменты времени”. Процесс прыгает из 0 в 1 и обратно с одной и той же “интенсивностью” λ , то есть плотность распределения длин t интервалов между скачками равна $f(t) = -\lambda e^{\lambda t}$.

Доказать, что вероятность того, что на заданном конечном интервале времени было бесконечное число скачков, равна нулю.

8.4.2 От дискретного пространства к континууму

Белый шум Пусть каждой точке $k \in Z$, сопоставлена гауссова случайная величина $\xi_\epsilon(k)$ причем все эти величины одинаково распределены с

$$E\xi_\epsilon(k) = 0, E\xi_\epsilon^2(k) = \epsilon$$

Каждому отрезку $I = (a, b) \subset R$ сопоставим случайную величину

$$\xi_\epsilon(I) = \sum \xi_\epsilon(k),$$

где сумма по k таким, что $k\epsilon \in I$. При $\epsilon \rightarrow 0$ случайные величины $\xi_\epsilon(I)$ сходятся по распределению (то есть распределения сходятся) к некоторой случайной величине $\xi(I)$, имеющей гауссово распределение с $E\xi(I) = 0, E\xi^2(I) = b - a$. Также для любой интегрируемой функции $f(x)$ определим

$$\xi_\epsilon(f) = \sum f(k\epsilon)\xi_\epsilon(k)$$

В пределе $\epsilon \rightarrow 0$ это сходится по распределению к гауссовой величине $\xi(f)$ с

$$E\xi(f) = 0, D\xi(f) = \lim D\xi_\epsilon(f) = \int f^2(x)dx$$

а вся система этих гауссовых величин имеет корреляционные функции

$$E\xi(f)\xi(g) = \left(\int, g \right) = \int fgdx$$

Броуновское движение (винеровский процесс) Рассмотрим случайное блуждание $\xi_\epsilon(t), \epsilon > 0$, на $Z_\epsilon = \{k\epsilon : k \in Z\}$ с $\xi_\epsilon(t) = 0$, а $t = \tau\epsilon, \tau \in Z$, которое делает скачки на $\pm\epsilon$ с вероятностью $\frac{1}{2}$ за периоды времени ϵ . А потом перейти к пределу. Но лучше сначала определить блуждание с дискретным временем $k\epsilon, k \in Z$, на R . Именно, если частица в момент времени $k\epsilon$ находится в точке x , то в момент $(k+1)\epsilon$ она будет в точке $x + \xi_k$, где ξ_k гауссовы, взаимно независимы и

$$E\xi_k = 0, D\xi_k = \epsilon$$

Предел называется броуновским движением и обозначается $w(t)$ и называется иногда стандартным. Таким образом, для интервалов $I = (t_1, t_2), t_1 < t_2$, величины $\xi(I) = w(t_2) - w(t_1)$ гауссовы с нулевым средним, дисперсией $|I|$ и независимы для непересекающихся интервалов.

Процесс $\sigma w(t)$ будет таким же процессом, но с дисперсией σ^2 .

Броуновское движение со сносом Оно определяется как

$$\xi(t) = \xi_0 + \mu t + \sigma w(t) \iff d\xi(t)$$

или простейшим стохастическим уравнением

$$d\xi(t) = \mu dt + \sigma dw(t), \xi(0) = \xi_0$$

При этом интегрирование последнего делается стандартным определением интеграла

$$\int_0^t \mu dt + \int_0^t \sigma dw(t) = \lim_{N \rightarrow \infty} (\mu t + \sigma \sum_{k=1}^N (w(\frac{kt}{N}) - w(\frac{(k-1)t}{N}))) = \mu t + w(t) - w(0)$$

Стохастические интегралы Стохастический интеграл Ито от детерминированной функции $f(t)$ есть гауссова случайная функция и может быть подсчитан интегрированием по частям. Например,

$$\int_0^t f(s)dw(s) = \tag{60}$$

Стохастический интеграл от случайной функции $\eta(t)$ формально определяется как предел некоторых частичных сумм, например для $0 = s_0 < s_1 < \dots < s_N = t$,

$$\int_0^t \eta(s)dw(s) = \lim \sum_{k=1}^{k=N} \eta(s_{k-1})(w(s_k) - w(s_{k-1}))$$

но предел может иметь иной смысл. Например, если $\eta(s) = w(s)$, то предел в обычном смысле не существует, но если воспользоваться простой алгебраической формулой

$$\sum_{k=1}^{k=N} w(s_{k-1})(w(s_k) - w(s_{k-1})) = \frac{1}{2}w^2(s_N) - \frac{1}{2} \sum_{k=1}^{k=N} (w(s_k) - w(s_{k-1}))^2,$$

то последняя сумма сходится (по распределению) к t

$$\int_0^t w(s)dw(s) = \frac{1}{2}w^2(t) - \frac{1}{2}t$$

Уравнение теплопроводности Простая выкладка показывает, что гауссово распределение

$$G(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$

при $t \neq 0$ удовлетворяет уравнению теплопроводности

$$\frac{d}{dt}G = \frac{1}{2} \frac{d^2}{dx^2}G$$

Удовлетворяет этому уравнению также функция

$$u(x, t) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} f(y)dy = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} f(y)dy \tag{61}$$

При этом $u(x, 0) = f(x)$, а если $f(x) \geq 0$, то правая часть имеет смысл плотности распределения частицы, которая при $t = 0$ имеет случайное расположение с плотностью $f(x)$.

Процесс Орнштейна-Уленбека Определяется уравнением

$$d\xi(t) = -k\xi(t)dt + fdt + \sigma dw(t) \quad (62)$$

что есть упрощенный вариант **геометрического броуновского движения**, определяемого уравнением

$$d\xi(t) = \mu\xi(t)dt + \sigma\xi(t)dw(t)$$

Явное решение уравнения (62) имеет следующий вид. Можно считать $f = 0$, так как от f можно избавиться заменой $\xi = \eta + \frac{f}{k}$. Тогда сделаем следующую замену $q(\xi(t), t) = \xi(t)e^{kt}$, что даст

$$dq = ke^{kt}\xi dt + e^{kt}d\xi = e^{kt}(k\xi dt + d\xi) = e^{kt}\sigma dw(t)$$

Откуда интегрируя по $[0, t]$ получим как

$$q(t) - q(0) = \sigma \int_0^t e^{ks} dw(s) \iff \xi(t) = e^{-kt}\xi(0) + \sigma \int_0^t e^{k(s-t)} dw(s) = e^{-kt}\xi(0) + \frac{\sigma}{\sqrt{2k}} w(1 - e^{2kt})$$

Кусочно детерминированные Марковские процессы

9 Квантовые

9.1 На конечном множестве

Самосопряженность и унитарность Стандартное Любое комплексное линейное пространство L размерности N , можно отождествить с множеством комплексных функций $\psi(i)$ на конечном множестве $\{1, \dots, N\}$. Вводя скалярное произведение и норму

$$(\psi_1, \psi_2) = \sum_{i=1}^N \psi_1(i)\psi_2^*(i), \quad \|\psi\| = \sqrt{(\psi, \psi)}$$

мы делаем это пространство метрическим с расстоянием $\rho(\psi_1, \psi_2) = \|\psi_1 - \psi_2\|$, которое называется гильбертовым пространством $l_2 = l_2(\{1, \dots, N\})$. Функции $e_k = e_k(j) = \delta_{kj} = 1$, если $k = j$, и 0 если $k \neq j$, образуют его базис. Более того, они ортогональны и нормализованы: $(e_k, e_l) = \delta_{kl}$. Говорят, что они образуют **ортонормированный базис**.

Спектральная теорема для самосопряженных операторов Пусть A линейный оператор в L , тогда линейный оператор A^* в L называется сопряженным к A , если для всех ψ_1, ψ_2

$$(A^*\psi_1, \psi_2) = (\psi_1, A\psi_2)$$

Легко показать существование и единственность такого оператора для любого A , а также что его матрица в ортогональном базисе будет (a_{ji}^*) , если a_{ij} - матрица A в этом базисе.

Оператор A называется **самосопряженным**, если $A = A^*$, то есть если $a_{ij} = a_{ji}^*$ для всех i, j . Как и любой линейный оператор, самосопряженный оператор A имеет собственный вектор ψ_1 (пусть его норма равна 1) с некоторым собственным значением λ_1 . Это собственное значение вещественно так как

$$(\lambda_1\psi_1, \psi_1) = (A\psi_1, \psi_1) = (\psi_1, A\psi_1) = (\psi_1, \lambda_1\psi_1) \implies \lambda_1 = \lambda_1^*$$

Обозначим L_1^\perp - множество всех ψ , ортогональных к ψ_1 , то есть $(\psi, \psi_1) = 0$. Тогда L_1^\perp - линейное подпространство размерности $N - 1$. Оно инвариантно относительно A , так как $(\psi_1, A\psi) = (\psi_1, \lambda_1\psi) = 0$, если $(\psi_1, \psi) = 0$.

Забудем теперь о ψ_1 и продолжим аналогичную процедуру в L_1^\perp , Мы построим ψ_2, ψ_2 , и далее по индукции построим собственные векторы ψ_1, \dots, ψ_N - ортогональные и имеющие вещественные собственные значения $\lambda_1, \dots, \lambda_N$. Тогда в этом ортогональном базисе A диагонален, то есть будет оператором умножения в $l_2\{1, \dots, N\}$.

Говорят, что самосопряженный оператор имеет **простой спектр**, если кратности всех собственных значений равны 1 (то есть число собственных значений равно N).

Унитарные операторы Линейный оператор U называется **унитарным**, если выполнено одно из следующих эквивалентных условий:

- U обратим и $U^{-1} = U^*$,

b)

$$(U\psi_1, U\psi_2) = (U^*U\psi_1, \psi_2) = (\psi_1, \psi_2)$$

Таким образом он сохраняет скалярные произведения, а значит ортогональность и норму. Благодаря последнему, его собственные значения принадлежат единичному кругу, то есть имеют вид $e^{i\lambda}$ с вещественными λ , а собственные вектора ортогональны. Даже больше, если ψ ортогонален собственному значению ψ_1 , то $U\psi$ также ортогонален ψ_1 , так как $(\psi_1, U\psi) = (U^{-1}\psi_1, \psi) = (e^{-i\lambda}\psi_1, \psi) = 0$. Поэтому существует $H = H^*$ такой что

$$U = e^{iH}$$

Матрицы e^{itH} , $t \in R$, образуют **унитарную группу**. В квантовой физике (конечно в бесконечномерном варианте) эта группа определяет квантовую динамику (где t - время) а H называется Гамильтонианом.

Операторная (или матричная) алгебра

След След $Tr A$ квадратной матрицы A определяется как сумма ее диагональных элементов. Простейшие свойства следа:

$$Tr(AB) = \sum_{i,j} a_{ij}b_{ji} = \sum_{i,j} b_{ij}a_{ji} = Tr(BA),$$

откуда следует и инвариантность относительно выбора базиса

$$Tr(CAC^{-1}) = Tr(C^{-1}CA) = Tr A$$

Состояния Рассмотрим множество всех $(N \times N)$ -матриц $\mathbf{M} = \mathbf{M}_N$. Оно образует **алгебру**, то есть линейное пространство с умножением, удовлетворяющим следующей аксиоме **дистрибутивности**

$$A(B + C) = AB + AC, (B + C)A = BA + CA$$

Любой линейный функционал $L = L(M)$ на этой алгебре может быть записан в виде

$$L(M) = Tr(AM)$$

для некоторой матрицы A . Действительно, он определен своими значениями λ_{ij} на базисных матрицах E_{ij} со всеми нулевыми элементами кроме a_{ij} . Определив матрицу $\Lambda = (\lambda_{ij})$, для любой матрицы $M = (m_{ij})$ имеем:

$$L(M) = \sum_{i,j} \lambda_{ij}m_{ij} = \sum_{i,j} \lambda'_{ji}m_{ij} = \sum_j \sum_i \lambda'_{ji}m_{ij} = Tr \Lambda' M$$

Кроме того. любой линейный функционал может быть записан в виде

$$L(M) = \sum_i (Mg_i, g_j),$$

где g_i - некоторые векторы, а сумма конечна.

Самосопряженный оператор называется **положительным** (**положительно определенным**) $A > 0$, и **неотрицательным** $A \geq 0$, если все его собственные значения положительны (неотрицательны). Или: $A \geq 0$ эквивалентно любому из следующих условий

$$\exists B : A = BB^* \Leftrightarrow (A\psi, \psi) \geq 0, \forall \psi \in R^N$$

Состояние ϕ на \mathbf{M} есть линейный функционал на \mathbf{M} со следующими свойствами: 1) положительность, то есть $\phi(AA^*) \geq 0$ для всех $A \in \mathbf{M}$, 2) нормировка - $\phi(E) = 1$, где E - единичная матрица в \mathbf{M} . Примеры: **чистые** состояния определяются одним вектором $\psi \in C^N$

$$\phi(A) = (A\psi, \psi),$$

где $\|\psi\|^2 = (\psi, \psi) = 1$. Остальные состояния называются **смешанными**.

Любое состояние может быть представлено в виде

$$\phi(A) = Tr(\rho A), \quad (63)$$

где $\rho \in \mathbf{M}$ и более того $\rho \geq 0, Tr\rho = 1$. Докажем положительность:

$$\rho = CC^* \rightarrow Tr(\rho BB^*) = Tr(CC^* BB^*) = Tr(C^* BB^* C) = Tr((B^* C)^* B^* C) \geq 0$$

Обратно: если $Tr(\rho A)$ есть состояние, то $\rho \geq 0$. Ясно что ρ для данного состояния единственно.

Для чистого состояния $\rho = P_\psi$, где P_ψ - ортогональный проектор на нормализованный вектор ψ , то есть $P_\psi x = (x, \psi)\psi$.

Любое смешанное состояние может быть записано как

$$\rho(A) = \sum_i (A\psi_i, \psi_i),$$

где сумма конечна и

$$\sum_i \|\psi_i\|^2 = \sum_i (\psi_i, \psi_i) = 1$$

Действительно, диагонализуем

$$\rho = U^{-1} \left(\sum c_i P_i \right) U,$$

где P_i - ортогональный проектор на единичный вектор f_i . Полагая $\phi_i = \sqrt{c_i} f_i$, мы имеем

$$Tr(\rho A) = Tr\left(\left(\sum c_i P_i\right) U A U^{-1}\right) = \sum (\phi_i, U A U^{-1} \phi_i) = \sum (U^{-1} \phi_i, A U^{-1} \phi_i)$$

Квантовая эволюция Чистые квантовые состояния обычно представляются нормированными векторами ψ . А если они рассматриваются как комплексные функции $\psi(x)$ на конечном (или счетном) множестве X , то они называются **волновыми функциями**. Всегда предполагается, что числа $\psi(x)\psi^*(x) = |\psi(x)|^2$ могут интерпретироваться как вероятности того, что система находится в точке x .

Наблюдаемые (физические переменные величины) представляются операторами. Средним значением наблюдаемой A в состоянии ψ называется число

$$\langle A \rangle_\psi = (A\psi, \psi)$$

Их эволюция во времени может быть определена двумя эквивалентными представлениями.

1) **Представление Шредингера**, где во времени изменяются состояния, а наблюдаемые не меняются. Пусть заданы самосопряженный оператор H (называемый **Гамильтонианом**) и унитарная группа $U^t = e^{itH}$, задающая эволюцию вектора $\psi(0)$ (то есть чистого состояния $\rho(0)$) как

$$\psi(t) = U^t \psi(0),$$

Тогда имеет место уравнение (называемое **уравнением Шредингера**)

$$\frac{d}{dt} \psi(t) = iH \psi(t), \rho(t) = U^{-t} \rho(0) U^t$$

2) **Представление Гейзенберга**, где состояния не меняются, но меняются наблюдаемые. Любая унитарная группа U^t определяет группу автоморфизмов алгебры операторов

$$\alpha_t(A) = A_t = U^t A U^{-t}$$

и удовлетворяет **уравнению Гейзенберга**

$$\frac{d}{dt} A_t = it[H, A_t]$$

Средние значения изменяются в обоих представлениях одинаково

$$(A\psi(t), \psi(t)) = (\alpha_t(A)\psi, \psi)$$

9.2 Квантовые частицы

9.3 Квантовые поля

References

- [1] Hausdorff F.. Set theory. 2nd edition. 1962.
- [2] Landau E. Foundations of analysis. Chelsey Publ. 1951.
- [3] Bourbaki N. A panorama of pure mathematics (as seen by N. Bourbaki), 1982.
- [4] Vinberg E. B. A Course in Algebra. 2003.
- [5] Brualdi R., Cvetkovic D. A Combinatorial approach to matrix theory and its applications. 2009.
- [6] Gantmakher F. R. Theory of matrices. 2004. Moscow. (Chapter 15).
- [7] Dollard J., Friedman Ch. Product Integration with Application to differential equations. 1979. Cambridge.
- [8] Slavik A. Product Integration, its history and applications. 2007. Prague.
- [9] Malyshev V.A. Non-relativistic classical mechanics of point particles: shortest elementary introduction. Structure of Mathematical Physics, 2018, № 1, pp. 121-157.
- [10] Malyshev V.A. Classical Microscopic Electrodynamics: short introduction. Structure of Mathematical Physics, 2018, № 1, pp. 159-184.