МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Физический факультет

Кафедра высшей математики физического факультета

А. П. Ульянов

ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА ДЛЯ СТУДЕНТОВ-ФИЗИКОВ

Часть 1. Миникурс

Учебное пособие
по курсу основ математического анализа

Новосибирск
~~2013~~ 2016

УДК: 510
ББК: В14я73-1 + В151.54я73-1
У 517

Ульянов А. П. Основы математического анализа для студентов-физиков. Часть 1. Миникурс. Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2013.

ISBN ??

Пособие содержит конспект лекций, прочитанных автором для студентов 1-го курса физического и геолого-геофизического факультетов НГУ. Ввиду большого объёма курса и разнообразия материала пособие разделено на несколько частей. Часть 1 нацелена на быстрое овладение наиболее востребованными навыками. Она включает следующие темы: введение; производные и интегралы; комплексные числа; разложения по степеням; теоремы о среднем.

Пособие предназначено для студентов первого курса физического и геолого-геофизического факультетов НГУ.

Рецензент
проф. В.А. Александров

Учебное пособие подготовлено в рамках реализации Программы развития НИУ-НГУ на 2009-2018 гг.

Версия от 23 января 2016 г.

Глава 0. СТАРТОВЫЕ ПОЗИЦИИ

лекция 1
03.09.15

0.1. Вещественные числа

Десятичные дроби. Развитие понятия числа имеет свою долгую историю. Шаг от натуральных чисел к рациональным, или дробям, сделан в глубокой древности. Отрицательные числа менее наглядны и потому приживались с трудом. Однако ещё шумерские, древнекитайские и тем более древнегреческие геометры знали об иррациональных числах, а также отыскивали рациональные приближения весьма высокой точности к самым важным из них, например, π, $\sqrt{2}$, золотое сечение $\varphi = (1+\sqrt{5})/2$.

Постепенно универсальным средством записи приближённо известных вещественных чисел стали десятичные дроби. (В Китае с IV века до н.э., а в Европе лишь с конца XVI века.) Десятичные дроби вам известны со школы, как и соответствие вещественных чисел и точек на прямой. При кажущейся простоте здесь спрятано немало логических тонкостей, однако в начале курса не следует углубляться в них. Stevin 1585

Числовые множества и системы. Понятие множества лежит в основе формальной математики с конца XIX века. Его не определяют через другие понятия. **Множеством** называют любую совокупность объектов, набор, семейство, класс, рассматриваемых как единое целое.

Пример. Перечислим основные числовые системы и напомним стандартные обозначения этих множеств.

- множество $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ всех натуральных чисел;
- множество $\mathbb{Z}$ всех целых чисел;
- множество $\mathbb{Q}$ всех рациональных чисел;
- множество $\mathbb{R}$ всех вещественных чисел;
- множество $\mathbb{C}$ всех комплексных чисел.

Часть математиков считают и учат, что натуральные числа начинаются с 1. Ну и пусть! Эти разногласия не имеют значения в анализе.

Объекты в множестве называют его **элементами**; говорят, что они **принадлежат** множеству или **лежат** в нём, что тоже самое, и используют специальный значок принадлежности $\in$. Любая часть множества называется его **подмножеством**, что указывают значком включения $\subseteq$, стилизованным из знака неравенства $\leqslant$. Используют также значок $\subset$; иногда при этом равенство исключено, а иногда по неаккуратности допустимо.

Поскольку мы занимаемся вещественным анализом, комплексные числа появляются лишь в эпизодах. Значит, назовём **числовым множеством** любое $X \subseteq \mathbb{R}$. Хорошие числовые множества изображают в виде множеств точек на прямой, но самые плохие сложно даже изобразить.

Один из основных способов определения подмножества некоторого множества — указать условие на элементы, выполнение которого соответствует принадлежности подмножеству. Для этих целей имеется специальная форма записи; например, $\{x \in \mathbb{R} \mid x > 0\}$ есть множество положительных вещественных чисел, обозначаемое через $\mathbb{R}_+$ или $\mathbb{R}_{>0}$.

Промежутки. Множество всех вещественных чисел, лежащих между двумя границами $a, b \in \mathbb{R}$, где $a < b$, называют конечным **промежутком**. Они бывают четырёх типов. Классифицирующий признак — включены ли в промежуток сами точки a и b:

- **замкнутым** промежутком или **отрезком** называют
$$[a, b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\};$$
- **открытым** промежутком или **интервалом** называют
$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\};$$
- **полуоткрытым** промежутком или **полуинтервалом** называют
$$[a, b) = \{x \in \mathbb{R} \mid a \leqslant x < b\},$$
$$(a, b] = \{x \in \mathbb{R} \mid a < x \leqslant b\}.$$

В этих обозначениях квадратная скобочка указывает на включение соответствующей границы, а круглая — на её исключение. Для интервалов встречается менее удачное обозначение $]a, b[$, и аналогично для полуинтервалов. Способов графического изображения промежутков также несколько.

рисунки

Замкнутые и открытые бесконечные промежутки
$$[a, +\infty), \quad (-\infty, b], \quad (a, +\infty), \quad (-\infty, b)$$
определяются аналогично, но с одной границей вместо двух. Наконец, $\mathbb{R} = (-\infty, +\infty)$ также считается бесконечным промежутком. Бесконечности исключаются, поскольку они не лежат в $\mathbb{R}$.

0.2. Функции

Эволюция понятия функции. Функции — главный объект изучения в математическом анализе. Понятие функции формировалось не одну сотню лет. Полезно задуматься об основных этапах его развития:

(0) функция как плоская линия (график);
(1) функция как зависимость, выражаемая формулой;
(2) функция как любая зависимость между переменными;
(3) функция как отображение числовых множеств.

С античности важной частью геометрии было изучение плоских линий. Восхищались тогда каждой по отдельности. Аналитическая геометрия Декарта резко расширила возможности, наводя мосты между Descartes 163 линиями и формулами, доступными для алгебраических манипуляций. Самого слова *функция* в математическом смысле ещё не было, хоть бы и в латыни. В тогдашних важнейших задачах о линиях мы теперь видим задачи о функциях, а в их решениях — зачатки интегрального и дифференциального исчислений, составивших начала анализа.

В этих исчислениях понятие функции постепенно уходило от представления о линиях. Функцией Лейбниц назвал зависимость, выража- Leibniz 1673 емую формулой: она всегда служила определённой математической цели, то есть выполняла полезную функцию. Заданные формулами функции заняли центральное место в анализе с выходом эпохального Euler 1748 учебника Эйлера.

Хотя представления о функции как о любой зависимости между переменными величинами восходят к XIV веку, во времена Эйлера они ещё вызывали ожесточённые споры. Даже определяя понятие функции как произвольной зависимости, все ведущие математики фактически думали лишь об аналитических выражениях (формулах). Дальнейшее развитие математического анализа изменило точку зрения и Dirichlet 183 функция стала любым правилом, связывающим переменные. Формальное выражение это определение получило лишь с введением языка теории множеств: функция стала отображением.

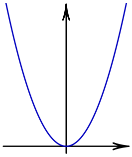
Определение. **Функция** $f\colon X \to Y$ состоит из двух множеств, **области определения** X и **области значений** Y, и правила, сопоставляющего каждому элементу x из X элемент y из Y. Это правило обозначают через $y = f(x)$ или $x \mapsto f(x)$.

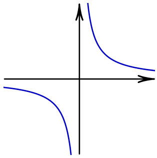
При этом y называют **образом** x, а x — **прообразом** y.

Расширение самого понятия привело к появлению в математике различных функций, которые шокируют интуицию, воспитанную на примерах из ранних эпох, или ныне из школьной программы. За небольшими исключениями, всё изучаемое в нашем курсе добыто наукой задолго до построения формализма теории множеств, часто называемого «основаниями математики». Наиболее ходовые приложения анализа в естественных науках главным образом используют функции как формулы (теория) и как приближённо известные зависимости (результаты наблюдений и экспериментов).

Отсюда проистекает печально известный конфликт интересов между анализом как стройной и строгой математической дисциплиной и анализом как инструментом, необходимым для постижения основ естественных наук. На переднем крае науки иногда бывает нужна вся сила современного анализа, но студенту, сталкивающемуся с анализом впервые, формальный подход несёт много бед. Поэтому в лекционной части нашего курса во главу угла ставится развитие интуитивного понимания; с формальной точки зрения остаётся немало пробелов и неточностей, до заполнения которых студенту ещё надо дозреть.

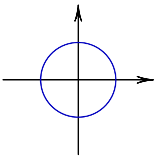
Формулы и графики. Итак, понятие функции развилось из понятия линии на плоскости. Теперь, наоборот, линии определяют посредством функций, как их графики. Рассмотрим простые примеры явлений, о которых следует помнить в этой связи.

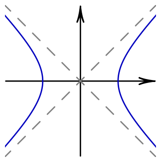




Примеры. Подробно изучаются в школе функции $y = x^2$ и $y = 1/x$. Здесь **явно** указана зависимость y от x. Графиками являются парабола и гипербола соответственно.

Второй пример показывает, что функция $f(x)$ может быть определена не для всех вещественных значений x, а лишь на части их; множество тех чисел, где функция определена, и называют её областью определения.

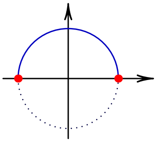


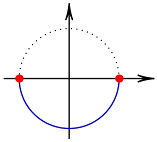


Примеры. Уравнения

$$x^2 + y^2 = 1, \qquad x^2 - y^2 = 1$$

задают линии — соответственно окружность и гиперболу. Но задают ли они функции? Нет: в обоих случаях имеется такое значение x, которому соответствует более одного значения y.



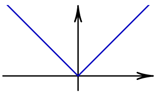


В явной же формуле $y = \pm\sqrt{1-x^2}$ знак $\pm$ указывает на выбор одного из двух значений. Таким образом, функция соответствует не всему графику, а лишь его **ветви**. Иногда эту деталь не упоминают, говоря о выражениях наподобие $x^2+y^2=1$ как о **неявных** функциях. Condorcet 17
Для более сложных неявных функций явное выражение зависимости, как правило, невозможно.

Пример. Функция $y = |x|$ определяется двумя различными явными формулами в зависимости от знака x:

$$|x| = \begin{cases} x & \text{при } x \geqslant 0, \\ -x & \text{при } x < 0. \end{cases}$$

Упражнение. *Задайте $|x|$ единой явной формулой.*



Большие возможности для задания линий открываются с использованием двух функций $x(t)$, $y(t)$ для координат точки. В таком случае говорят о **параметрическом** задании линии, а также параметрическом задании функции $y(x)$. Тогда, как и при неявном задании, а то и чаще, бывает несколько значений y для одного значения x. Нам придётся иметь дело и с такими функциями.

примеры?

Операции с функциями. Арифметические операции применяются к функциям, как и к их значениям:

$$(f+g)(x) = f(x) + g(x)$$

и аналогично для вычитания, умножения и деления. При этом может уменьшаться область определения.

Композицией функций называют подстановку одной функции вместо аргумента другой. Например, если $f(x) = x^2$ и $g(x) = \sin x$, то $f(g(x)) = \sin^2 x$ и $g(f(x)) = \sin(x^2)$. При этом тоже может уменьшаться область определения.

Важный частный случай композиции — функции вида $f(\alpha x + \beta)$ с постоянными α и β. График такой функции получается из графика

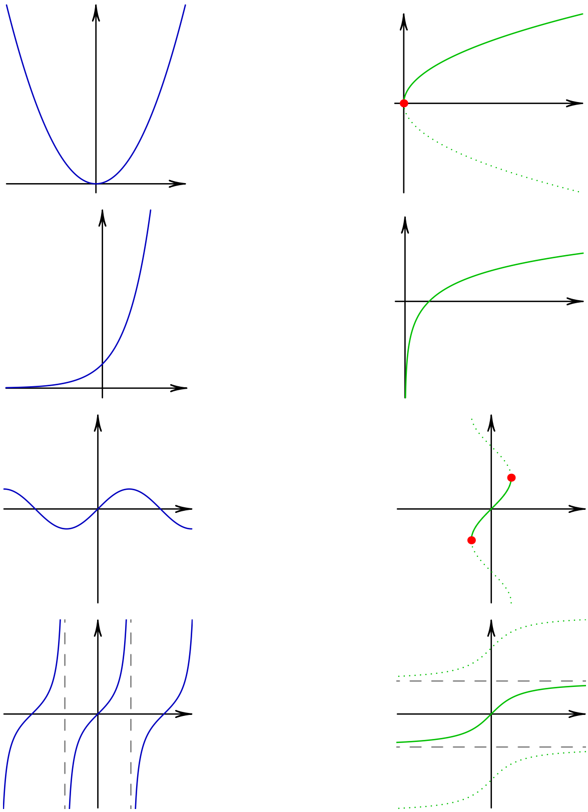
$f(x)$ преобразованиями горизонтального сдвига, сжатия или растяжения, а также отражения, которые нужно чётко усвоить.

рисунок

Функция $f^{-1}(x)$ **обратна** к функции $f(x)$, если

$$f^{-1}(f(x)) = f(f^{-1}(x)) = x.$$

Тогда и $f(x)$ обратна к $f^{-1}(x)$. Здесь также следует следить за областями определения и выбором ветви. Примеры пар обратных функций, известных со школы: x^2 и $\sqrt{x}$; a^x и $\log_a x$; $\sin x$ и $\arcsin x$; $\operatorname{tg} x$ и $\operatorname{arctg} x$.



лекция 2
07.09.15

Элементарные функции. Функцию называют **элементарной**, если она представляется конечным числом операций из **основных элементарных** функций. Нужно разобраться, какие функции считать основными, а также все ли перечисленные выше операции допускать.

Список основных элементарных функций безусловно должен включать **степенные** функции $f(x) = x^n$ для натуральных n. Тогда в элементарные функции попадут все полиномы и **рациональные** функции, то есть частные полиномов. Минималист, желающий сократить список основных элементарных функций, может без ущерба, но не надолго, оставить из степенных единственную функцию $f(x) = x$.

Свободно пользуясь переходом к обратным функциям, мы получаем **алгебраические** функции. Здесь есть новые степенные функции рисунки

$$f(x) = x^{1/n} = \sqrt[n]{x}; \qquad f(x) = x^{p/q} = \sqrt[q]{x^p}.$$

Более того, здесь есть функции наподобие

$$f(x) = \sqrt{\frac{1-x}{1+x}}; \qquad f(x) = \sqrt[3]{1+\sqrt{x}} + \sqrt[3]{1-\sqrt{x}}.$$

Кажется, это не так плохо, но серьёзная проблема таится в том, что лишь простейшие алгебраические уравнения решаются явно. Например, обратная функция к весьма простой полиномиальной функции $y = f(x) = x^5 + x$ явного алгебраического выражения не имеет и выражается лишь неявным уравнением $y^5 + y = x$.

Поэтому все алгебраические функции включать в элементарные с точки зрения анализа неразумно. Включим лишь степенные в число основных и далее запретим огульное взятие обратных функций.

Неалгебраические функции называют **трансцендентными** (от латинского *превосходящими силу*, а именно, силу алгебры). Элементарными среди них считаются: **показательные** функции a^x; **логарифмические** функции $\log_a x$; **тригонометрические** функции $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$ и им обратные; наконец, **гиперболические** функции $\operatorname{sh} x$, $\operatorname{ch} x$, $\operatorname{th} x$, $\operatorname{cth} x$ и им обратные. При этом в основные элементарные функции достаточно зачислить только по одной из показательных и логарифмических, рисунки обычно e^x и $\ln x$; тригонометрическую при большом желании можно оставить любую одну; гиперболические же и вовсе лишние, ибо они очень просто выражаются через e^x, а им обратные — через $\ln x$ и $\sqrt{x}$.

Итак, основными элементарными функциями будем считать

$$x^{p/q},\ e^x,\ \ln x,\ \sin x,\ \cos x,\ \operatorname{tg} x,\ \arcsin x,\ \operatorname{arctg} x,$$

а каждую элементарную функцию получать из них с помощью конечного числа арифметических операций и композиции. Избыточность списка в тригонометрической части вызвана соображениями удобства.

Исторически, эти основные функции имеют различную природу, что отражается и на их изучении в школе. Позже мы увидим в математическом анализе глубинную причину важности элементарных трансцендентных функций и фактически единый источник их появления (реванш минимализма). Школьные знания останутся нужны для проработки примеров и закрепления навыков.

С другой стороны, выделение элементарных функций в отдельный класс является пережитком эпохи Эйлера. Огромное количество неэлементарных функций, применяемых теперь в науке и технике, называют **специальными**. Принципиально они не отличаются от элементарных трансцендентных, а всего лишь имеют более сложные свойства. Очень немногие специальные функции появятся в этом курсе.

Упражнение. *Элементарны ли функции* $|x|$, $x^{\sqrt{2}}$, x^x*?*

0.3. Непрерывность и точки разрыва

Наивная непрерывность. Возьмём любую основную элементарную функцию $f(x)$ и число a из области её определения. Интересно поведение $f(x)$ при приближении x к a. Во всех случаях можно сказать, что $f(x)$ приближается к $f(a)$, и написать $f(x) \approx f(a)$ при $x \approx a$. Распространена другая запись, математически более точная и в то же время внешне нехитрая: $f(x) \to f(a)$ при $x \to a$.

Это важнейшее свойство функций называют **непрерывностью** в точке a. Подчеркнём наивность такого представления, поскольку настоящее формальное определение непрерывности мы рассмотрим позже; проблема заключена в слове «приближается» и разрешается лишь скрупулёзным изучением понятия предела:

$$f(a) = \lim_{x \to a} f(x).$$

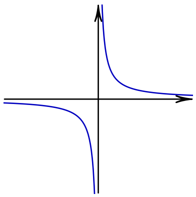
Тем не менее, наивные представления способны принести ощутимые и содержательные результаты, быстро нужные физикам. Это соответствует и историческому развитию предмета: непрерывность движения лежала в основе рассуждений Ньютона при создании основ анализа вместе с его немедленными первыми приложениями к физике; сходна была и философская позиция Лейбница. Формализацией анализа, и понятия предела в первую очередь, математики серьёзно озаботились только в XIX веке, хотя необходимость её осознавали оба основателя.

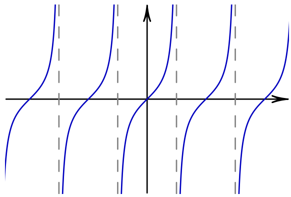
Теорема. *Всякая элементарная функция непрерывна в каждой точке своей области определения.*

Схема доказательства. Отдельно проверяется непрерывность основных элементарных функций. Далее иерархическое строение элементарных функций позволяет заключить, что не только основные, но и вообще все элементарные функции непрерывны во всех тех точках, где они определены. Для этого нужно убедиться, что если две функции непрерывны, то непрерывны также их сумма, разность, произведение, частное и композиция. Единственная проблема возникает при делении на функцию, приближающуюся к нулю, но в таких точках значение дроби не определено. $\square$

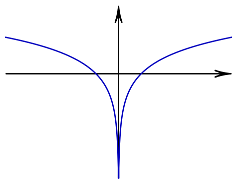
Всякая функция, заданная единой формулой, элементарна, а значит, непрерывна. Однако функция, взятая наугад как отображение, не будет непрерывна ни в одной точке! Непрерывность следует осознать как свойство хороших функций. Здесь ярок конфликт интересов теории и практики: доказывая теорему, непрерывность абстрактно взятой функции необходимо выводить из каких-то других свойств; имея же конкретную функцию, для установления её непрерывности следует сперва присмотреться к её строению и выявить шаги, могущие нарушить непрерывность. Такие операции в анализе есть.

Происхождение разрывов. Точку, в которой функция **разрывна**, то есть не является непрерывной, называют её **точкой разрыва**. При этом функция должна быть определена вблизи изучаемой точки, иначе и вопроса не возникает. Как появляются такие точки при работе с элементарным материалом?





Точки разрыва элементарных функций обычно возникают в дробях, если знаменатель приближается к нулю. Таковы разрывы функций $1/x$ и $\operatorname{tg} x = \sin x/\cos x$. Возможны также фокусы с областью определения, как в случае функции $\ln(x^2)$, имеющей разрыв в точке $x = 0$.



Другой несложный способ: задать функцию разными формулами слева и справа от определённой точки (примеры ниже).

Упражнение. *Почему эти способы не дают противоречия с утверждаемой выше непрерывностью всякой элементарной функции на всей своей области определения?*

Весьма общий способ получения новых функций, в том числе разрывных в некоторых точках своей области определения, доставляет предельный переход по другой переменной:

$$f(x) = \lim_{y \to b} g(x, y).$$

Такой переход может быть спрятан внутри разных операций, изучением которых собственно и занят математический анализ; при этом нередко вопрос сохранения непрерывности или появления разрыва является центральным. До изучения самих таких предельных переходов нам ещё далеко, но давайте безотлагательно познакомимся с классификацией изолированных точек разрыва.

Типы точек разрыва. Накладывая дополнительное условие $x < a$ либо $x > a$, получают **односторонние пределы** функции при $x \to a$. Для них имеются специальные обозначения $f(a-0)$ и $f(a+0)$, а также названия **предел слева** и **предел справа**.

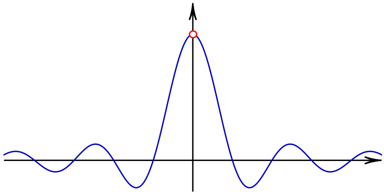
Пример. Функция $f(x) = 1/x$ не имеет предела при $x \to 0$. Односторонние же пределы существуют, хотя они бесконечны: $f(0-0) = -\infty$ и $f(0+0) = +\infty$.

Точки разрыва традиционно разбивают на типы согласно односторонним пределам при $x \to a$ безотносительно значения при $x = a$:

- в точке **устранимого** разрыва оба односторонних предела существуют, конечны и равны;
- в точке разрыва **первого рода** оба односторонних предела существуют, конечны, но не равны;
- в точке разрыва **второго рода** хотя бы один из односторонних пределов бесконечен или не существует.

Устранимые разрывы так названы потому, что непрерывности в такой точке можно добиться, доопределяя или переопределяя функцию в этой точке: просто полагаем $f(a) = f(a \pm 0)$. Часто такие новые значения весьма интересны; сравните с грядущим определением производной.

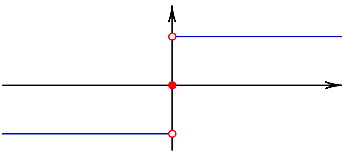
Пример. Функция $\frac{\sin x}{x}$ имеет устранимый разрыв в точке $x = 0$. Рисунок растянут в 10 раз по вертикали.



Пример. Часто встречается разрывная в нуле функция $\operatorname{sgn} x$, вычисляющая знак числа x:

$$\operatorname{sgn} x = \begin{cases} 1 & \text{при } x > 0, \\ 0 & \text{при } x = 0, \\ -1 & \text{при } x < 0. \end{cases}$$

Разрыв здесь имеет характер конечного **скачка**, то есть первого рода.



Пример. Нижняя и верхняя целые части числа — монотонно неубывающие целочисленные функции, определённые правилами

$$n \leqslant x < n + 1 \Rightarrow \lfloor x \rfloor = n,$$
$$n < x \leqslant n + 1 \Rightarrow \lceil x \rceil = n + 1.$$

Эти функции имеют разрывы первого рода в каждой целой точке.

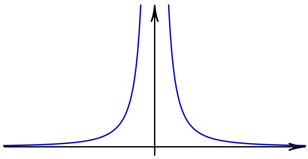
В данных примерах значение функции в точке скачка определено. Видно, что оно может как совпадать с одним из односторонних пределов, так и отличаться от обоих. В следующем примере, оставленном в качестве упражнения, значение в точке скачка не определено.

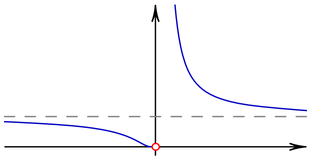
Упражнение. *Даже элементарная функция может иметь скачок. Исследуйте функцию* $f(x) = \operatorname{arctg} \frac{1}{x}$.

пилы?

Точки разрыва второго рода более разнообразны. Остаётся сожалеть, что все они собраны в один так называемый род. Мы уже обратили внимание на то, что функция $1/x$ имеет неравные бесконечные односторонние пределы при $x \to 0$. Несколько разновидностей разрыва возникают у её композиций с другими основными функциями.

Пример. Функция $1/x^2$ имеет равные бесконечные односторонние пределы при $x \to 0$.

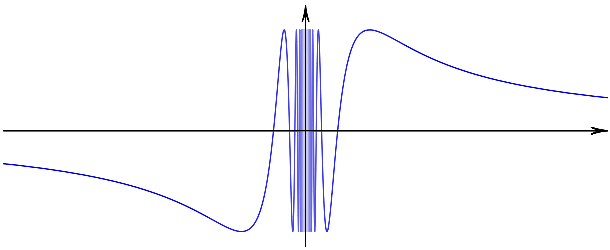




Пример. Функция $e^{1/x}$ имеет конечный предел слева и бесконечный предел справа при $x \to 0$.

Пример. В курсах анализа популярен пример функции $f(x) = \sin \frac{1}{x}$. Её график плох вблизи одной точки $x = 0$, при приближении к которой эта функция не имеет предела ни с одной из сторон, бешено колеблясь между значениями -1 и 1. Значит, даже элементарная функция может иметь настолько «плохой» график, что его трудно назвать линией.

правка графика 07.09.15



Последняя функция весьма экзотична на первый взгляд. Однако, точка разрыва у неё единственна, как и в остальных примерах этой серии. Есть функции и похуже, у которых точки разрыва сгущаются.

Dirichlet 1837

Пример. **Функция Дирихле**, заданная правилом

$$f(x) = \begin{cases} 1 & \text{если } x \text{ рационально}, \\ 0 & \text{если } x \text{ иррационально}, \end{cases}$$

разрывна в каждой точке!

0.4. Предпосылки возникновения анализа

Активизация наук в Европе XVI века принесла много новых математических задач. Достижения древнегреческих и арабских математиков вышли в печатном виде, и вскоре десятки европейских учёных начали развивать их методы и изобретать новые. В математике, прежде полностью излагавшейся словами, появлялось всё больше обозначений, теперь привычных всем со школы.

Три важнейшие задачи. Многие разрозненные работы первой половины XVII века считаются предвестниками появления дифференциального и интегрального исчисления, то есть начал анализа. Среди важнейших геометрических задач того времени выделим два типа: рисунки

(A) Уравнение касательной к известной линии в точке.

(B) Площадь под известной линией на некотором отрезке.

По востребованности, сначала в астрономии и навигации, а потом и в самой математике, им не уступала ещё одна группа задач:

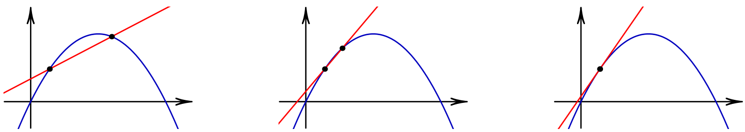
(C) Вычисление значений функций с высокой точностью.

Прежде всего речь о функциях $\sqrt{x}$, $\sin x$ и $\ln x$.

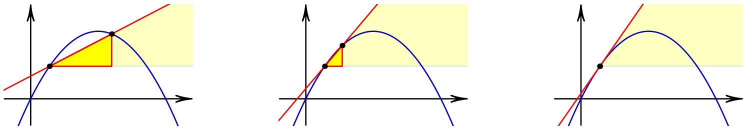
Решение задачи о касательной привело к понятиям производной и дифференциала, а решение задачи о площади — к понятию интеграла. Первоначально эти задачи и вызревавшие понятия не были связаны друг с другом. Обнаружение их связи стало большим толчком к прогрессу.

Физически, или вернее механически, задача о касательной интерпретируется как нахождение скорости движения $v(t)$ по известному положению $x(t)$, а задача о площади — как нахождение пройденного пути по известной скорости.

Задача о касательной. Выберем хорошую линию. Прямую, пересекающую линию в двух точках, называют **секущей**. Представим себе, что эти две точки приближаются друг к другу и, наконец, сливаются в одну. При этом секущая становится **касательной**. Говорят также, что касательная есть предельное положение секущей.



Задача состоит в том, чтобы по известному уравнению линии определить уравнение касательной в любой её точке. Для этого достаточно знать угол наклона касательной. Тангенс угла наклона секущей выражают геометрически как отношение малых **приращений** координат. Предел этой величины при стремлении приращений к нулю и есть искомый тангенс угла наклона касательной.



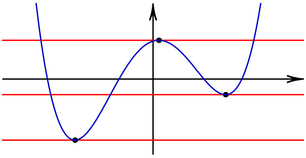
Зависимость его от точки касания на графике функции называют **производной** этой функции:

$$\frac{f(x)-f(a)}{x-a} \to f'(a).$$

Ньютон и даже много позже Эйлер определяли производную, обходясь без явного понятия предела. Все функции тогда были непрерывными, но отношение приращений как функция $x \to a$ имеет разрыв в точке $x = a$, часто устранимый. Число, устраняющее разрыв, и считали значением $f'(a)$ производной в этой точке, хотя сам термин *производная* появился позже, у Лагранжа.

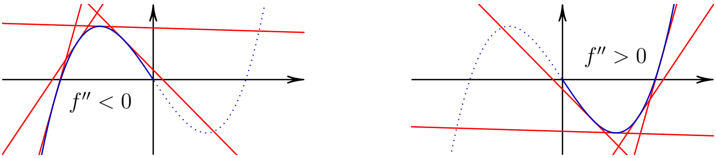
В кинематической интерпретации, если линия представляет график переменной величины, то отношение приращений выражает среднюю скорость за соответствующий промежуток времени. Предельным переходом мы получаем мгновенную скорость. Итак, производная переменной величины выражает скорость её изменения (относительно независимой переменной, которая не обязана быть именно временем).

Задача о касательной к графику функции содержит в себе как частный случай задачу о максимумах и минимумах этой функции, поскольку в таких точках касательные горизонтальны, а мгновенная скорость изменения функции нулевая. Сюда же примыкает задача об определении промежутков возрастания и убывания функции: зная производную, мы решаем задачу, находя промежутки, на которых производная соответственно положительна и отрицательна.

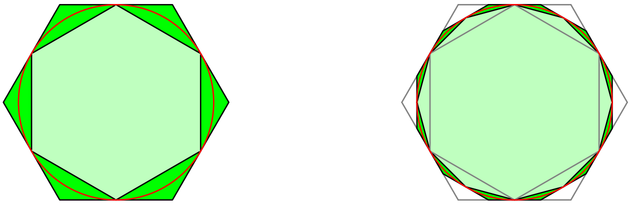


Скорость изменения скорости называют ускорением. Если скорость выражена как производная координаты, то ускорение есть производная от производной; её для краткости называют **второй производной**.

Знаку второй производной соответствует направление отклонения графика исходной функции от касательной вблизи точки касания. Поэтому рассмотрение второй производной необходимо для более точного построения графика, когда оно выполняется вручную.



Задача о площади. Вычислением площадей криволинейных фигур занимались ещё античные математики; больше всех своими успехами в этом деле прославился Архимед. Древний метод исчерпывания состоит в приближении фигуры с любой точностью вписанными и описанными многоугольниками и содержит в себе идею предельного перехода. Метод был строго обоснован Евдоксом, но громоздок. Широко известно приближение таким способом площади единичного круга, то есть числа π.



Упражнение. *Вычислите площади правильных 12-угольников, вписанных в единичную окружность и описанных около неё.*

Метод неделимых, также известный ещё Архимеду, намного мощнее, но долго вызывал опасения ввиду отсутствия строгого обоснования. К середине XVII века успехи метода неделимых были впечатляющими в вычислении не только площадей, но и объёмов, длин линий, площадей поверхностей, центров тяжести тел. Из него выросла идея бесконечно малой величины.

Приближения площади под графиком функции используют разбиения фигуры на тонкие вертикальные полоски. В простейшем варианте каждую полоску заменяют на прямоугольник «почти той же» высоты.

Площади этих прямоугольников складывают и получают приближённое значение площади фигуры. Для повышения точности приближения переходят к более тонким полоскам. Практические вычислительные методы используют вместо прямоугольников более качественные приближения иными фигурами с точно известными площадями.

Полученную сумму называют **интегральной** от латинского корня, означающего «целый, единый». Предел интегральных сумм при повышении их точности называют определённым интегралом. Ранее задачу вычисления площади криволинейной фигуры называли задачей о квадратуре этой фигуры, поэтому и сейчас выражения, содержащие интегралы, не приведённые или не приводящиеся к другому, элементарному виду, называют выражениями в квадратурах.

Другие геометрические и механические задачи этого типа также приводят к интегральным суммам иного вида. Мы к ним ещё вернёмся.

Приближённые вычисления. Ни у одной функции, кроме рациональных, и полиномиальных в частности, нельзя точно найти любое значение даже при точно известных (рациональных) аргументах. Случается, что отдельные значения найти проще, или же можно вычислить некоторые с желаемой точностью один раз и занести в таблицу. С помощью таблицы затем можно приблизить значения функции в других точках, если заменить функцию на более простую, чаще всего полиномиальную, имеющую те же значения в табличных точках. Этот метод называется **интерполяцией**.

Для повышения точности интерполяции приходится повышать степень приближающего полинома. При малом отклонении x от табличной точки, в которую мы сейчас для упрощения формул поместим начало координат, слагаемые x^n высоких степеней становятся очень маленькими. Отсюда возникла идея, что любую (хорошую) функцию можно представить суммой вида

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

с числовыми коэффициентами a_k. Такие суммы с *бесконечным* количеством слагаемых называют **степенными рядами**, а такие представления функций — их разложениями в степенной ряд.

Эта идея разложения, вместе с лёгкостью дифференцирования и интегрирования степенных слагаемых, принесла множество открытий, составивших самые начала анализа и быстро прославивших его. Вскоре мы познакомимся с важнейшими разложениями и общей формулой, ограничиваясь приближёнными конечными представлениями. Теории бесконечных рядов мы коснёмся позже.

Глава 1. ПРОИЗВОДНЫЕ И ИНТЕГРАЛЫ

1.1. Правила отыскания производных

Определение. **Производной** функции $f(x)$ в точке a называется предел лекция 2+ 07.09.15

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{\delta \to 0} \frac{f(a + \delta) - f(a)}{\delta}.$$

Само понятие предела пока остаётся на интуитивном уровне, оказывающемся впрочем вполне достаточным для простых вычислений.

Вместо малого $\delta = x - a$ в данном определении по старой традиции часто пишут малое Δx, называемое приращением аргумента. Числитель же представляет собой приращение функции, так что определение можно записать иначе:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}(x).$$

Поэтому Лейбниц ввёл удобное обозначение $f' = \frac{df}{dx}$.

Каков здесь смысл числителя и знаменателя по отдельности? Вопрос долго вызывал ожесточённые споры не только математиков, но и философов. Чёткий математический ответ мы увидим позже.

Текущая наша задача: понять, как находить производные элементарных функций. Стратегия состоит в том, чтобы сводить задачу для сложной функции к более простым при помощи нескольких общих правил, а производные основных элементарных функций, через которые выразятся все прочие, найти и крепко запомнить. Эта стратегия действительно работает, и в результате получится

Теорема. *Производная всякой элементарной функции элементарна.*

Оставшая часть этого раздела, за исключением минимума примеров, составляет «доказательство» этой теоремы, состоящее из многих шагов. Результатом каждого шага является важное правило; все правила необходимо знать и научиться споро применять в задачах.

Арифметические правила. Правило дифференцирования суммы,

$$(f(x) + g(x))' = f'(x) + g'(x),$$

выражает тот факт, что скорость изменения суммы величин равна сумме скоростей их изменения. Формально, поскольку

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

это правило прямо следует из арифметических свойств предела.

Абсолютно аналогично правило дифференцирования разности:

$$(f(x) - g(x))' = f'(x) - g'(x).$$

Leibniz 1677

Правило дифференцирования произведения (**правило Лейбница**):

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

Это правило легко представить наглядно. Считаем множители длинами сторон переменного прямоугольника, скорости изменения которых известны. С какой скоростью изменяется его площадь $S(x)$?

С точностью до самого маленького прямоугольничка имеем

$$\frac{\Delta S}{\Delta x} = \frac{\Delta f}{\Delta x} \cdot g + f \cdot \frac{\Delta g}{\Delta x},$$

откуда предельным переходом и получаем правило Лейбница.

Применяя правило для произведения к $1 = g(x) \cdot \frac{1}{g(x)}$, находим вспомогательное правило:

$$0 = 1' = g'(x) \cdot \left(\frac{1}{g(x)}\right) + g(x) \cdot \left(\frac{1}{g(x)}\right)' \quad \Longrightarrow \quad \left(\frac{1}{g(x)}\right)' = -\frac{g'(x)}{g(x)^2}.$$

Подставляя теперь $1/g$ вместо g в правило для произведения, найдём правило для частного:

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}.$$

Упражнение. *Напишите и обоснуйте правило дифференцирования тройного произведения* $f(x) \cdot g(x) \cdot h(x)$.

Производные степенных функций. Прежде всего, производная постоянной функции равна нулю, ибо равно нулю её приращение.

Производная функции $f(x) = x$ равна 1, ибо $\Delta f = \Delta x$.

Возьмём $f(x) = x^2$. Тогда

$$\Delta f = (x + \Delta x)^2 - x^2 = 2x \cdot \Delta x + (\Delta x)^2.$$

Значит, $\frac{\Delta f}{\Delta x} = 2x + \Delta x$ и в пределе при $\Delta x \to 0$ видим $(x^2)' = 2x$. Этот же ответ моментально получается по правилу дифференцирования произведения:

$$(x \cdot x)' = x' \cdot x + x \cdot x' = 1 \cdot x + x \cdot 1 = 2x.$$

Аналогично для $f(x) = x^3$ найдём

$$(x \cdot x \cdot x)' = x' \cdot x \cdot x + x \cdot x' \cdot x + x \cdot x \cdot x' = 3x^2.$$

Аналогично для $f(x) = x^n$ с натуральным показателем n получим n слагаемых, в каждом из которых x' умножается $n-1$ раз на x. Поэтому

$$(x^n)' = nx^{n-1}.$$

Для $f(x) = x^{-n}$ с положительным целым n найдём

$$f'(x) = \left(\frac{1}{x^n}\right)' = -\frac{(x^n)'}{(x^n)^2} = -nx^{-n-1}.$$

Значит, предыдущее правило верно для всех целых n. Вместе с арифметическими правилами, его достаточно для дифференцирования полиномиальных и рациональных функций.

Теперь подставим в определение $f(x) = \sqrt{x}$. Небольшим трюком

$$\Delta f = \sqrt{x + \Delta x} - \sqrt{x} = \frac{x + \Delta x - x}{\sqrt{x + \Delta x} + \sqrt{x}}$$

преодолеваем проблемы при вычислении предела. Значит, $f'(x) = \frac{1}{2\sqrt{x}}$, что также подходит в предыдущее правило при $n = 1/2$.

Ниже мы установим правило

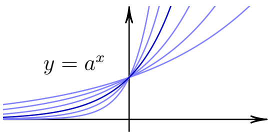
$$(x^\alpha)' = \alpha x^{\alpha - 1}$$

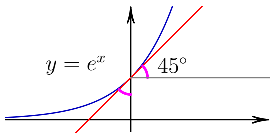
для любых вещественных показателей α, поняв сперва, что такое x^α.

Производная показательной функции. Для $f(x) = a^x$ находим

$$\frac{\Delta f}{\Delta x} = \frac{a^{x+\Delta x} - a^x}{\Delta x} = a^x \cdot \frac{a^{\Delta x} - 1}{\Delta x}.$$

Предел выделенной дроби при $\Delta x \to 0$ по определению равен $f'(0)$.





Выбор загадочного числа e в качестве основания стандартной показательной функции вызван тем, что именно при этом основании производная в точке $x = 0$ равна 1. Отсюда несложно найти её значение и при прочих основаниях (упражнение). Поэтому

$$(e^x)' = e^x, \quad (a^x)' = a^x \ln a.$$

Производные тригонометрических функций. Запишем приращения функций $\cos x$ и $\sin x$, пользуясь формулами для косинуса и синуса суммы. После перегруппировки получим

$$\frac{\Delta(\cos x)}{\Delta x} = \cos x \cdot \frac{\cos \Delta x - 1}{\Delta x} - \sin x \cdot \frac{\sin \Delta x}{\Delta x},$$
$$\frac{\Delta(\sin x)}{\Delta x} = \sin x \cdot \frac{\cos \Delta x - 1}{\Delta x} + \cos x \cdot \frac{\sin \Delta x}{\Delta x}.$$

Значит, чтобы найти $(\cos x)'$ и $(\sin x)'$, достаточно вычислить выделенные пределы — между прочим, по определению равные искомым производным, взятым при $x = 0$. Ситуация напоминает экспоненту; неслучайность этого нам довольно скоро откроется.

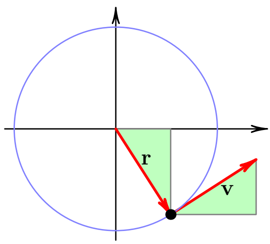
Значения ключевых пределов интуитивно ясны из картинки:

$$\frac{\cos \Delta x - 1}{\Delta x} \to 0, \quad \frac{\sin \Delta x}{\Delta x} \to 1 \quad \text{при } \Delta x \to 0.$$

Поэтому

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

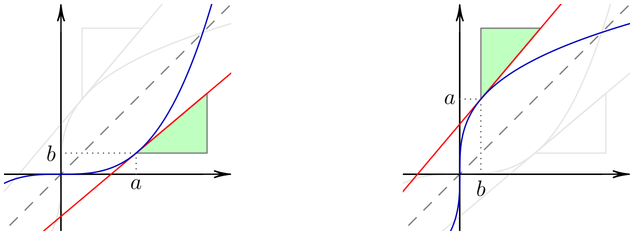
Есть и другие способы найти нужные пределы и сами производные. Например, рассмотрим равномерное вращение вокруг начала координат. Пара функций $(\cos t, \sin t)$ параметризует единичную окружность, проходимую точкой с единичной угловой скоростью. Вектор скорости $\mathbf{v}(t)$ имеет единичную длину и перпендикулярен радиус-вектору $\mathbf{r}(t) = (\cos t, \sin t)$. Значит, $\mathbf{v}(t) = \pm(-\sin t, \cos t)$. Правильный знак определяем, глядя на направление вращения. Компоненты скорости по осям равны искомым производным.



Выразив $\operatorname{tg} x$ и $\operatorname{ctg} x$ как отношения $\sin x$ и $\cos x$, по правилу дифференцирования частного находим

$$(\operatorname{tg} x)' = 1 + \operatorname{tg}^2 x = 1/\cos^2 x, \quad (\operatorname{ctg} x)' = -1 - \operatorname{ctg}^2 x = -1/\sin^2 x.$$

Производная обратной функции. Графики функции $y = f(x)$ и обратной к ней функции $y = f^{-1}(x)$, или $x = f(y)$ говоря иначе, зеркально симметричны относительно прямой $y = x$. лекция 3 10.09.15



Отсюда геометрически находим, что тангенсы углов наклона касательных в точках (a, b) и (b, a), где $b = f(a)$, обратны друг другу. Поскольку эти тангенсы равны производным $f'(a)$ и $(f^{-1}(b))'$, получаем правило дифференцирования обратной функции:

$$(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}.$$

С помощью этого правила находятся производные всех оставшихся основных элементарных функций.

$$(e^x)' = e^x \quad\Longrightarrow\quad (\ln x)' = \frac{1}{e^{\ln x}} = \frac{1}{x};$$

$$(\operatorname{tg} x)' = 1 + \operatorname{tg}^2 x \quad\Longrightarrow\quad (\operatorname{arctg} x)' = \frac{1}{1 + \operatorname{tg}^2(\operatorname{arctg} x)} = \frac{1}{1 + x^2};$$

$$(\sin x)' = \cos x \quad\Longrightarrow\quad (\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}.$$

Производная композиции функций. Это правило называют также правилом дифференцирования сложной функции или правилом цепочки (англ. *chain rule*). Со штрихами оно записывается как

$$f(g(x))' = f'_g(g(x)) \cdot g'(x).$$

Нижний индекс g указывает на то, что соответствующая производная берётся, будто всё выражение $g(x)$ является одной буквой, от которой функция f и зависит. В обозначениях Лейбница, $f'(x) = \frac{df}{dx}$ и т. п., правило цепочки запоминается проще: Leibniz 1677

$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}.$$

Цепочка нередко бывает длиннее; скажем, для $f(g(h(x)))'$ получится
$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dh} \cdot \frac{dh}{dx}.$$

Не вдаваясь в формальности, правило цепочки можно интуитивно объяснить тем, что на уровне приращений равенство
$$\frac{\Delta f}{\Delta x} = \frac{\Delta f}{\Delta g} \cdot \frac{\Delta g}{\Delta x}$$
справедливо, лишь бы $\Delta g \neq 0$.

Вот простой пример применения правила цепочки:
$$(\sin^2 x)' = 2\sin x \cdot \frac{d(\sin x)}{dx} = 2\sin x \cos x.$$

Возьмём теперь показательную функцию x^α, где число α, наконец, произвольное вещественное, а $x > 0$. Разумно определять эту функцию как $e^{\alpha \ln x}$. Применим правило цепочки:
$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \cdot (\alpha \ln x)' = x^\alpha \cdot \alpha/x = \alpha x^{\alpha - 1}.$$

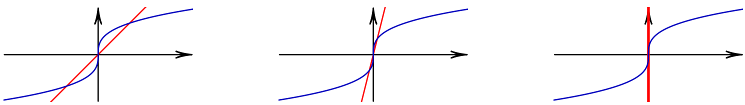
Упражнение. *Дана хорошая функция $f(x)$. Найдите функцию $u(x)$, для которой $u'(x) = f'(x)/f(x)$.*

Упражнение. *С помощью функции $u(x)$, найденной в предыдущем упражнении, получите формулы*
$$\frac{(f(x)g(x))'}{f(x)g(x)} = \frac{f'(x)}{f(x)} + \frac{g'(x)}{g(x)}; \qquad \frac{(f(x)/g(x))'}{f(x)/g(x)} = \frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}.$$
Сравните с правилами дифференцирования произведения и частного, которые не используйте при выводе.

Особые случаи. Рассмотрим ситуации, когда непрерывная функция не имеет производной в одной точке $x = a$. Это могут быть:

- точки, где касательная вертикальна;
- граничные точки;
- точки излома;
- точки с колебательным подходом;
- точки с сочетанием этих проблем, включая точки возврата.

Когда пределом хорд графика является вертикальная касательная, иногда говорят, что производная существует, но она бесконечна.



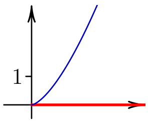
Пример. Функция $f(x) = \sqrt[3]{x}$, определённая для всех вещественных x, имеет бесконечную производную в точке $x = 0$.

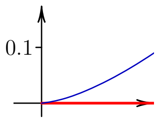
Для граничной точки a области определения функции отношение приращений

$$k(x) = \frac{f(x) - f(a)}{x - a},$$

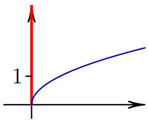
оно же наклон хорды, определено при $x \to a$ лишь с одной стороны. Тогда можно рассмотреть его односторонний предел. Если он существует и конечен, то говорят об **односторонней** производной, а именно, о производной слева $f'(a-0)$ при $x < a$ и о производной справа $f'(a+0)$ при $x > a$. Иногда для сокращения оговорок удобно включать случаи существования односторонних производных в крайних точках в дифференцируемость, но часто их исключают, рассматривая дифференцируемость только внутри области определения.

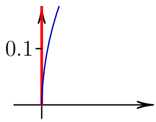
Пример. Функция $f(x) = \sqrt{x^3}$, определённая для неотрицательных x, имеет правую производную $f'(0+0) = 0$.





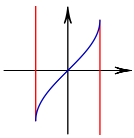
Пример. Функция $f(x) = \sqrt{x}$ имеет бесконечную правую производную $f'(0+0) = +\infty$.





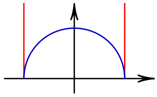
Пример. Функция $f(x) = \arcsin x$ имеет бесконечные односторонние производные на границах области определения:

$$f'(-1+0) = f'(1-0) = +\infty.$$



Пример. Функция $f(x) = \sqrt{1 - x^2}$ имеет бесконечные односторонние производные на границах области определения:

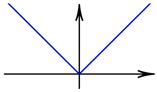
$$f'(-1+0) = +\infty, \qquad f'(1-0) = -\infty.$$



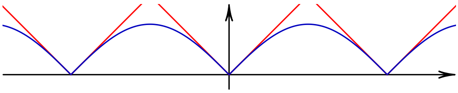
Односторонние производные могут оказаться полезны и внутри области определения, когда отношение $k(x)$ определено при $x \neq a$, но не имеет (двустороннего) предела.

Пример. Функция $f(x) = |x|$ имеет точку излома $x = 0$ с односторонними производными $f'(0 \pm 0) = \pm 1$. Отметим кстати, что

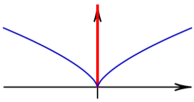
$$\operatorname{sgn} x = \tfrac{1}{2}(f'(x+0) + f'(x-0)).$$



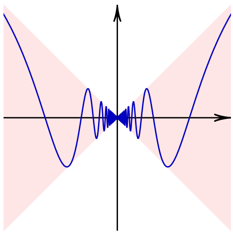
Пример. Функция $f(x) = |\sin x|$ имеет точку излома $x = \pi n$ для каждого целого числа n.



Пример. Функция $f(x) = x^{2/3}$, определённая для всех вещественных x, имеет точку возврата $x = 0$ с бесконечными, но не равными односторонними производными $f'(0 \pm 0) = \pm\infty$.



Пример. Функция $f(x) = x \sin \frac{1}{x}$, продолженная по непрерывности значением $f(0) = 0$, не имеет односторонних производных в особой точке. При $x \to a = 0$ наклон хорды колеблется от -1 до 1.



1.2. Поиск первообразных

Определение. Функция $F(x)$ называется **первообразной** функции $f(x)$, если $F'(x) = f(x)$.

Обычно говорят о первообразной на промежутке, так что условие должно быть выполнено во всех внутренних точках промежутка.

Пример. Для $f(x) = \sin x$ угадываем первообразную $F(x) = -\cos x$. Инакомыслящий способен предложить иную первообразную $1 - \cos x$, и также будет прав.

Лемма. *Все первообразные любой функции на промежутке отличаются друг от друга лишь прибавлением постоянной.*

Доказательство. Возьмём две первообразные одной и той же функции. Их разность является первообразной функции, тождественно равной нулю, а это может быть только константа. □

Упражнение. *Зачем в лемме предполагается, что область определения — промежуток?*

Определение. Совокупность всех первообразных функции $f(x)$ называется её **неопределённым интегралом**.

Неопределённый интеграл функции $f(x)$ обозначают через

$$\int f(x)\,dx = F(x) + C.$$

Происхождение этих значков и названий (интеграл и дифференциал) проясняется чуть позже с установлением связи между понятием первообразной как обратным к понятию производной и задачей о вычислении площади под графиком.

Начальная таблица первообразных. Первообразные некоторых самых простых функций подбираются обыкновенным наблюдением таблицы производных.

Например, деля равенство $(x^\alpha)' = \alpha x^{\alpha-1}$ на α и затем заменяя α на $\alpha + 1$, приходим к

$$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C.$$

Эта формула непригодна при $\alpha = -1$, но в этом случае мы находим

$$\int \frac{dx}{x} = \ln x + C.$$

Чтобы не ограничиваться здесь лишь положительными значениями x, заметим, что при $x < 0$ выполнено $\ln(-x)' = 1/x$. Значит,

$$\int \frac{dx}{x} = \ln|x| + C$$

для всех $x \neq 0$. Очень часто учебники и справочники приводят этот интеграл именно так. Однако правая часть равенства перечисляет не все первообразные для $1/x$. Причина в том, что разрыв функции при $x = 0$ позволяет прибавлять к $\ln|x|$ две различные постоянные слева и справа от разрыва. Эта проблема уходит, когда от первообразной на промежутке требуют непрерывности и даже дифференцируемости, что обычно и делают.

Первообразную показательной функции увидеть легче всего:

$$\int e^x\,dx = e^x + C; \qquad \int a^x\,dx = \frac{a^x}{\ln a} + C, \quad a > 0,\ a \neq 1.$$

Для синуса и косинуса необходимо помнить о минусе, и тут источник частых ошибок:

$$\int \sin x\,dx = -\cos x + C; \qquad \int \cos x\,dx = \sin x + C.$$

Первообразные функций $\operatorname{tg} x$ и $\ln x$ одним взглядом не находятся, да и остальных основных элементарных функций тоже. Однако стоит дополнить таблицу наиболее употребительных первообразных двумя лёгкими находками:

$$\int \frac{dx}{1+x^2} = \operatorname{arctg} x + C; \qquad \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C.$$

Общие правила. По аналогии со стратегией для производных, попытаемся найти правила, позволяющие отыскивать первообразные, разбивая функции на более простые составляющие. Составляться они могут посредством четырёх арифметических операций и композиции. Каждое из общих правил дифференцирования может соответствовать правилу интегрирования. Однако на деле не всё так хорошо.

Не вызывает затруднений лишь правило для суммы функций. Оно просто переходит в правило для суммы:

$$\int (f(x) + g(x))\,dx = \int f(x)\,dx + \int g(x)\,dx.$$

Столь же незамысловато правило для умножения на постоянную:

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

Удобно объединять эти два свойства в одно,

$$\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx,$$

называемое **линейностью** интеграла.

Из правила для композиции по сути получается правило замены переменной интегрирования. Из правила для произведения получается правило «интегрирования по частям». К ним мы сейчас перейдём. Из правила для частного ничего полезного для первообразных не получается. В итоге, в отличие от производных, первообразная элементарной функции не обязательно элементарна. Например, не существует элементарной первообразной у функции e^x/x.

Для облегчения применений в науке и технике, раньше составлялись таблицы интегралов, заполнявшие толстые справочники. Теперь эти книги в основном вытеснены компьютерными системами символьных вычислений. Однако отыскание первообразных вручную на основе оставших двух правил изобилует элементами несложного творчества и потому увлекательно и развивающе.

Замена переменной интегрирования. Интегрирование функции $f(x)$ иногда упрощается, если перейти от x к другой переменной, найдя подходящую функциональную зависимость $x = x(t)$ или $u = u(x)$. Fermat 1657

Правило для первого типа замены есть

$$\int f(x)\,dx = \int f(x(t))\,x'(t)\,dt.$$

Формально оно получается из формулы $dx/dt = x'(t)$, хотя сейчас мы ещё далеки от понимания того, что же такое dx в отдельности.

Второй тип замены является применением этого же правила в обратную сторону (со сменой букв):

$$\int f(u(x))\,u'(x)\,dx = \int f(u)\,du.$$

Его конкретный вид обычно обнаруживают, идентифицируя фрагмент $u(x)$ или $u'(x)$.

Пример. Возьмём $f(x) = (x+3)^{-2}$. Находим в таблице известных нам первообразных наиболее похожую функцию x^{-2} и делаем замену

$$u = x + 3 \quad \Longrightarrow \quad x = u - 3, \quad dx = du.$$

Тогда

$$\int \frac{dx}{(x+3)^2} = \int \frac{du}{u^2} = -\frac{1}{u} + C = -\frac{1}{x+3} + C.$$

Пример. Возьмём $f(x) = \sin 5x$. Напрашивается замена

$$u = 5x, \quad du = 5dx.$$

Тогда

$$\int \sin 5x\, dx = \tfrac{1}{5} \int \sin u\, du = -\tfrac{1}{5} \cos u + C = -\tfrac{1}{5} \cos 5x + C.$$

В обоих разобранных примерах замена задаётся линейной функцией $u(x) = \alpha x + \beta$. Это наиболее частый вид замены, поэтому полезно выписать формулу специально для него:

$$\int f(x)\, dx = F(x) + C \qquad \Longrightarrow \qquad \int f(\alpha x + \beta)\, dx = \frac{1}{\alpha} F(\alpha x + \beta) + C.$$

Пример. Возьмём $f(x) = \operatorname{tg} x = \sin x / \cos x$. Заметим по таблице производных, что $\sin x = u'(x)$ для $u(x) = -\cos x$. Тогда

$$\int \operatorname{tg} x\, dx = \int \frac{\sin x}{\cos x}\, dx = -\int \frac{d(\cos x)}{\cos x} = -\ln|\cos x| + C.$$

Как всегда, с развитием навыка некоторые промежуточные шаги не выписываются, поэтому здесь замена на u нигде явно не указана.

Пример. Возьмём $f(x) = \sqrt{1 - x^2}$. Здесь тот случай, когда подынтегральное выражение упрощается заменой

$$x = \sin t, \quad \sqrt{1 - x^2} = \cos t, \quad dx = \cos t\, dt.$$

Модуль не нужен, ибо изначально $|x| \leqslant 1$ и потому можно считать, что $|t| \leqslant \pi/2$, а тогда $\cos t \geqslant 0$.

Первообразной для $\cos^2 t$ в нашей таблице ещё нет, но помогает тригонометрия: $\cos^2 t = \frac{1}{2}(1 + \cos 2t)$. Итак,

$$\begin{aligned}\int \sqrt{1 - x^2}\, dx &= \int \cos^2 t\, dt = \tfrac{1}{2} \int (1 + \cos 2t)\, dt \\ &= \tfrac{1}{2} t + \tfrac{1}{4} \sin 2t + C = \tfrac{1}{2} \arcsin x + \tfrac{1}{2} x \sqrt{1 - x^2} + C.\end{aligned}$$

Интегрирование по частям. Перегруппируем формулу дифференцирования произведения:

$$(fg)' = f'g + fg' \qquad \Longleftrightarrow \qquad fg' = (fg)' - f'g.$$

Теперь возьмём первообразную:

$$\int fg'\, dx = fg - \int f'g\, dx.$$

Fermat 1657 Это и есть формула интегрирования по частям, хотя буквы чаще пишут иные:

$$\int u\,dv = uv - \int v\,du.$$

Название формулы объясняется тем, что для её применения нужно представить исходную функцию в виде произведения двух функций особым образом; при этом один сомножитель — одна часть — должен сразу интегрироваться, а оставшийся после такого преобразования интеграл может оказаться проще исходного.

Пример. Возьмём $f(x) = x\cos x$. Выделение сразу интегрируемой части здесь похоже на задачу о первообразной для $\operatorname{tg} x$. Берём

$$u = x,\ dv = \cos x\,dx \quad\Longrightarrow\quad du = dx,\ v = \sin x.$$

Тогда по формуле получаем

$$\int x\cos x\,dx = \int x\,d(\sin x) = x\sin x - \int \sin x\,dx = x\sin x + \cos x + C.$$

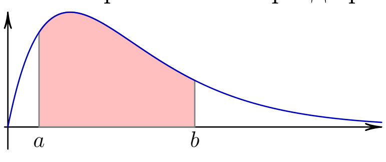
Упражнение. *Почему в проделанном вычислении нет нужды учитывать константу на первом шаге, то есть брать $v = \sin x + c$?*

Пример. Возьмём $f(x) = \ln x$. На том же уровне находим

$$\int \ln x\,dx = x\,\ln x - \int x\,d(\ln x) = x\ln x - \int x\,\frac{dx}{x} = x\ln x - x + C.$$

1.3. Определённый интеграл

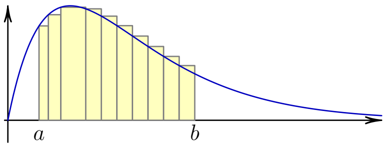
Определённый интеграл есть число, выражающее первоначально площадь криволинейной трапеции, а впоследствии и многие другие величины. Как исторически, так и в наших физических курсах, это прежде всего геометрические и механические величины. Разумеется, актуальные современные приложения гораздо разнообразнее.



Интегральные суммы. Определённый интеграл от функции f на отрезке $[a, b]$ вводят как предел интегральных сумм

$$\sum_{(a)}^{(b)} f(x_i)\,\Delta x_i.$$

Точки $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ разбивают отрезок на короткие кусочки $[x_i, x_{i+1}]$ длины $\Delta x_i = x_{i+1} - x_i$. За высоту прямоугольника, приближающего полоску криволинейной трапеции на таком кусочке, принимают $f(x_i)$. Поэтому каждое слагаемое интегральной суммы есть площадь тонкого прямоугольника.

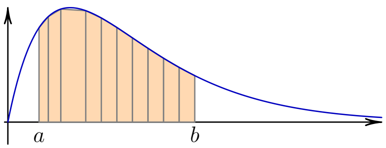


При предельном переходе устремляют к нулю наибольшую ширину полоски. В случае успеха, а для непрерывных функций он гарантирован развитой позже теорией, полученный предел обозначают через

$$\int_a^b f(x)\,dx.$$

Лейбниц (1675) первым употребил значок интеграла как стилизацию первой буквы латинского слова *summa*, а за значок дифференциала взял первую букву слова *differentia*, означающего *разность*. Термин *интеграл* ввёл Бернулли (1690). За a и b закрепились неудачные названия: это соответственно **нижний** и **верхний пределы интегрирования**; проблема тут в употреблении слова *предел* в совершенно ином смысле, чем при предельном переходе.

Приближение прямоугольниками лишь самое грубое, самое простое. Для повышения точности вычислений используют другие приближения. Следующее по простоте — метод трапеций.

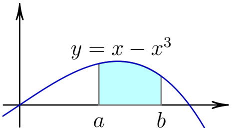


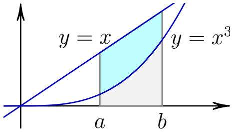
В этом курсе мы избегаем вопросов приближённого интегрирования; наша задача — изучить аналитические методы, а не численные.

Основные свойства. Определение определённого интеграла и арифметические свойства предела сразу влекут свойство **линейности**

$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx,$$

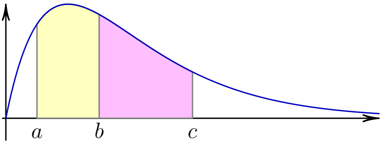
которого мы, конечно, ожидаем от площади, опираясь на принцип Кавальери.





Составляя теперь две криволинейные трапеции в одну бок о бок, видим второе естественное свойство **аддитивности**, также формально опирающееся на свойства предела:

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$



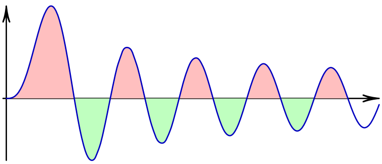
При том, ограничение $a \leqslant b \leqslant c$ можно снять, принимая соглашение

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx \quad \text{при } a > b.$$

Свойство **монотонности** утверждает, что

$$f(x) \geqslant 0 \text{ на отрезке } [a, b] \quad \Longrightarrow \quad \int_a^b f(x)\,dx \geqslant 0 \quad \text{при } a \leqslant b.$$

Если же $f(x) < 0$ на каком-то участке, то подграфик становится надграфиком, а площадь там отрицательна.



лекция 4 14.09.15

Фундаментальная теорема анализа. Наша ближайшая цель — научиться вычислять как можно больше определённых интегралов, не прибегая к построению интегральных сумм. Средства для этого даёт интегральное исчисление.

Нас будет интересовать площадь не одной фиксированной криволинейной трапеции, а её зависимость от меняющегося верхнего предела:

$$F(b) = \int_a^b f(x)\,dx.$$

рисунок

Тут уже хочется заменить b на x и написать функцию

$$F(x) = \int_a^x f(x)\,dx.$$

Однако такая запись некорректна, ибо одна буква несёт разный смысл под интегралом и вне его. Некорректности избегают, заменяя переменную интегрирования на любую свободную букву:

$$F(x) = \int_a^x f(t)\,dt = \int_a^x f(\xi)\,d\xi = \int_a^x f(\tilde{x})\,d\tilde{x} = \int_a^x f(\text{ъ})\,d\text{ъ}.$$

Происходящее здесь понятнее, если в исходной записи сперва заменить x на t и только затем b на x. К таким выражениям прилагают длинное название: **интегралы с переменным верхним пределом**.

Gregory 1668
Barrow 1670

Теорема. *Если функция f непрерывна на отрезке $[a, b]$, то*

(1) *площадь подграфика*

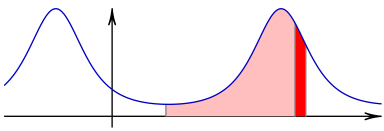
$$S(x) = \int_a^x f(t)\,dt$$

является её первообразной на этом отрезке;

(2) *для всякой её первообразной F выполнена формула*

$$\boxed{F(b) - F(a) = \int_a^b f(x)\,dx.}$$

Содержание второй части теоремы практически повсеместно называют **формулой Ньютона — Лейбница**, хотя доказательства её получили их предшественники Грегори и Барроу. Формула связывает операцию дифференцирования, обращением которой является взятие первообразной, с операцией интегрирования в смысле квадратур. Поэтому, она позволяет заменить сложный процесс с интегральными суммами на действия с первообразными в случаях, когда аналитические выражения последних известны (и не обязательно элементарны).



Доказательство. Возьмём маленькое $\delta > 0$ и рассмотрим разность

$$S(x+\delta) - S(x) = \int_a^{x+\delta} f(x)\,dx - \int_a^x f(x)\,dx = \int_x^{x+\delta} f(x)\,dx.$$

Это площадь очень узкой криволинейной трапеции ширины δ, поэтому предел отношения $\frac{1}{\delta}\big(S(x+\delta)-S(x)\big)$ при $\delta \to 0$ равен её высоте $f(x)$. Значит, $S'(x) = f(x)$.

Поскольку разность $F(x) - S(x)$ первообразных постоянна, второе утверждение теоремы следует. $\square$

подробнее

Упражнение. *Выведите основные свойства линейности, аддитивности и монотонности определённого интеграла из фундаментальной теоремы.*

Выражение для первообразной бывает громоздким, и тогда может оказаться неразумно записывать его целиком дважды лишь для того, чтобы подставить разные буквы. Часто $F(b)-F(a)$ пишут компактнее:

$$\int_a^b f(x)\,dx = F(x)\Big|_{x=a}^{x=b}.$$

Вместо вертикальной черты, или вместе с ней, иногда пишут большие скобки вокруг $F(x)$.

Ещё раз о первообразной степенной функции. Вычислим

$$\int_1^x \frac{dt}{t} \quad \text{для } x > 0.$$

Можно выбрать любую первообразную; берём $F(t) = \ln t$. По формуле не Ньютона и не Лейбница, искомое число равно $\ln x - \ln 1 = \ln x$.

Проиллюстрируем тут компактную запись:

$$\int_1^x \frac{dt}{t} = \big[\ln t\big]\Big|_{t=1}^{t=x} = \ln x - \ln 1 = \ln x.$$

Теперь при $\alpha \neq 0$ вычислим

$$\int_1^x t^\alpha \frac{dt}{t} = \frac{x^\alpha - 1}{\alpha}.$$

Чему равен предел этого выражения при $\alpha \to 0$? Мы уже встречали этот предел, только буквы входили иначе. Он равен $\ln x$. Тем самым, исключительный случай в формуле интегрирования степенной функции, приводящий к логарифму, интерпретируется как предельный.

рисунок?

Простейшие преобразования. Когда нужно вычислить конкретный определённый интеграл, лучше сперва попробовать найти первообразную, «взяв» соответствующий неопределённый интеграл. Здесь мы не будем входить в детали: это отрабатывается на семинарах и самостоятельно, причём требует значительных вложений времени.

Можно преобразовывать и определённый интеграл, но тогда необходимо следить за изменениями пределов интегрирования. При введении новой переменной $u = u(x)$ нижний предел $x = a$ заменяют на $u(x) = u(a)$, а при подстановке $x = x(t)$ приходится решать уравнение $a = x(t)$, и аналогично для верхнего предела. Всё это зачастую лишняя работа, а также дополнительный источник ошибок.

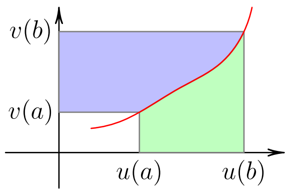
Интегрирование по частям ещё хуже в этом смысле. Значки для нижних и верхних пределов излишне загромождают саму формулу,

$$\int_a^b u(x)v'(x)\,dx = \big[u(x)v(x)\big]\Big|_{x=a}^{x=b} - \int_a^b u'(x)v(x)\,dx,$$

и что уж говорить о вычислениях. Однако, немного иной записи

$$\int_{x=a}^{x=b} u(x)\,dv(x) + \int_{x=a}^{x=b} v(x)\,du(x) = \big[u(x)v(x)\big]\Big|_{x=a}^{x=b},$$

оказывается, можно сопоставить простую картинку.



Линия здесь задана параметрически:

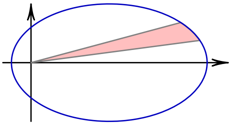
$$u = u(x), \quad v = v(x).$$

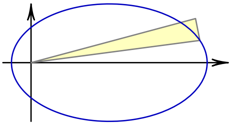
Площади над и под линией выражаются двумя интегралами из левой части формулы. Их сумма равна разности площадей прямоугольников, выражаемых произведениями из правой части.

1.4. Интегралы в геометрии и физике

Определённый интеграл применяется для вычисления не только площадей под графиками, но и многих других величин. Рассмотрим некоторые из приложений.

Площадь в полярных координатах. Найдём формулу для вычисления площади сектора, заметаемого радиус-вектором движущейся точки. Траектория задаётся в полярных координатах уравнением $\rho = \rho(\varphi)$, угол меняется в пределах $\alpha \leqslant \varphi \leqslant \beta$.





На участке от φ до $\varphi + \delta$ площадь сектора приближённо равна площади треугольника:

$$S(\varphi + \delta) - S(\varphi) = \tfrac{1}{2}\rho(\varphi)\rho(\varphi + \delta)\sin\delta + o(\delta).$$

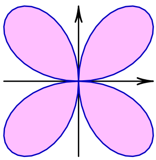
Последнее слагаемое обозначает малую величину, которой будем пренебрегать. Пользуясь эквивалентностью $\sin\delta \sim \delta$, а также непрерывностью траектории $\rho(\varphi + \delta) \sim \rho(\varphi)$, упростим формулу до

$$S(\varphi + \delta) - S(\varphi) = \tfrac{1}{2}\rho(\varphi)^2\delta + o(\delta).$$

Поэтому $S'(\varphi) = \frac{1}{2}\rho(\varphi)^2$ и тогда площадь сектора равна

$$S = \tfrac{1}{2}\int_\alpha^\beta \rho(\varphi)^2\,d\varphi.$$

Пример. Вычислим площадь, ограниченную линией $\rho = \sin 2\varphi$, позволяя значения $\rho < 0$.



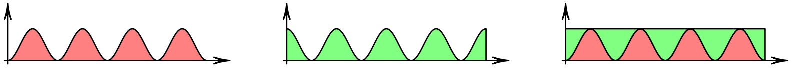
Подставив данную функцию в формулу, можно затем честно найти первообразную (она встречается часто) и подставить пределы, но можно схитрить. Поскольку $\sin x = \cos(x - \frac{\pi}{2})$, а интегрирование идёт по всему периоду 2π, имеем

$$\int_0^{2\pi} \sin^2 2\varphi\,d\varphi = \int_0^{2\pi} \cos^2 2\varphi\,d\varphi,$$

так что оба эти интеграла равны своей полусумме

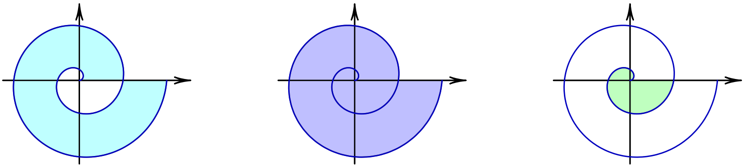
$$\tfrac{1}{2}\int_0^{2\pi} (\cos^2 2\varphi + \sin^2 2\varphi)\,d\varphi = \tfrac{1}{2}\int_0^{2\pi} d\varphi = \pi.$$

Значит, $S = \pi/2$.



Пример (спираль Архимеда). Вычислим площадь между первым и вторым витками спирали, заданной уравнением $\rho = \varphi$. Здесь нужна разность двух интегралов:

$$S = \frac{1}{2}\int_{2\pi}^{4\pi} \rho^2\, d\varphi - \frac{1}{2}\int_{0}^{2\pi} \rho^2\, d\varphi = \cdots = 8\pi^3.$$



Длина линии. Найдём теперь формулы для вычисления длины линии в декартовых и полярных координатах.

При радиус-векторе движущейся точки $\mathbf{r}(t) = (x(t), y(t))$ скорость равна $\mathbf{r}'(t) = (x'(t), y'(t))$, а по абсолютной величине

$$|\mathbf{r}'(t)| = \sqrt{x'(t)^2 + y'(t)^2}.$$

Тогда длина участка траектории от $t = a$ до $t = b$ равна

$$L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

Когда линия задана иначе, приходится сначала отыскивать способ параметризовать её, что означает представить её в виде $\mathbf{r} = \mathbf{r}(t)$.

По аналогичной формуле вычисляют длины линий в пространстве:

$$L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt.$$

Однако, длины многих весьма обычных линий как в пространстве, так и на плоскости часто не выражаются в элементарных функциях. Некоторые выражаются громоздко. Например, неэлементарные эллиптические функции появились в математике при попытках найти длину эллипса, а также возникают во многих других задачах.

Искать длину линии можно и в полярных координатах. При малом δ радиус-векторы $\mathbf{r}(\varphi)$ и $\mathbf{r}(\varphi + \delta)$ почти параллельны, поэтому длина маленького отрезка также приближённо находится по теореме Пифагора, только катеты лежат на луче из начала координат и дуге окружности с центром в начале.

Длина первого катета равна

$$|\rho(\varphi+\delta)-\rho(\varphi)| = |\rho'(\varphi)\delta| + o(\delta).$$

Длина второго равна доле длины окружности, соответствующей углу δ, а именно, $\rho(\varphi)\delta$. Значит,

$$L(\varphi+\delta)-L(\varphi) = \sqrt{\rho'(\varphi)^2+\rho(\varphi)^2}\,|\delta| + o(\delta).$$

Поэтому длина участка линии равна

$$L = \int_\alpha^\beta \sqrt{\rho'(\varphi)^2+\rho(\varphi)^2}\,d\varphi.$$

Пример (кардиоида). Вычислим длину линии, заданной уравнением $\rho = 1-\cos\varphi$. Находим

$$\sqrt{(\rho')^2+\rho^2} = \sqrt{2-2\cos\varphi} = 2|\sin(\varphi/2)|.$$

При $0 \leqslant \varphi \leqslant 2\pi$ модуль пропадает. Ответ необычно прост: $L = 8$.

Пример (астроида). Вычислим длину линии, заданной уравнением

$$x^{2/3}+y^{2/3}=1.$$

Вспоминая равенство $\cos^2 t+\sin^2 t = 1$, находим параметризацию

$$x=\cos^3 t,\quad y=\sin^3 t,\quad 0\leqslant t\leqslant 2\pi.$$

Упрощение корня из формулы для длины даст под интегралом модуль, от которого избавляемся стандартным образом:

$$L = 3\int_0^{2\pi} |\cos t\sin t|\,dt = 3\cdot 4\int_0^{\pi/2} \cos t\sin t\,dt = \dots = 6.$$

Упражнение. *Получите формулы для вычисления длины линии в цилиндрических и сферических координатах.*

Упражнение. *Найдите первообразную для функции $\sqrt{1-\cos x}$ из примера с кардиоидой. Подсказка: тут есть подвох!*

Фигуры вращения. Теперь рассмотрим две объёмные задачи. На плоскости возьмём график неотрицательной функции $y = f(x)$ на отрезке $[a, b]$ и станем вращать его в пространстве вокруг оси Ox. Образуется **фигура вращения** — поверхность с уравнением $y^2 + z^2 = f(x)^2$. Этому классу принадлежат некоторые «школьные» фигуры: цилиндры, конусы, сферы.

Чтобы вычислить объём тела вращения, по сути применим древний метод Архимеда. Разбиению отрезка $[a, b]$ оси Ox точками $x = x_k$ в пространстве соответствует разбиение тела на тонкие круглые блины плоскостями $x = x_k$. Объём блина малой толщины δ приближённо равен произведению площади основания на толщину:

рисунок

$$V(x+\delta) - V(x) = \pi f(x)^2\delta + o(\delta).$$

Поэтому объём всего тела вращения равен

а вокруг Oy?

$$V = \pi \int_a^b f(x)^2\,dx.$$

Чтобы вычислить площадь поверхности вращения, а при этом всегда подразумевается лишь боковая поверхность, составим её из боковых поверхностей тонких блинов. На участке постоянного радиуса $f(x)$ и толщины δ площадь равна $2\pi f(x)\delta$. При изменении радиуса боковая поверхность тела вращения образует некоторый угол $\alpha(x)$ с осью Ox. Тогда площадь находится делением предыдущего выражения на $\cos\alpha(x)$, что равносильно умножению на $\sqrt{1+\operatorname{tg}^2\alpha(x)}$. При этом $\operatorname{tg}\alpha(x) = f'(x)$. Итак, получаем формулу для вычисления площади поверхности тела вращения:

рисунок

$$S = 2\pi \int_a^b f(x)\sqrt{1+f'(x)^2}\,dx.$$

В этих задачах об объёме и площади поверхности, благодаря вращательной симметрии, скрыто предварительное интегрирование: площади и периметры круговых сечений уже известны. Для других тел такие задачи требуют кратного интегрирования; в нашем курсе это делается во втором семестре.

Масса и моменты. Возьмём стержень переменной удельной (линейной) плотности $\varrho(x)$. На его маленьком кусочке плотность приближённо постоянна, поэтому масса кусочка от x до $x+\delta$ равна $\varrho(x)\delta + o(\delta)$. Тогда масса отрезка $[a,b]$ стержня есть

$$m = \int_a^b \varrho(x)\,dx.$$

Помимо массы, в механике нужны и другие характеристики тела, чаще всего — моменты статические и моменты инерции, выражаемые интегралами

$$\int_a^b (x-c)\varrho(x)\,dx, \qquad \int_a^b (x-c)^2\varrho(x)\,dx.$$

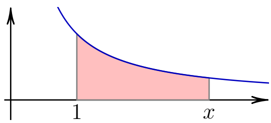
Из равенства статического момента нулю находят координату центра масс стержня:

$$c = \frac{1}{m}\int_a^b x\varrho(x)\,dx.$$

Для массы, распределённой по двумерным пластинам и трёхмерным телам, ситуация аналогична, но опять же требуется кратное интегрирование. Лишь для самых простых тел, благодаря их симметрии, удаётся обойтись более простыми средствами.

1.5. Новые функции

Трансцендентные элементарные функции. Интегрирование является, в частности, важным средством введения новых функций. Представим на время, что мы не знаем никаких основных функций, кроме степенных, x^n с целым n, но знаем об арифметических операциях, композиции и обратной функции. На этой основе с помощью интегрирования можно построить все основные трансцендентные элементарные функции: показательную и логарифмическую, тригонометрические и им обратные. Доказательства того, что все эти функции «хорошие», будут следовать автоматически, когда мы установим некоторые общие «улучшающие» свойства интеграла. лекция 5 17.09.15



Действительно, рассмотрим функцию Saint-Vincent

$$L(x) = \int_1^x \frac{dt}{t}.$$

По-старинному, это квадратура гиперболы. Найдём

$$L(ab)=\int_1^{ab}\frac{dt}{t}=\int_1^{a}\frac{dt}{t}+\int_a^{ab}\frac{dt}{t}=\int_1^{a}\frac{dt}{t}+\int_1^{b}\frac{du}{u}.$$

Второе слагаемое преобразовано заменой $t=au$. Получаем функциональное уравнение

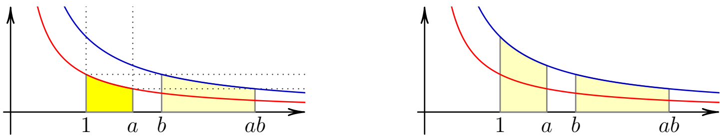
$$L(ab)=L(a)+L(b).$$

Именно это уравнение (стремление облегчить вычисление произведений) исторически явилось источником логарифмической функции. Вопрос об основании логарифмов решается позже, но он не столь важен, ибо смена основания сводится к умножению на константу.

Обнаруженному только что в результате замены равенству

$$\int_a^{ab}\frac{dt}{t}=\int_1^{b}\frac{du}{u}=\int_1^{b}\frac{dt}{t}$$

можно дать изящное геометрическое истолкование. Две криволинейные трапеции, квадратуры которых здесь написаны, связаны преобразованием плоскости, состоящим из двух шагов: горизонтальное сжатие и вертикальное растяжение с одинаковыми коэффициентами.



Упражнение. *Убедитесь в этом и проследите, как по ходу и в итоге преобразования изменяются площади фигур.*

Итак, логарифмы теперь есть. Обратная функция объявляется показательной. Вся тригонометрия при этом уже на самом деле появляется, но чтобы разглядеть её, требуются комплексные числа.

Есть и другие пути к (сперва обратным) тригонометрическим функциям через квадратуры. Обозначим через $A(x)$ дугу единичной окружности, отсчитываемую вправо от оси Oy до точки с абсциссой x. По привычной тригонометрии $x=\sin A(x)$, так что $A(x)=\arcsin x$.

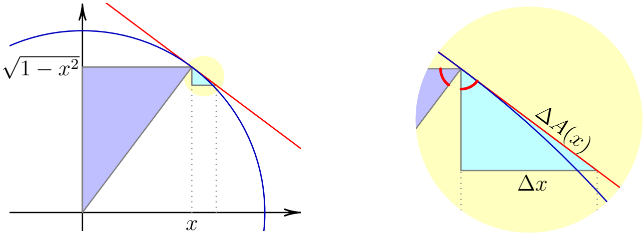
Newton 1669

Подобие прямоугольных треугольников даёт

$$\Delta A(x)\approx\frac{\Delta x}{\sqrt{1-x^2}}\qquad\Longrightarrow\qquad\arcsin x=\int_0^x\frac{dt}{\sqrt{1-t^2}}.$$

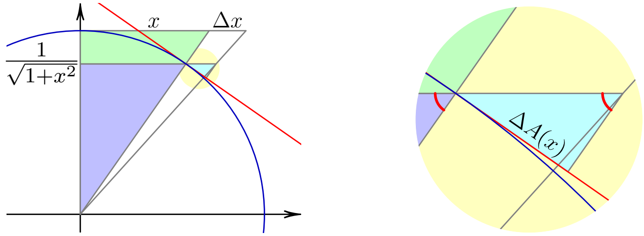
Упражнение. *Распознайте в этой же картинке равенство*

$$\int_0^x\sqrt{1-t^2}\,dt=\tfrac12\arcsin x+\tfrac12x\sqrt{1-x^2}.$$

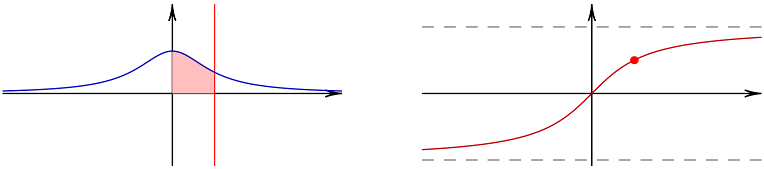


Немного модифицировав картинку, чтобы реализовать теперь соотношение $x = \operatorname{tg} A(x)$, аналогичным способом получим выражение арктангенса как квадратуры:

$$A(x) = \int_0^x \frac{dt}{1+t^2}.$$

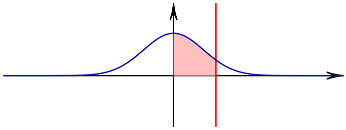
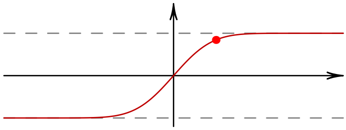


Замешанную здесь линию $y = \frac{1}{1+x^2}$ раньше называли **локоном Марии Аньези**, и с этим названием связана любопытная история.



Специальные функции. Разнообразные функции, на которые опираются приложения математики, в том числе в физике, называются **специальными**. Среди них много неэлементарных. Этот класс очень обширен и недостаточно хорошо систематизирован.

Специальные функции возникают различными путями. Находясь лишь в самом начале курса основ анализа, многие способы мы даже не можем сформулировать. Укажем лишь пару примеров, получаемых (в духе предыдущего подраздела) как неэлементарные первообразные элементарных функций.

Функция ошибок

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

широко применяется в статистике. С ней связаны интегралы Френеля

$$C(x) = \int_0^x \cos(t^2)\, dt, \qquad S(x) = \int_0^x \sin(t^2)\, dt,$$

возникающие в оптике и полезные при строительстве дорог.

$E(x)$?

Простейшие дифференциальные уравнения. Нахождение первообразных — это решение уравнения $y' = f(x)$ с известной функцией в правой части; целью при этом является найти неизвестную функцию $y = y(x)$. Более общее **дифференциальное уравнение** первого порядка выражает производную неизвестной функции через саму функцию и её аргумент: $y' = f(x, y)$. Ещё более сложные дифференциальные уравнения — богатейший источник новых функций.

Пример. Решение уравнения $y' = y$ можно угадать, поскольку мы знаем функцию, равную своей производной: $y = e^x$. Однако это не единственное решение: подходит также $y = Ce^x$ с любой постоянной C.

Аналогично угадываем решение уравнения $y' = ay$ с постоянным коэффициентом a в виде $y = Ce^{ax}$.

В общей постановке задача решения дифференциального уравнения $y' = f(x, y)$ слишком трудна. Сейчас мы изучим важный частный случай: уравнения вида $y' = f(x)\, g(y)$. При обычном интегрировании

$$\frac{dy}{dx} = f(x) \quad \Rightarrow \quad dy = f(x)\, dx \quad \Rightarrow \quad y = \int dy = \int f(x)\, dx.$$

Рассматриваемые уравнения интегрируются по аналогии:

$$\frac{dy}{dx} = f(x)\, g(y) \quad \Rightarrow \quad \frac{dy}{g(y)} = f(x)\, dx \quad \Rightarrow \quad \int \frac{dy}{g(y)} = \int f(x)\, dx.$$

Решение нередко получается лишь в виде неявной функции. Поскольку метод основан на разделении переменных x и y по двум частям равенства, такие уравнения называют **уравнениями с разделяющимися переменными**.

Пример. Решая таким путём уравнение $y' = ay$, получаем

$$\frac{dy}{y} = a\,dx \quad \Rightarrow \quad \ln|y| = ax + c \quad \Rightarrow \quad y = Ce^{ax}.$$

Решение $y = 0$ потеряно на первом же шаге ввиду деления на y, но войдёт в итоговый ответ, если допустить значение $C = 0$. Кроме того, допуская $C < 0$, мы избавляемся от модуля на y.

Пример. Решим уравнение $y' = 2xy$:

$$\frac{dy}{y} = 2x\,dx \quad \Rightarrow \quad \ln|y| = x^2 + c \quad \Rightarrow \quad y = Ce^{x^2}.$$

В первых примерах мы выписывали **общее решение** — выражение с произвольной константой, при каждом значении константы дающее конкретное решение уравнения. В следующем примере найдём единственное решение дифференциального уравнения, удовлетворяющее также **начальному условию**.

Пример (Радиоактивный распад). Доля атомных ядер радиоактивного вещества, распадающихся в единицу времени, постоянна для каждого изотопа. Математически этот закон природы выражает дифференциальное уравнение $\dot{y} = -ay$ на количество изотопа $y(t)$. Производные по времени принято обозначать точкой вместо штриха.

Общим решением является $y(t) = Ce^{-at}$. Начальным условием считаем начальное количество изотопа $y(0)$. Поэтому подстановкой $t = 0$ определяем значение $C = y(0)$.

Правильное применение определённых интегралов напрямую даёт ответ с учётом начального условия.

Пример (Закон теплообмена Ньютона). Скорость изменения температуры тела пропорциональна разности между его температурой $y(t)$ и температурой окружающей среды T, считающейся постоянной. Это даёт дифференциальное уравнение $\dot{y} = k(T - y)$. Разделяя переменные и интегрируя по промежутку времени от t_0 до t_1, находим

$$\begin{aligned}
\frac{dy}{y-T} = -k\,dt \quad &\Rightarrow \quad \int_{y(t_0)}^{y(t_1)} \frac{dy}{y-T} = -k\int_{t_0}^{t_1} dt \\
&\Rightarrow \quad \ln\frac{y(t_1) - T}{y(t_0) - T} = -k(t_1 - t_0) \\
&\Rightarrow \quad y(t) = T + (y(t_0) - T)e^{-k(t-t_0)}.
\end{aligned}$$

Обратите внимание на пределы интегрирования по y.

Упражнение. *Почему под логарифмом нет модуля?*

Пример (Движение в потенциальном поле). Подставим в закон сохранения энергии $T(v)+U(x) = E$ кинетическую энергию $\frac{1}{2}mv^2$ и выразим скорость $v = \dot{x}$. Получим дифференциальное уравнение

$$\dot{x} = \sqrt{\frac{2}{m}(E - U(x))}.$$

Переменные разделяются; решение получается как обратная функция $t(x)$, увы, не обязательно элементарная.

Глава 2. КОМПЛЕКСНЫЕ ЧИСЛА И ПРИЛОЖЕНИЯ

2.1. Комплексные числа

лекция 5+ 17.09.15

Комплексные числа возникают уже при решении квадратного уравнения $x^2+1=0$. Однако веками это оставалось невидимым: подводило простое заключение, что корней нет. С неизбежностью комплексных чисел столкнулся Кардано (1539), изучая кубические уравнения: уравнение $x^3=15x+4$ имеет корень $x=4$, но его вычисление по тайному тогда методу включает арифметические операции с числом $\sqrt{-121}$. Затем такие операции ещё долго считались мистикой и вызывали много путаницы, хотя Бомбелли (1572) подробно объяснил их в том же трактате, где он впервые в Европе оперировал с отрицательными числами.

Алгебраическая форма и арифметические операции. Каждое комплексное число представляют выражением $a+ib$, где специальный символ «мнимой единицы» i удовлетворяет основному равенству $i^2=-1$. При этом вещественные числа a и b называют соответственно **вещественной** и **мнимой частями** числа $a+ib$. Когда $b=0$, получается вещественное число a.

Сложение и вычитание комплексных чисел выполняются отдельно над вещественной и мнимой частями:

$$(a_1+ib_1)\pm(a_2+ib_2)=(a_1\pm a_2)+i(b_1\pm b_2).$$

Умножение также выполняется обычным раскрытием скобок с последующей заменой i^2 на -1:

$$(a_1+ib_1)(a_2+ib_2)=(a_1a_2-b_1b_2)+i(a_2b_1+a_1b_2).$$

Формулу для деления не назовёшь приятной. Для пояснения идеи, привлечём аналогию со школьным приёмом избавления от иррациональности в знаменателе путём домножения на сопряжённое число:

$$\frac{1}{5-\sqrt{3}}=\frac{5+\sqrt{3}}{(5-\sqrt{3})(5+\sqrt{3})}=\frac{5+\sqrt{3}}{22}.$$

В том же духе,

$$\frac{1}{a+ib}=\frac{a-ib}{(a+ib)(a-ib)}=\frac{a}{a^2+b^2}+i\frac{-b}{a^2+b^2}.$$

Поэтому деление возможно на любое комплексное число $a+ib\neq 0$ и

$$\frac{a_1+ib_1}{a_2+ib_2}=\frac{a_1a_2+b_1b_2}{a_2^2+b_2^2}+i\frac{a_2b_1-a_1b_2}{a_2^2+b_2^2},$$

но лучше помнить не формулу, а способ её получения.

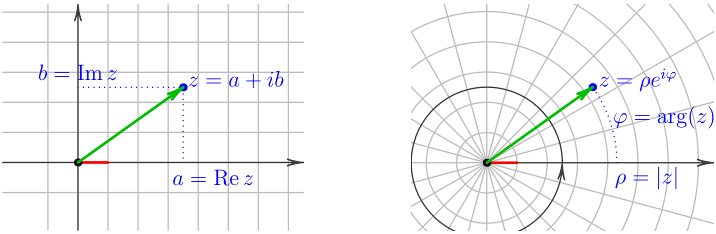
Комплексное число $a - ib$ называют **комплексно сопряжённым** к $z = a+ib$ и обозначают обычно через $\bar{z}$, но иногда (чаще в физике) через z^*.

Геометрическое представление. К концу XVIII века немало математиков осознали полезность комплексных чисел. Вессель и Арган независимо предложили изображать комплексные числа на плоскости: число $x + iy$ соответствует точке с декартовыми координатами (x, y) или её радиус-вектору.

рисунок

Сложение комплексных чисел, как легко видеть из формулы, согласуется с правилом параллелограмма для сложения векторов. Собственно, в этой связи норвежский геодезист Вессель первым и ввёл понятие вектора и правило параллелограмма. Арган интерпретировал умножение на i как поворот плоскости на угол $\pi/2$. Связь комплексных чисел и движений плоскости освещена чуть подробнее в конспекте моих лекций по алгебре и геометрии 2009 г.

гиперссылка



Тригонометрическая форма. Переходя от декартовых координат к полярным,

$$x = \rho \cos\varphi, \quad y = \rho \sin\varphi,$$

получаем тригонометрическую форму записи комплексных чисел:

$$z = x + iy = \rho(\cos\varphi + i\sin\varphi),$$

где $\rho = |z|$ называют **модулем** или абсолютной величиной z, а φ — **аргументом** z. При $z = 0$ неопределённость аргумента не мешает работе. В физике модуль называют амплитудой, а аргумент — фазой.

В этой форме очень удобно выполнять умножение и деление:

$$\begin{aligned}
&(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) \\
&\qquad = (\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2) \\
&\qquad = \cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2).
\end{aligned}$$

При умножении модули перемножаются, а аргументы складываются. Тогда при делении модули делятся, а аргументы вычитаются.

Однако тригонометрическую форму записи комплексных чисел следует считать лишь промежуточным шагом, поскольку мы тотчас заменим её на более компактную и эффективную.

Комплексная экспонента. Рассмотрим повнимательнее функцию

$$E(t) = \cos t + i \sin t.$$

Она имеет весьма интересную производную:

$$\begin{aligned}\frac{d}{dt}E(t) &= \frac{d}{dt}\cos t + i\frac{d}{dt}\sin t\\ &= -\sin t + i\cos t = i(\cos t + i\sin t) = iE(t).\end{aligned}$$

Единственным решением дифференциального уравнения $\dot{E}(t) = iE(t)$ с начальным условием $E(0) = 1$ должна быть функция $E(t) = e^{it}$. Свойство $E(s+t) = E(s)E(t)$, кодирующее формулы косинуса и синуса суммы, также говорит об экспоненциальном характере решения. Временное обозначение $E(t)$ теперь можно забыть.

Получилась знаменитая **формула Эйлера**

$$e^{it} = \cos t + i \sin t.$$

Подставляя $t = \pi$, видим не менее знаменитый вывод $e^{i\pi} = -1$, или

$$e^{i\pi} + 1 = 0.$$

Вместо тригонометрической формы записи комплексного числа гораздо удобнее экспоненциальная форма $z = \rho e^{i\varphi}$. Выгоды станут особенно очевидны при дальнейшем употреблении в физике, где выражение для фазы φ бывает достаточно громоздким, и выписывать его дважды, раз под косинусом и другой под синусом, расточительно.

Повторим формулы для умножения и деления:

$$\begin{gathered}(\rho_1 e^{i\varphi_1})(\rho_2 e^{i\varphi_2}) = (\rho_1\rho_2)e^{i(\varphi_1+\varphi_2)},\\ (\rho_1 e^{i\varphi_1})/(\rho_2 e^{i\varphi_2}) = (\rho_1/\rho_2)e^{i(\varphi_1-\varphi_2)}.\end{gathered}$$

Выразив косинус и синус через экспоненту, где окружно

$$\cos t = \frac{e^{it}+e^{-it}}{2}, \quad \sin t = \frac{e^{it}-e^{-it}}{2i},$$

отметим связь с одноимёнными гиперболическими функциями

$$\operatorname{ch} t = \frac{e^{t}+e^{-t}}{2}, \quad \operatorname{sh} t = \frac{e^{t}-e^{-t}}{2}.$$

Итак, косинусы и синусы суть лишь тени комплексной экспоненты — функции $e^{\lambda t}$ с произвольным комплексным $\lambda = \alpha + i\beta$:

$$e^{(\alpha+i\beta)t} = e^{\alpha t}e^{i\beta t} = e^{\alpha t}(\cos\beta t + i\sin\beta t).$$

Она очень пригодится нам в следующем разделе.

2.2. Уравнение свободных колебаний

Линейные уравнения второго порядка. Рассмотрим простую механическую систему: тело на пружинке. Закон Ньютона $F = m\ddot{x}$ и закон Гука для силы растянутой пружины дают уравнение $m\ddot{x} + kx = 0$. Ускорение есть производная скорости $\dot{x}$, и здесь мы впервые в нашем курсе видим вторую производную. Пропорциональное скорости сопротивление вносит ещё одно слагаемое.

Для математики характерно стремление к общности. Уступая ему, перейдём к абстрактным обозначениям для постоянных коэффициентов и будем изучать дифференциальное уравнение

$$(*) \qquad a\ddot{x}(t) + b\dot{x}(t) + cx(t) = 0,$$

где $a \neq 0$.

Прежде всего обратим внимание на структуру нашего уравнения: это линейная комбинация неизвестной функции и её производных; такие дифференциальные уравнения называют **линейными**. Более того, эта комбинация тождественно равна нулю; такие уравнения называют **однородными**. Теория линейных дифференциальных уравнений несравнимо проще теории нелинейных благодаря замечательному свойству всей совокупности решений.

Лемма (Принцип суперпозиции)**.** *Всякая линейная комбинация решений однородного линейного дифференциального уравнения также является его решением.*

Доказательство. Подставим комбинацию $x(t) = \alpha_1 x_1(t) + \alpha_2 x_2(t)$ в уравнение и перегруппируем:

$$a\ddot{x} + b\dot{x} + cx = \alpha_1(a\ddot{x}_1 + b\dot{x}_1 + cx_1) + \alpha_2(a\ddot{x}_2 + b\dot{x}_2 + cx_2).$$

Если $x_1(t)$ и $x_2(t)$ являются решениями, то обе скобки равны нулю. Отметим, что утверждение сохранит силу, даже если коэффициенты в $(*)$ зависят от t, причём с тем же доказательством. □

Следовательно, общее решение однородного линейного уравнения можно записать как линейную комбинацию какого-то набора решений с произвольными константами. Количество таких независимых **базисных** решений оказывается равным наивысшему порядку производной, входящей в дифференциальное уравнение, который называют **порядком** уравнения. Эти утверждения доказываются в курсах линейной алгебры и дифференциальных уравнений.

Опыт угадывания решений. В нашем случае независимых решений должно быть два. Научимся угадывать их.

Пример. Возьмём уравнение $\ddot{x} - x = 0$. Ему удовлетворяет любая функция, для которой $\dot{x} = x$, то есть $x_1(t) = e^t$. В то же время, годится и функция, для которой $\dot{x} = -x$, то есть $x_2(t) = e^{-t}$. Общее решение есть $x(t) = c_1 e^t + c_2 e^{-t}$.

Пример. Возьмём уравнение $\ddot{x} + x = 0$. Нам известны две функции с таким свойством: это $x_1(t) = \cos t$ и $x_2(t) = \sin t$. Общее решение есть $x(t) = c_1 \cos t + c_2 \sin t$.

Пример. Возьмём уравнение $\ddot{x} + \omega^2 x = 0$. Сюда подходят $x_1(t) = \cos \omega t$ и $x_2(t) = \sin \omega t$. Общее решение есть $x(t) = c_1 \cos \omega t + c_2 \sin \omega t$. Это уравнение колебаний без сопротивления, получаемое из закона Гука; константа $\omega = \sqrt{k/m}$ является «угловой скоростью».

Однако мы ведь уже знаем, что косинус и синус представляются через экспоненту! Подставляя в уравнение комплексные функции $z_+(t) = e^{i\omega t}$ и $z_-(t) = e^{-i\omega t}$, видим, что обе являются решениями. Тогда общее вещественное решение можно записать в виде

зачем?

$$x(t) = \mathfrak{Re}(z(t)) = \mathfrak{Re}(c_+ e^{i\omega t} + c_- e^{-i\omega t})$$

с комплексными константами. Чтобы понять переход отсюда к выражению общего вещественного решения через косинус и синус, сделайте следующее упражнение.

Упражнение. *Докажите, что комплексное решение*

$$z(t) = c_+ e^{i\omega t} + c_- e^{-i\omega t}$$

вещественно тогда и только тогда, когда c_+ и c_- сопряжены.

Характеристическое уравнение и решения. Теперь подставим экспоненту $x(t) = e^{\lambda t}$ в уравнение $(*)$, пока не зная коэффициент λ в показателе:

лекция 6
21.09.15

$$a\ddot{x} + b\dot{x} + cx = (a\lambda^2 + b\lambda + c)e^{\lambda t} = 0.$$

Поскольку $e^{\lambda t} \neq 0$, получим обычное квадратное уравнение на λ:

$$a\lambda^2 + b\lambda + c = 0.$$

Его называют **характеристическим уравнением** для исходного дифференциального уравнения, а его корни — **характеристическими корнями** или **числами**. Характеристическое уравнение выписывается непосредственно по исходному дифференциальному. Самой обычной ошибкой при этом, вероятно, является излишнее умножение c на λ.

Согласно типам корней, квадратные уравнения с вещественными
стиль коэффициентами делятся на три группы:

(1) два различных вещественных корня λ_1, λ_2;
(2) двойной вещественный корень λ;
(3) пара различных комплексно-сопряжённых корней $\alpha \pm i\beta$.

В первом случае сразу пишем общее решение

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

Во втором случае мы угадали только одно решение $x_1(t) = e^{\lambda t}$. Оказывается, что вторым решением является $x_2(t) = te^{\lambda t}$. Легко это проверить подстановкой в уравнение $(*)$ и не требовать больше. Итак, здесь общее решение

$$x(t) = e^{\lambda t}(c_1 + c_2 t).$$

Если желание увидеть причину появления $te^{\lambda t}$ всё-таки остаётся… Она достаточно красива. Отступим к случаю, когда характеристические корни различны, но близки: $\lambda_1 = \lambda$, а $\lambda_2 = \lambda + \delta$. Тогда

$$x(t) = \frac{e^{(\lambda+\delta)t} - e^{\lambda t}}{\delta}$$

является решением. Предел этого выражения при $\delta \to 0$ равен производной $e^{\lambda t}$ как функции от λ, а именно, $te^{\lambda t}$. Так что опять исключи-
гиперссылка тельный случай интерпретирован как предельный.

В третьем случае переходим от линейной комбинации комплексно-сопряжённых экспонент к косинусам и синусам и записываем общее вещественное решение

$$x(t) = e^{\alpha t}(c_1 \cos \beta t + c_2 \sin \beta t)$$

с вещественными константами.

Упражнение. *При каком условии все решения уравнения* $(*)$ *стремятся к нулю при* $t \to +\infty$*?*

Поскольку общее решение уравнения второго порядка включает два произвольных постоянных коэффициента, чтобы выделить единственное решение, необходимо наложить два условия. Чаще всего это

- **начальные условия**: значения $x(t)$ и $\dot{x}(t)$ в начальный момент,
- **граничные условия**: значения $x(t)$ в двух точках.

Пример. Решим уравнение с начальными условиями:

$$\ddot{x} + 5\dot{x} + 6x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1.$$

Характеристическое уравнение легко раскладывается на множители:

$$\lambda^2 + 5\lambda + 6 = (\lambda + 2)(\lambda + 3) = 0 \quad \Longrightarrow \quad \lambda_1 = -2,\ \lambda_2 = -3.$$

Общее решение данного дифференциального уравнения есть

$$x(t) = c_1 e^{-2t} + c_2 e^{-3t},$$

а его производная —

$$\dot{x}(t) = -2c_1 e^{-2t} - 3c_2 e^{-3t}.$$

Подставляя $t = 0$, видим систему из двух линейных уравнений

$$c_1 + c_2 = 0, \quad -2c_1 - 3c_2 = 1.$$

Решая систему, находим $c_1 = 1$ и $c_2 = -1$. Единственным решением задачи является

$$x(t) = e^{-2t} - e^{-3t}.$$

Пример. Решим уравнение с начальными условиями:

$$\ddot{x} + 6\dot{x} + 10x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1.$$

Характеристическое уравнение имеет комплексные корни:

$$\lambda^2 + 6\lambda + 10 = (\lambda + 3)^2 + 1 = 0 \quad \Longrightarrow \quad \lambda = -3 \pm i.$$

Общее решение данного дифференциального уравнения есть

$$x(t) = e^{-3t}(c_1 \cos t + c_2 \sin t),$$

а его производная —

$$\dot{x}(t) = e^{-3t}\big(c_1(-3\cos t - \sin t) + c_2(-3\sin t + \cos t)\big).$$

Подставляя $t = 0$, видим систему из двух линейных уравнений

$$c_1 = 0, \quad -3c_1 + c_2 = 1,$$

откуда $c_1 = 0$ и $c_2 = 1$. Единственным решением задачи является

$$x(t) = e^{-3t}\sin t.$$

2.3. Интегрирование рациональных функций

Постановка задачи и стратегия. Частное двух полиномов называют **рациональной функцией**. Мы уже видели первообразные от таких функций, например:

$$\int \frac{dx}{x^2} = -\frac{1}{x} + C, \quad \int \frac{dx}{x} = \ln|x| + C, \quad \int \frac{dx}{1+x^2} = \operatorname{arctg} x + C.$$

Более того, функции $\ln x$ и $\operatorname{arctg} x$ введены именно таким путём. Естественно поинтересоваться, какие ещё функции могут появиться при интегрировании рациональных.

Euler 1730

Теорема (Эйлер). *Первообразная всякой рациональной функции с известными корнями знаменателя выражается через рациональные функции, логарифм и арктангенс.*

Лейбниц и Бернулли считали, что существенно новых функций при интегрировании рациональных не возникает. Поскольку корни не любого полинома можно выразить через элементарные функции, Эйлер закрыл вопрос лишь условно; так он теперь и понимается. С другой стороны, в приложениях редко возникают знаменатели, корни которых выразить нельзя.

Остаток раздела занят доказательством этой теоремы. Оно состоит из двух совершенно разных частей: в алгебраической части мы решим задачу представления произвольной дроби суммой простейших слагаемых; в аналитической части нам останется найти первообразные для этих простейших.

Разложение дроби на простейшие в теории. Решаемая здесь задача имеет самостоятельную ценность, поскольку её применение к интегрированию — далеко не единственное.

Далее рациональные функции будем, для краткости, называть просто дробями. Дробь называют **правильной**, если степень её числителя ниже (меньше) степени знаменателя.

Лемма. *Всякая дробь однозначно представима как сумма полинома и правильной дроби.*

Доказательство. Поделим с остатком числитель на знаменатель. □

Дальше рассуждения опираются на основную теорему алгебры полиномов, а точнее, на её следствия: каждый полином с комплексными коэффициентами представи́м как произведение линейных; каждый полином с вещественными коэффициентами представи́м как произведение линейных и квадратичных. С теоретической точки зрения намного выгоднее работать с комплексными коэффициентами.

Рассмотрим на самом простом примере основную идею метода, позволяющего разобраться с любыми правильными дробями.

Пример. Попробуем представить дробь с квадратичным знаменателем, имеющим различные корни, как сумму двух более простых дробей с неизвестными числовыми коэффициентами:

$$\frac{1}{(x-a)(x-b)} = \frac{A}{x-a} + \frac{B}{x-b}.$$

Коэффициенты отыскивают многими способами. Самый обычный, он же топорный: привести правую часть к общему знаменателю и приравнять числители обеих частей как полиномы. Более быстрый, но не универсальный способ: умножить равенство на $x-a$, затем подставить $x = a$, и от правой части останется A в чистом виде; коэффициент B находим аналогично. Получим

$$\frac{1}{(x-a)(x-b)} = \frac{1}{a-b} \cdot \frac{1}{x-a} + \frac{1}{b-a} \cdot \frac{1}{x-b}.$$

Каков бы ни был знаменатель дроби, если среди его неприводимых делителей есть различные, то таким путём задача сводится к дробям, степени знаменателей которых ниже. После конечного числа шагов получится сумма конечного числа дробей, у каждой из которых знаменатель является степенью неприводимого полинома. Правильную дробь такого вида называют **примарной**.

Лемма. *При $x \to +\infty$:*

(1) *предел правильной дроби равен нулю;*
(2) *предел ненулевого полинома не равен нулю.*

Доказательство. Упражнение. □

Лемма. *Всякая правильная дробь представима как сумма примарных дробей.*

Доказательство. Дадим короткое и простое рассуждение, хотя оно сильно отличается от способа, применяемого на практике. Дробь $1/q(x)$ предста́вим как сумму примарных последовательностью шагов из примера выше. Умножив эту сумму на $p(x)$, получим разложение дроби $p(x)/q(x)$, слагаемые которого не обязаны быть правильными дробями, но мы выделим правильную часть из каждого слагаемого. Тогда полиномиальные части без знаменателей чудом сократятся, поскольку иначе предел всей суммы при $x \to +\infty$ не сможет равняться нулю. Получим разложение $p(x)/q(x)$ на примарные дроби. $\square$

Упражнение. *Разложите дроби*

$$\frac{1}{(x-a)(x-b)(x-c)}, \qquad \frac{p(x)}{(x-a)(x-b)(x-c)},$$

где корни различны.

Упражнение. *Разложите почти произвольную правильную дробь $p(x)/q(x)$, где все корни знаменателя различны (нет кратных).*

Если же в примарной дроби степень числителя не только ниже степени знаменателя, но даже ниже степени неприводимого делителя знаменателя, то дробь называют **простейшей**.

Пример. Дробь $x/(x+1)^2$ примарная, но не простейшая.

Лемма. *Всякая примарная дробь однозначно представима как сумма простейших дробей.*

Доказательство. Замаскированная схема Горнера. $\square$

Наконец, комбинируя леммы, получаем алгебраический результат.

Теорема. *Всякая рациональная функция (однозначно) представима как сумма полинома и простейших дробей.*

Упражнение. *Докажите однозначность.*

Вещественный случай уже не красив, но сейчас именно он требуется на практике. Доказывать заново тут почти ничего не нужно; достаточно лишь обеспечить возможность отделения квадратичных сомножителей знаменателя, имеющих комплексно-сопряжённые корни. Однако их можно отделить, пользуясь комплексными коэффициентами, а потом собрать попарно в вещественные дроби. Либо можно исключить комплексные числа из всех промежуточных шагов, найдя сразу в общем виде неизвестные коэффициенты разложения

$$\frac{1}{(x-a)(x^2+bx+c)} = \frac{A}{x-a} + \frac{Bx+C}{x^2+bx+c},$$

а также ещё одного варианта — с двумя различными квадратичными сомножителями. Однако это не практично.

Разложение дроби на простейшие на практике. Разложение данной рациональной функции на простейшие дроби выполняют, исходя из знания того, какие возможны знаменатели и какие им соответствуют числители. Конечно, предварительно требуется разложить знаменатель на неприводимые множители. Затем, раскладывая дробь, важно не забыть ни одного слагаемого. Когда всё верно, количество неизвестных коэффициентов должно равняться степени общего знаменателя. В случае неправильной «догадки» метод неопределённых коэффициентов укажет на ошибку: попытки найти коэффициенты неизменно будут безуспешны.

Пример. Раскладываем данную дробь, исходя из состава знаменателя:
$$\frac{x^2+2x-3}{(x+1)^2(x^2-x+1)} = \frac{A}{(x+1)^2} + \frac{B}{x+1} + \frac{Cx+D}{x^2-x+1}.$$
Приведя к общему знаменателю, приравняем коэффициенты в числителях при одинаковых степенях x. Получим систему четырёх линейных уравнений, из которой и найдём коэффициенты. Некоторые коэффициенты существенно быстрее находить, как в первом примере; этот способ можно хитро применить и здесь.

Интегрирование простейших дробей. На этом этапе комплексный подход гораздо проще вещественного, ибо в комплексном случае всякая простейшая дробь имеет вид $1/(x-a)^k$ с точностью до числового коэффициента. Замена переменной сводит поиск первообразной к дроби $1/x^k$, ответ для которой известен. И сказке конец.

В этой аккуратной схеме самой существенной проблемой для нас сейчас является тот факт, что мы не рассматривали комплексную функцию $\ln z$. Поэтому сложно понять, например, что такое $\ln(x-i)$. Заполнить этот пробел должным образом не так просто; обычно это делается только в комплексном анализе. Видимо, лучшее, что можно сказать сейчас: мы знаем о комплексной экспоненте, а логарифм есть обратная функция. Отсюда, $\ln z = \ln|z| + i\arg z$.

Интересно также спросить: а где же арктангенс? Он ведь не возник. Ответ заключается в замечательном равенстве
$$\operatorname{arctg} x = \frac{i}{2}\ln\left(\frac{i+x}{i-x}\right).$$

Упражнение. *«Получите» это равенство, разложив $1/(x^2+1)$ на простейшие дроби с комплексными коэффициентами, затем интегрируя (модуль под логарифмом не нужен, но постоянная интегрирования потребуется).*

Упражнение. *Зная теперь секрет комплексного логарифма, объясните модуль под логарифмом в привычной первообразной для* $1/x$.

Формула понижения. Для тех, кто чувствует слишком сильный дискомфорт при чтении нескольких предыдущих абзацев, рассмотрим чисто вещественный поиск первообразных для оставшихся в вещественном случае простейших дробей.

Итак, требуется взять интегралы

$$I_k = \int \frac{x\,dx}{(1+x^2)^k}, \qquad J_k = \int \frac{dx}{(1+x^2)^k}.$$

Первый из них легко берётся заменой $u = 1 + x^2$. Ответ рационален при $k > 1$ и содержит логарифм при $k = 1$:

$$I_k = -\frac{1}{k} \cdot \frac{1}{(1+x^2)^{k-1}} + C, \qquad I_1 = \frac{1}{2}\ln(1+x^2) + C.$$

Также мы знаем уже, что $J_1 = \operatorname{arctg} x + C$. Вычисление J_k при $k > 1$
23.01.16 посложнее. По формуле интегрирования по частям,

$$\begin{aligned} J_k = \int \frac{dx}{(1+x^2)^k} &= \frac{x}{(1+x^2)^k} - \int x\, d\Big(\frac{1}{(1+x^2)^k}\Big) \\ &= \frac{x}{(1+x^2)^k} + 2k \int \frac{x^2\,dx}{(1+x^2)^{k+1}} \\ &= \frac{x}{(1+x^2)^k} + 2k(J_k - J_{k+1}). \end{aligned}$$

Найдя отсюда J_{k+1}, увидим **формулу понижения**, сводящую J_{k+1} к J_k. Применяя её $k-1$ раз, теоретически можно выразить J_k через рациональные функции и арктангенс.

Любопытные читатели могут найти в больших таблицах интегралов много формул понижения для разнообразных функций, а некоторые даже будут получены на семинарах.

$x^2 + bx + c$?

2.4. Рационализация некоторых интегралов

лекция 7
24.09.15

Интегралы от некоторых классов функций, построенных из рациональных функций $R(u, v)$ двух переменных подстановками нерациональных, могут быть сведены к интегрированию рациональных. Мы рассмотрим их весьма кратко. Знание деталей многочисленных подстановок не заслуживает причисления к базовым навыкам. Подробное их рассмотрение в нашем курсе излишне ещё и потому, что в тех случаях, когда нужны не самые простые подстановки, в реальности прибегают к компьютерным системам символьных вычислений.

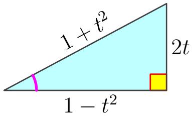
Тригонометрия. Интеграл от всякой функции вида $R(\cos x, \sin x)$ сводится к интегралу от рациональной функции с помощью универсальной подстановки $t = \operatorname{tg} \frac{x}{2}$. Тогда из различных тригонометрических формул получим

$$\cos x = \frac{1-t^2}{1+t^2}, \quad \sin x = \frac{2t}{1+t^2}, \quad dx = \frac{2dt}{1+t^2} = 2d(\operatorname{arctg} t).$$

Чтобы облегчить запоминание этих выражений, отметим прямоугольный треугольник с гипотенузой $1+t^2$ и катетами $1-t^2$ и $2t$. Формула

$$\operatorname{tg} 2\alpha = \frac{2 \operatorname{tg} \alpha}{1 - \operatorname{tg}^2 \alpha}$$

с заменой $x = 2\alpha$ даёт именно $t = \operatorname{tg} \frac{x}{2}$.



Как всегда с универсальными методами, во многих наиболее распространённых случаях этот подход оказывается не самым простым из возможных. Быстрее к цели ведут подстановки, пригодные в отдельных случаях (не)чётности:

- подстановка $t = \cos x$, когда $R(\cos x, -\sin x) = -R(\cos x, \sin x)$;
- подстановка $t = \sin x$, когда $R(-\cos x, \sin x) = -R(\cos x, \sin x)$;
- подстановка $t = \operatorname{tg} x$, когда $R(-\cos x, -\sin x) = R(\cos x, \sin x)$.

Пример. Чтобы проинтегрировать функцию $\sin^3 x$, преобразуем

$$\sin^3 x \, dx = \sin^2 x \cdot \sin x \, dx = (\cos^2 x - 1) \cdot d(\cos x).$$

Подстановка $t = \cos x$ сводит интеграл к рациональному.

По этой же причине подстановка $t = \cos x$ работает для всех интегралов первого из отмеченных типов: вынося один синус, получаем функцию, рационально выражающуюся через косинус. Второй тип аналогичен, только меняются роли синуса и косинуса.

Пример. Чтобы проинтегрировать функцию $1/\cos x \sin x$, преобразуем

$$\frac{dx}{\cos x \sin x} = \frac{\cos x}{\sin x} \cdot \frac{dx}{\cos^2 x} = \frac{d(\operatorname{tg} x)}{\operatorname{tg} x}.$$

По этой же причине подстановка $t = \operatorname{tg} x$ работает для всех интегралов третьего из отмеченных типов: подынтегральная функция преобразуется к рациональной функции от

$$\frac{\sin x}{\cos x} = \operatorname{tg} x \quad \text{и} \quad \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x.$$

Любопытно отметить, что любую рациональную функцию можно представить как сумму функций трёх указанных типов симметрии — косинусы и синусы в качестве аргументов тут ни при чём — благодаря тождеству

$$\begin{aligned} 2R(u,v) &= \big(R(u,v) - R(u,-v)\big) \\ &\quad + \big(R(u,-v) - R(-u,-v)\big) \\ &\quad + \big(R(-u,-v) + R(u,v)\big). \end{aligned}$$

Тогда каждую скобку можно проинтегрировать, применяя соответствующую подстановку.

Аналогичный подход применим для функций вида $R(\operatorname{ch} t, \operatorname{sh} t)$, хотя в этом случае проще положить $t = e^x$.

Квадратичные иррациональности. Выразим $\sin x = \sqrt{1 - \cos^2 x}$, игнорируя знаки. Ввиду замены $t = \cos x$, рассмотренный выше класс функций вида $R(\cos x, \sin x)$ с точки зрения задачи отыскания первообразных равносилен классу функций вида $R(x, \sqrt{1 - x^2})$. При наличии иррациональности $\sqrt{x^2 - 1}$ или $\sqrt{x^2 + 1}$ аналогичным образом помогают гиперболические подстановки $x = \operatorname{ch} t$ и $x = \operatorname{sh} t$.

Функции более общего класса $R(x, \sqrt{ax^2 + bx + c})$ сводятся к уже перечисленным с помощью несложных преобразований. Кроме того, для них существуют специальные подстановки Эйлера, на которые мы не будем тратить время, как и на другие (немногие) общие классы иррациональностей, интегрируемых в элементарных функциях.

Сводимость таких интегралов к интегралам от рациональных функций и несводимость интегралов от более сложных функций $R(x, y(x))$, где переменные связаны алгебраическим уравнением $y^2 = P(x)$, имеет геометрическую природу. В случае $P(x) = 1 - x^2$ задача связана с окружностью и тригонометрией. Универсальная подстановка пришла из рациональной параметризации окружности:

$$(x, y) = \left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right).$$

При повышении степени полинома $P(x)$ линия, задаваемая уравнением $y^2 = P(x)$, усложняется и не обязана допускать рациональной параметризации; таковая редко возможна и слабо предсказуема.

Богатая и красивая математика связана со случаем полиномов $P(x)$ третьей и четвёртой степени. Такие интегралы впервые появились в геометрической задаче о длине дуги эллипса и потому называются эллиптическими. Эллиптическая функция также управляет истинным движением маятника.

Глава 3. РАЗЛОЖЕНИЯ ПО СТЕПЕНЯМ

3.1. Бином Ньютона

лекция 7+
24.09.15

Бином до Ньютона. Сумму из двух слагаемых называют **биномом**, особенно в тех случаях, когда рассматривается её степень. Достаточно изучить бином $1 + x$, ибо остальные сводятся к нему подстановками.

Правила раскрытия натуральных степеней бинома были известны ещё в X веке арабским и персидским математикам:

$$\begin{aligned}
(1+x)^2 &= 1 + 2x + x^2,\\
(1+x)^3 &= 1 + 3x + 3x^2 + x^3,\\
(1+x)^4 &= 1 + 4x + 6x^2 + 4x^3 + x^4,\\
(1+x)^5 &= 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5.
\end{aligned}$$

Коэффициенты в разложении бинома называют **биномиальными коэффициентами**. Эти числа очень часто встречаются в математике и имеют специальное обозначение. Коэффициент при x^k в раскрытии $(1 + x)^n$ обозначается через $\binom{n}{k}$, хотя старые русскоязычные учебники используют C_n^k. Вот стандартная, самая компактная формула:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad \text{где } k! = 1 \cdot 2 \cdot 3 \times \dots \times (k-1) \cdot k,$$

то есть произведение всех целых чисел от 1 до k включительно, называется **факториалом** натурального числа k и встречается само по себе ещё чаще, чем $\binom{n}{k}$. Составим небольшую таблицу для запоминания:

k	0	1	2	3	4	5	6	7	8
$k!$	1	1	2	6	24	120	720	5040	40320

Факториал нуля считается равным 1, как и любое пустое произведение. Тогда $\binom{n}{0} = 1$. Отметим также простой случай $\binom{n}{1} = n$.

Упражнение. *Раскрывая тождество $(1 + x)^n = (1 + x)(1 + x)^{n-1}$, установите рекуррентную формулу*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Если слишком трудно, начните с небольшого значения n.

Упражнение. *Теперь установите эту же формулу для выражения $\binom{n}{k}$ через факториалы. Это и докажет, что оно действительно даёт биномиальные коэффициенты.*

Для вычисления $\binom{n}{k}$ полезна видоизменённая формула. Сократим $(n-k)!$ с такими же множителями, входящими в $n!$, и получим

$$\binom{n}{k} = \frac{n}{1} \cdot \frac{n-1}{2} \times \dots \times \frac{n-k+1}{k} = \prod_{1 \leqslant s \leqslant k} \frac{n-s+1}{s}.$$

Большой значок справа означает произведение. В частности,

$$\binom{n}{k} = \binom{n}{k-1} \cdot \frac{n-k+1}{k}.$$

Геометрическая прогрессия. Со школы известно правило суммирования геометрической прогрессии. Здесь мы ограничимся случаем

рисунок

$$1 + x + x^2 + x^3 + x^4 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x},$$

причём сделаем очень важное для всего остатка этого раздела предположение, что x близко к нулю. Отбрасывая тогда маленькое слагаемое в правой части, получаем приближение

Viète 1593

$$\frac{1}{1-x} \approx 1 + x + x^2 + x^3 + x^4 + \dots + x^n.$$

Подставляя сюда $-x$ вместо x, находим

$$\frac{1}{1+x} \approx 1 - x + x^2 - x^3 + x^4 \mp \dots + (-1)^n x^n.$$

Бином Ньютона. В юности Ньютон заметил, и очень гордился этим всю жизнь, что указанная выше формула для $\binom{n}{k}$ как произведение дробей сохраняет смысл при замене натурального n любым числом α. Таким образом, он получил для малых x приближение

Newton 1665
Gregory 1670

$$(1+x)^\alpha \approx 1 + \alpha x + \frac{\alpha}{1} \cdot \frac{\alpha - 1}{2} \cdot x^2 + \frac{\alpha}{1} \cdot \frac{\alpha - 1}{2} \cdot \frac{\alpha - 2}{3} \cdot x^3 + \dots,$$

называемое **биномом Ньютона**.

Существенная разница между натуральным показателем и прочими в том, что при натуральном $\alpha = n$ приближение становится точным равенством после n шагов, в левой-то части полином, а при прочих α точное равенство достигается лишь предельным переходом к сумме бесконечного числа слагаемых. Такие суммы называют **рядами**; о них в нашем курсе ещё будет сказано немало, но позже.

Пример. При $\alpha = -1$ бином Ньютона превращается в бесконечную геометрическую прогрессию.

Пример. При $\alpha = 1/2$ имеем $\binom{\alpha}{1} = 1/2$, $\binom{\alpha}{2} = -1/8$, $\binom{\alpha}{3} = 1/16$. Поэтому Briggs 1624

$$\sqrt{1+x} = (1+x)^{1/2} \approx 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 + \dots .$$ правка 22.01.15

Не имея ещё возможности доказать найденную им общую формулу для степени бинома, Ньютон для начала проверил, что она не противоречит равенству $\sqrt{1+x} \cdot \sqrt{1+x} = 1 + x$.

Упражнение. *Повторите эту проверку с точностью до x^3, умножая разложение на себя.*

Пример. При $\alpha = -1/2$ имеем $\binom{\alpha}{1} = -1/2$, $\binom{\alpha}{2} = 3/8$, $\binom{\alpha}{3} = -5/16$. Нас будет интересовать бином, содержащий $-x^2$ вместо x. Итак,

$$\frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-1/2} \approx 1 - \tfrac{1}{2}(-x^2) + \tfrac{3}{8}(-x^2)^2 - \tfrac{5}{16}(-x^2)^3 + \dots$$
$$= 1 + \tfrac{1}{2}x^2 + \tfrac{3}{8}x^4 + \tfrac{5}{16}x^6 + \dots .$$

Пример (Кинетическая энергия). Едва ли кто-то не слышал о формуле Эйнштейна $E = mc^2$. Согласно устаревшему или упрощённому изложению релятивизма, частица массы покоя m_0, движущаяся со скоростью $v \ll c$, имеет релятивистскую массу $m = m_0/\sqrt{1 - v^2/c^2}$. Напишем её энергию, приблизив релятивистский фактор его разложением из предыдущего упражнения с $x = v/c \ll 1$:

$$E \approx m_0c^2 + \frac{1}{2}m_0c^2\left(\frac{v}{c}\right)^2 + \frac{3}{8}m_0c^2\left(\frac{v}{c}\right)^4 + \dots$$
$$= m_0c^2 + \frac{1}{2}m_0v^2 + \frac{3}{8}m_0v^2\left(\frac{v}{c}\right)^2 + \dots .$$

При $v \ll c$ последнее выписанное слагаемое уже настолько мало, что в нерелятивисткой механике оно не заметно, а опущенные ещё меньше. Доминирующее первое слагаемое присутствует всегда как фон, при любой скорости, и поэтому обычно тоже не заметно. Значит, заметно лишь второе слагаемое, которое и назвали кинетической энергией.

3.2. Приближение функций полиномами

Нельзя ли приближённо изобразить ход любой функции $y = f(x)$ на некотором участке при помощи других, более простых? Чаще всего пользуются первым приближением — касательной к графику:

$$f(x) \approx f(x_0) + f'(x_0) \cdot (x - x_0).$$

Для упрощения формул будем считать, что $x_0 = 0$, чего всегда можно добиться сдвигом независимой переменной.

рисунок

Приближения получше дают полиномы более высоких степеней:

$$f(x) \approx A_0 + A_1 x,$$
$$f(x) \approx A_0 + A_1 x + A_2 x^2,$$
$$f(x) \approx A_0 + A_1 x + A_2 x^2 + A_3 x^3,$$

и так далее. Вблизи $x = 0$ переход от первого приближения ко второму заключается в добавлении относительно небольшой поправки $A_2 x^2$, и далее аналогично. Идея в том, что при малых x каждая следующая поправка значительно меньше предыдущих. Briggs 1624

Формула Тейлора. Коэффициенты A_k вычисляются, исходя из наблюдения, что наилучшее приближение из полиномов фиксированной степени даёт тот, у которого больше всего (старших) производных в точке $x = 0$ совпадают с производными того же порядка приближаемой функции $f(x)$. К обоснованию этого наблюдения мы вернёмся позже, а сейчас найдём реально применяемую формулу для A_k.

Мы уже знаем, что $A_0 = f(0)$ и $A_1 = f'(0)$. Дифференцируя полином $A_0 + A_1 x + A_2 x^2$ дважды, записываем наше требование $f''(0) = 2A_2$ и находим A_2. Таким же образом дифференцируя k раз приближение степени k, найдём общую формулу Maclaurin 17

$$A_k = \frac{f^{(k)}(0)}{k!}.$$

Числитель обозначает значение производной порядка k функции $f(x)$ в точке $x = 0$, а при $k = 0$, значение самой функции $f(x)$.

Итак, вблизи нуля приближение порядка n есть

$$f(x) \approx \sum_{0 \leqslant k \leqslant n} \frac{f^{(k)}(0)}{k!} \cdot x^k.$$

Сдвигая переменную, получаем **полином Тейлора** степени n, дающий наилучшее приближение порядка n вблизи точки $x = x_0$:

Gregory 1671
Newton 1692
Bernoulli 169
Taylor 1712

$$f(x) \approx \sum_{0 \leqslant k \leqslant n} \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k.$$

Однако все самые важные приближения, которые нужно крепко запомнить, делаются при x, близком к нулю. Лёгкость получения общей формулы для приближения вблизи $x = x_0$ из частного случая $x_0 = 0$ это лишь одна причина. Существеннее то, что важнейшим на практике, и в физике как нигде более, является именно случай очень малого параметра x.

Имея уже общий рецепт, далее мы часто не будем заботиться о степени приближающего полинома. В формулах эта небрежность указывается многоточием. При желании быть аккуратнее можно указывать порядок малости игнорируемых слагаемых как $o(x^n)$, например:

$$(1-x^2)^{-1/2} = 1 + \tfrac{1}{2}x^2 + o(x^2).$$

Удобное и строгое обозначение $o(x^n)$ изобрели только в XIX веке. Осваивать его мы начнём в следующем разделе.

Ради демонстрации, пока что мы будем выписывать разложения, удерживая несколько слагаемых — заведомо больше, чем требуется при решении большинства физических задач. Однако первые шаги важнейших разложений нужны везде и их надо запомнить.

Простейшие разложения. Простые выражения для производных

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

при $x = 0$ позволяют найти полиномы Тейлора функций e^x, $\cos x$ и $\sin x$ любой степени непосредственно по формуле Тейлора. Дроби в коэффициентах это просто факториалы в знаменателе, а в разложениях $\cos x$ и $\sin x$ также наблюдается чередование плюсов и минусов:

$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + \frac{x^7}{5040} + \ldots;$$

$$\cos x \approx 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \ldots;$$

$$\sin x \approx x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \ldots.$$

Следует отметить, что приближения $\cos x$, $\sin x$, а также $\operatorname{arctg} x$ любого порядка были известны загадочной индийской школе математиков и астрономов в конце XIV века, причём изложены были в стихах. Сейчас доподлинно не известно, как они тогда были найдены.

Madhava 1400

Упражнение. *Найдите коэффициенты в полиномиальном приближении экспоненты, исходя лишь из свойства* $(e^x)' = e^x$ *и не пользуясь формулой Тейлора. Аналогично для* $\cos x$ *и* $\sin x$ *используйте дифференциальное уравнение* $f'' = -f$.

Euler 1740

Именно решая это уравнение разными способами, Эйлер увидел формулу $e^{ix} = \cos x + i \sin x$.

Упражнение. *Найдите разложения гиперболических функций* $\operatorname{ch} x$ *и* $\operatorname{sh} x$.

Выведем теперь бином Ньютона из формулы Тейлора:

$$
\begin{aligned}
f(x) &= (1+x)^{\alpha},\\
f'(x) &= \alpha(1+x)^{\alpha-1},\\
f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2},\\
f'''(x) &= \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3};
\end{aligned}
$$

таким образом, в значениях производных в точке $x = 0$ накапливается нужное произведение, а после деления на $k!$ из формулы мы и получаем коэффициент $\binom{\alpha}{k}$.

Разложение первообразных. Логарифм, арксинус и арктангенс были разложены раньше, чем экспонента, которую тогда даже не выделяли как основную функцию. Не было ещё ни понятия интеграла, ни формулы Тейлора. Эти разложения стали важными открытиями на пути к основам анализа.

После того, как мы получим их, лишь одна из основных элементарных функций останется ещё не разложенной. Искать разложение $\operatorname{tg} x$ напрямую по формуле Тейлора затруднительно, поскольку старшие производные быстро становятся страшноваты. Эта задача решается более сложными методами и служит им прекрасной иллюстрацией.

Функция $\ln x$ вблизи нуля определена лишь с одной стороны, для $x > 0$, и стремится к $-\infty$ при $x \to 0$. Поэтому она там не приближается полиномом. Актуально её приближение полиномом вблизи $x = 1$, что равносильно приближению функции $\ln(1 + x)$ при малом x. Хотя его можно получить напрямую из формулы Тейлора, полезно увидеть другой, исторически первый подход.

Вспомним, что логарифм определён как первообразная, и сдвинем переменную:

$$\ln(1+x) = \int_0^x \frac{dt}{1+t}.$$

Смелая идея Меркатора заключалась в том, чтобы подставить под интеграл известное разложение функции $(1 + t)^{-1}$, а затем отдельно проинтегрировать каждое слагаемое. В современной записи, Mercator 165

$$
\begin{aligned}
\ln(1+x) &= \int_0^x \bigl(1 - t + t^2 - t^3 + t^4 - t^5 + o(t^5)\bigr)\,dt\\
&= x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \tfrac{1}{5}x^5 - \tfrac{1}{6}x^6 + o(x^6).
\end{aligned}
$$

Упражнение. *Найдите это разложение непосредственно из формулы Тейлора, дифференцируя* $\ln(1 + x)$ *необходимое число раз.*

Ньютон поступил аналогично с другой функцией: ln 2

$$\arcsin x = \int_0^x \frac{dt}{\sqrt{1-t^2}} = \int_0^x \left(1 + \tfrac{1}{2}t^2 + \tfrac{3}{8}t^4 + \tfrac{5}{16}t^6 + o(t^7)\right) dt$$
$$= x + \tfrac{1}{6}x^3 + \tfrac{3}{40}x^5 + \tfrac{5}{112}x^7 + o(x^8).$$

Хотя это разложение редко требует точности выше x^3, всё же можно указать формулу для коэффициентов при любой степени:

$$\tfrac{1}{6} = \tfrac{1}{2} \cdot \tfrac{1}{3}, \quad \tfrac{3}{40} = \tfrac{1\cdot 3}{2\cdot 4} \cdot \tfrac{1}{5}, \quad \tfrac{5}{112} = \tfrac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6} \cdot \tfrac{1}{7}.$$

Gregory 1671 Здесь же следует упомянуть очень простое разложение

$$\operatorname{arctg} x = \int_0^x \frac{dt}{1+t^2} = \int_0^x \left(1 - t^2 + t^4 - t^6 + o(t^7)\right) dt$$
$$= x - \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 - \tfrac{1}{7}x^7 + o(x^8).$$

Упражнение. *Почему в последних двух разложениях указана точность на единицу выше последней выписанной степени?*

Интегрирование разложений элементарных функций с тем же успехом даёт разложения и неэлементарных первообразных.

Пример. Разложим неэлементарную функцию

$$F(x) = \int_0^x e^{-t^2}\, dt = \int_0^x \left(1 - t^2 + \tfrac{1}{2}t^4 - \tfrac{1}{6}t^6 + o(t^7)\right)$$
$$= x - \tfrac{1}{3}x^3 + \tfrac{1}{10}x^5 - \tfrac{1}{42}x^7 + o(x^8).$$

Функция $\operatorname{erf} x = \frac{2}{\sqrt{\pi}} F(x)$ важна в теории вероятностей, математической статистике, а также их приложениях во всех науках, от точных до гуманитарных. Её называют **функцией ошибок**.

3.3. Асимптотические сравнения

Зачастую физики делят (положительные) величины на три группы: много меньше единицы, порядка единицы, много больше единицы. Им соответствуют значки

$$\varepsilon \ll 1, \quad n \sim 1, \quad N \gg 1.$$

Это связано с необходимостью заменять исходно более точные, но чересчур математически сложные уравнения приближёнными. И в математическом анализе регулярно работают с приближениями чисел и функций, причём нередко мы хотим отслеживать их точность.

Самым удобным известным способом указания точности является o-символика, представленная в этом разделе. Широко распространены лишь символы «o-малое» и «O-большое».

Асимптотическая эквивалентность. Чтобы сравнивать поведение функций $f(x)$ и $g(x)$ при $x \to a$, изучают их частное, то есть функцию $\tau(x)$, определяемую равенством $f(x) = \tau(x)g(x)$. При обычном употреблении изучаемой техники можно избегать ритуального повторения уточняющих слов *при* $x \to a$, договорившись, что местами они относятся целиком к рассуждению. Значение a часто ясно из контекста, в особенности из-за доминирования на практике резко контрастирующих случаев $x \to 0$ и $x \to +\infty$.

Итак, простейший случай: если $\tau(x) \to 1$ при $x \to a$, то $f(x)$ и $g(x)$ называют **эквивалентными** при $x \to a$ и пишут $f(x) \sim g(x)$ при $x \to a$.

Примеры. При $x \to 0$ имеем важнейшие эквивалентности $\sin x \sim x$, $e^x \sim 1 + x$, $\ln(1 + x) \sim x$, $\cos x \sim 1 - \frac{1}{2}x^2$, $\operatorname{tg} x \sim x$. выделить!

Эквивалентностями очень часто пользуются, проводя приближённые оценки и вычисления. В них явно не указывается точность, но подразумевается, что при необходимости её можно повысить; например, $\sin x \sim x - \frac{1}{6}x^3$.

Пример (Маятник). Задача о колебании подвешенного тела приводит к дифференциальному уравнению $\ddot{x} + \omega^2 \sin x = 0$, которое не решается в элементарных функциях. Здесь $x(t)$ обозначает угол отклонения маятника от положения равновесия. Однако если колебания очень малы, то $\sin x \sim x$ и уравнение заменяется на более простое, $\ddot{x} + \omega^2 x = 0$. Его решения вам уже известны.

лекция 8
28.09.15

Бесконечно малые. Случай конечного и ненулевого предела функции $\tau(x)$ не интересен, поскольку это фактически эквивалентность с точностью до умножения на константу. Содержательны случаи нулевого и бесконечного пределов. Подробнее мы рассмотрим первый из них. Ещё есть случай, когда $\tau(x)$ не имеет никакого предела. Это может означать, что мы пытаемся сравнивать слишком разные функции.

Определение. Функцию $f(x)$ называют **ограниченной** при $x \to a$ и пишут $f(x) = O(1)$ при $x \to a$, если $|f(x)|$ вблизи $x = a$ меньше некоторой постоянной.

Сумма и произведение ограниченных функций ограничены.

Определение. Функцию $f(x)$ называют **бесконечно малой** при $x \to a$ и пишут $f(x) = o(1)$ при $x \to a$, если $f(x) \to 0$ при $x \to a$.

Сумма и произведение бесконечно малых бесконечно малы.

Ради сравнений с постоянной новый символ бы не вводили. Настоящая область его применения в анализе — сравнение бесконечно малых между собой. Если $f(x) \to 0$, $g(x) \to 0$, и при этом $\tau(x) \to 0$, то функцию $f(x)$ называют бесконечно малой относительно $g(x)$ и пишут $f(x) = o(g(x))$. Поэтому,

$$f(x) = o(g(x)) \Leftrightarrow f(x) = o(1) \cdot g(x).$$

Действительно, тут левая часть означает, что $\tau(x) \to 0$, а правая — что $\tau(x) = o(1)$. Широкая двусторонняя стрелочка указывает логическую равносильность и в дальнейшем будет спокойно использоваться.

Относительную ограниченность записывают как

$$f(x) = O(g(x)) \Leftrightarrow f(x) = O(1) \cdot g(x).$$

Это означает, что $\tau(x)$ ограничена вблизи $x = a$, но речи о её пределе при $x \to a$ не идёт; он может и не существовать.

Упражнение. *Известно, что $f(x) = O(g(x))$ и $g(x) = O(f(x))$ при $x \to a$. Можно ли утверждать, что $f(x) \sim g(x)$ при $x \to a$ с точностью до умножения на константу?*

Пример. Функция $f(x) = 1 - \cos x$ при $x \to 0$ стремится к нулю быстрее чем x. Поэтому $f(x) = o(x)$. Пользуясь полиномом Тейлора, мы видим более точно, что $f(x) \sim x^2/2$ при малом x. В частности, $f(x) = O(x^2)$.

рисунок

Соотношение $f(x) = o(g(x))$ показывает, что $f(x)$ имеет более высокий **порядок малости** чем $g(x)$. Соотношение $f(x) = O(g(x))$ показывает, что порядок малости $f(x)$ не ниже порядка малости $g(x)$.

Понятие ограниченности проще понятия бесконечной малости. Первое не зависит от понятия предела, а второе зависит — а точнее, равносильно ему. Запишем их в форме, подчёркивающей это:

$$f(x) = O(g(x)) \text{ при } x \to a \Leftrightarrow |f(x)| \leqslant K|g(x)| \text{ вблизи } x = a$$

для некоторой постоянной K, указывать которую обычно не требуется;

$$f(x) = o(g(x)) \text{ при } x \to a \Leftrightarrow |f(x)| \leqslant \varepsilon|g(x)| \text{ вблизи } x = a$$

рисунки

для любой постоянной ε. Чем меньше ε, тем ближе приходится подходить к точке $x = a$, чтобы неравенство выполнялось. Здесь заложена та же идея, что и в понятии предела, которое нам ещё предстоит формально записать.

Самый типичный случай — сравнение при $x \to 0$ со степенной функцией x^n. Показатели степени могут быть и нецелыми, хотя такие сравнения в анализе используются редко.

Примеры. Перепишем замечательные пределы при $x \to 0$ на языке бесконечно малых:

$$
\begin{aligned}
\frac{\sin x}{x} - 1 \to 0 \quad &\Longrightarrow \quad \sin x = x + o(x);\\
\frac{1-\cos x}{x^2} - \frac{1}{2} \to 0 \quad &\Longrightarrow \quad \cos x = 1 - \tfrac{1}{2}x^2 + o(x^2);\\
\frac{e^x - 1}{x} - 1 \to 0 \quad &\Longrightarrow \quad e^x = 1 + x + o(x);\\
\frac{\ln(1+x)}{x} - 1 \to 0 \quad &\Longrightarrow \quad \ln(1+x) = x + o(x);\\
\frac{(1+x)^\alpha - 1}{x} - \alpha \to 0 \quad &\Longrightarrow \quad (1+x)^\alpha = 1 + \alpha x + o(x).
\end{aligned}
$$

Разумеется, мы видим здесь начальные шаги разложений Тейлора.

Преобразования. Нельзя обращаться с асимптотическими сравнениями $f(x) = o(g(x))$ и $f(x) = O(g(x))$ как с обычными равенствами. Например, $x^2 = o(x)$ и $x^3 = o(x)$ при $x \to 0$, но $x^2 \neq x^3$.

Иногда предлагают бороться с путаницей, считая $o(g(x))$ и $O(g(x))$ множествами функций и заменяя значок равенства на принадлежность: $f(x) \in o(g(x))$. Этот подход тоже ограничен, ибо он не приносит желаемого удобства в обращении с o-символикой для упрощения выражений, а ведь это приложение интересует нас больше всего.

Лучше понимать под $o(g(x))$ произвольную функцию бесконечно малую относительно $g(x)$. Цепочка преобразований выражений, содержащих такие «произвольности» и традиционно связываемых знаком равенства, как правило необратима: помимо простой алгебры, она включает шаги, на которых мы отбрасываем излишние детали, упрятывая их внутрь o-малых или O-больших. Попробуем первое время ставить вместо равенств особые необратимые стрелочки.

Примеры. Выпишем несколько типичных преобразований.

- Сложение: $o(x^2) + o(x^3) \rightarrowtail o(x^2)$.
- Умножение на степень: $x^3 \cdot o(x) \rightarrowtail o(x^4)$.
- Перемножение: $O(x^3) \cdot o(x) \rightarrowtail o(x^4)$.
- Сокращение: $o(x^5)/x^2 \rightarrowtail o(x^3)$.
- Ослабление: $o(x^2) \rightarrowtail O(x^2)$.

Пример. Цепочка поглощения слагаемого:

$$x^3 + o(x^2) \rightarrowtail o(x^2) + o(x^2) \rightarrowtail o(x^2).$$

Пример. Цепочка поглощения множителя через эквивалентность:

$$\sin^2 x \cdot o(x^3) \sim x^2 \cdot o(x^3) \rightarrowtail o(x^5).$$

Пример. Раскрытие скобок:

$$\begin{aligned}\big(4x + o(x)\big)\big(x^2 + o(x^2)\big) &= 4x^3 + x^2 \cdot o(x) + 4x \cdot o(x^2) + o(x) \cdot o(x^2)\\ &\rightarrowtail 4x^3 + o(x^3) + o(x^3) + o(x^3)\\ &\rightarrowtail 4x^3 + o(x^3).\end{aligned}$$

Несложно обобщить преобразования из первого примера.

Лемма. *Имеются правила сложения и перемножения O-больших:*

$$O(f) + O(g) \rightarrowtail O(f+g), \quad O(f) \cdot O(g) \rightarrowtail O(f \cdot g).$$

Доказательство. Сведём вопрос к ограниченности суммы и произведения ограниченных функций. Вынесем f и g, сложим две ограниченных функции инкогнито внутри $O(1)$, затем вернём $f+g$ обратно:

$$O(f) + O(g) = f \cdot O(1) + g \cdot O(1) = (f+g) \cdot O(1) = O(f+g).$$

Таким же способом выполняем перемножение. □

Лемма. *Имеются правила сложения и перемножения o-малых:*

$$o(f) + o(g) \rightarrowtail o(f+g), \quad o(f) \cdot o(g) \rightarrowtail o(f \cdot g).$$

Доказательство. Полностью аналогично предыдущей лемме. □

Упражнение. *Установите правило сокращения степени:*

$$o(x^n)/x^m \rightarrowtail o(x^{n-m}) \text{ для } n \geqslant m.$$

Упражнение. *Что можно сказать об $o(x^n)/o(x^m)$?*

Лемма. *Условия $f(x) \sim g(x)$ и $f(x) = g(x) + o(g(x))$ равносильны.*

Доказательство. Запишем простое преобразование

$$f(x) = \tau(x)g(x) = g(x) + \big(\tau(x) - 1\big)g(x).$$

Утверждение сводится к равносильности $\tau(x) \to 1$ и $\tau(x) - 1 \to 0$. □

Лемма. *Если $f(x) = o(g(x))$ и $g(x) \sim h(x)$, то $f(x) = o(h(x))$.*

Доказательство. По условию, $f/g \to 0$ и $g/h \to 1$. Перемножая, получаем $f/h \to 0$. □

Пример. Типичное применение последней леммы: $o(\operatorname{tg} x) \rightarrowtail o(x)$.

Бесконечно большие. Выбор обозначения для бесконечно малых и относительно ограниченных функций красиво, но исторически неверно, объясняют сильным внешним сходством греческой буквы омикрон с цифрой 0, а также тем, что в греческом есть «противоположная» буква омега. И её привлекают служить противоположной цели.

Определение. Функцию $f(x)$ называют **бесконечно большой** при $x \to a$, но очень редко пишут $f(x) = \omega(1)$, если $|f(x)| \to \infty$ при $x \to a$.

Если $|f(x)| \to \infty$, $|g(x)| \to \infty$, и при этом $|f(x)/g(x)| \to \infty$, то функцию $f(x)$ называют бесконечно большой относительно $g(x)$, но, увы, очень редко пишут $f(x) = \omega(g(x))$. Итак,

$$f(x) = \omega(g(x)) \Leftrightarrow f(x) = \omega(1)g(x).$$

Соотношение $|f(x)/g(x)| \to \infty$ показывает, что $f(x)$ имеет более высокий **порядок роста** чем $g(x)$.

Пример. Все функции x^α с $\alpha > 0$, а также $\ln x$ и e^x стремятся к $+\infty$ при $x \to +\infty$. При этом $e^x = \omega(x^n)$ и $x^n = \omega(\ln x)$ для любого n.

Сравнения бесконечно больших между собой важны не только в анализе, но и в информатике и теории алгоритмов.

рисунок: шкала?

Определения относительно бесконечно малых и больших никак не используют сами функции, а проверяют лишь предел их отношения (частного). Можно расширить область применимости этой терминологии, отбрасывая лишние предположения $f(x) \to 0$ и $g(x) \to 0$, и аналогично для бесконечности. Тогда получим

$$f(x) = o(g(x)) \text{ при } x \to a \Leftrightarrow g(x) = \omega(f(x)) \text{ при } x \to a.$$

Единственная возникающая при этом проблема — психологическая: относительно бесконечно малой тогда может стать функция, являющаяся абсолютно бесконечно большой, и наоборот.

Пример. При $x \to 0$ имеем $x^{-1} = \omega(1)$ и $x^{-1} = o(x^{-2})$.

3.4. Методы разложения функций

Допустим, нам известны разложения функций $f(x)$ и $g(x)$. Как разложить функции, полученные из них с помощью обычных операций? Применяемые для этого формулы и методы сильно разнятся по сложности, которая зависит от операции. Суммы, произведения и композиции раскладываются бесхитростно и даже рутинно, особенно при хорошем навыке обращения с o-малыми. Частные и обратные функции практичнее раскладывать окольными путями, опирающимися, впрочем, на бесхитростные.

Суммы и разности. Разложить сумму и разность двух функций легко, напрямую применяя эти операции к двум разложениям. Это сводится к раскрытию скобок, собиранию подобных слагаемых, а также отбрасыванию лишних.

Пример. Разложим функцию $\frac{1}{2}\ln\frac{1+x}{1-x}$ по степепям x. Заменим логарифм частного на разность логарифмов и разложим их поточнее:

$$\begin{aligned}\ln(1+x) &= x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 + \tfrac{1}{5}x^5 - \tfrac{1}{6}x^6 + \tfrac{1}{7}x^7 + o(x^7),\\ \ln(1-x) &= -x - \tfrac{1}{2}x^2 - \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 - \tfrac{1}{5}x^5 - \tfrac{1}{6}x^6 - \tfrac{1}{7}x^7 + o(x^7).\end{aligned}$$

В полуразности остаются только нечётные слагаемые:

$$\frac{1}{2}\ln\frac{1+x}{1-x} = x + \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 + \tfrac{1}{7}x^7 + o(x^7).$$

Упражнение. *Сравнивая разложение* $\operatorname{arctg} x$ *и разложение* $\frac{1}{2}\ln\frac{1+x}{1-x}$, *найдите простую формулу, связывающую эти функции.*

Точность разложения не может повыситься при арифметических операциях. Точность результата операции равна наименьшей из точностей исходных разложений (операндов).

Пример. Если известно, что

$$f(x) = 1 + 2x + 3x^2 + o(x^2), \quad g(x) = 1 + x + o(x),$$

то для $f(x) - g(x)$ можно написать лишь разложение $x + o(x)$, тогда как $x + 3x^2 + o(x^2)$ написать нельзя, ибо коэффициент при x^2 в разложении $g(x)$ не известен. К такому же заключению приводят правила поглощения бесконечно малыми: $x + 3x^2 + o(x^2) - o(x) \rightarrowtail x + o(x)$.

Произведения. Произведения раскладываются также напрямую, но раскрытие скобок при этом требует чуть больше труда. Кроме того, оно даёт много лишних слагаемых.

Пример. Разложим до x^5 функцию $e^x \sin x$. Выписываем оба множителя с такой же точностью:

$$\begin{aligned} e^x &= 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \tfrac{1}{24}x^4 + \tfrac{1}{120}x^5 + o(x^5),\\ \sin x &= x - \tfrac{1}{6}x^3 + \tfrac{1}{120}x^5 + o(x^5).\end{aligned}$$

Слагаемые, появляющиеся в раскрытии перемноженных скобок, удобно учитывать от низких степеней к более высоким, на ходу собирая подобные. Находим таким образом

$$x + x^2 + \left(\tfrac{1}{2} - \tfrac{1}{6}\right)x^3 + \left(\tfrac{1}{6} - \tfrac{1}{6}\right)x^4 + \left(\tfrac{1}{24} - \tfrac{1}{12} + \tfrac{1}{120}\right)x^5 + o(x^5).$$

Отслеживать коэффициенты при старших степенях уже не нужно, ведь мы всё равно отбросим их, согласно требованиям задачи. Остаётся упростить дроби, и в итоге

$$e^x \sin x = x + x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{30}x^5 + o(x^5).$$

Упражнение. *Найдите это разложение непосредственно из формулы Тейлора, дифференцируя $e^x \sin x$ необходимое число раз.*

Композиции. Простейшие случаи разложения композиции несложно заметить в первых же примерах с биномами: там мы перешли от $f(x) = (1-x)^{-1}$ к $f(-x)$ и от $f(x) = (1+x)^{-1/2}$ к $f(-x^2)$.

Разложение композиции $f(g(x))$ находится из разложения $f(x)$ подстановкой разложения $g(x)$ вместо x. Затем полученное сложное выражение нужно упростить, раскрывая скобки. При этом используют все те же приёмы, что и для разложения сумм и произведений.

Пример. Разложим до x^3 функцию $e^{g(x)}$, где $g(x) = x - \frac{1}{2}x^2$. Начинаем с подстановки $g(x)$ в разложение экспоненты:

$$\begin{aligned} e^{g(x)} &= 1 + g(x) + \tfrac{1}{2}g(x)^2 + \tfrac{1}{6}g(x)^3 + o\big(g(x)^3\big) \\ &= 1 + \big(x - \tfrac{1}{2}x^2\big) + \tfrac{1}{2}\big(x - \tfrac{1}{2}x^2\big)^2 + \tfrac{1}{6}\big(x - \tfrac{1}{2}x^2\big)^3 + o(x^3). \end{aligned}$$

Заметьте упрощение в o-малом на основе эквивалентности $g(x) \sim x$. Теперь раскрываем скобки и собираем подобные:

$$\begin{aligned} \exp\big(x - \tfrac{1}{2}x^2\big) &= 1 + x + \left(\tfrac{1}{2} - \tfrac{1}{2}\right)x^2 + \left(\tfrac{1}{6} - \tfrac{1}{2}\right)x^3 + o(x^3) \\ &= 1 + x - \tfrac{1}{3}x^3 + o(x^3). \end{aligned}$$

В первом примере специально взят бином, чтобы показать все действия на минимальном нетривиальном примере. Однако при более сложной внутренней функции, заменяемой её разложением, действия изменяются не столь существенно, лишь добавляются манипуляции с o-малыми, уже применявшиеся при разложении произведений.

Пример. Разложим до x^4 функцию $h(x) = \ln\big(x + \sqrt{1+x^2}\big)$. Здесь нужно использовать разложение $\ln(1+x)$, поэтому заменим сперва внутреннюю функцию на её разложение

$$x + \sqrt{1+x^2} = 1 + x + \tfrac{1}{2}x^2 - \tfrac{1}{8}x^4 + o(x^4) = 1 + g(x)$$

и далее будем подставлять в скобки разложение $g(x)$:

$$
\begin{aligned}
h(x) &= g(x) - \tfrac{1}{2}g(x)^2 + \tfrac{1}{3}g(x)^3 - \tfrac{1}{4}g(x)^4 + o\big(g(x)^4\big) \\
&= \big(x + \tfrac{1}{2}x^2 - \tfrac{1}{8}x^4 + o(x^4)\big) - \tfrac{1}{2}\big(\square\big)^2 + \tfrac{1}{3}\big(\square\big)^3 - \tfrac{1}{4}\big(\square\big)^4 + o(x^4) \\
&= x + \big(\tfrac{1}{2} - \tfrac{1}{2}\big)x^2 + \big(-\tfrac{1}{2} + \tfrac{1}{3}\big)x^3 + \big(-\tfrac{1}{8} - \tfrac{1}{8} + \tfrac{1}{2} - \tfrac{1}{4}\big)x^4 + o(x^4) \\
&= x - \tfrac{1}{6}x^3 + o(x^4).
\end{aligned}
$$

Все удержанные чётные коэффициенты благополучно занулились.

Частные. Разложить частное $h(x) = f(x)/g(x)$ в общем виде не удаётся: получается сложная и непрактичная формула. Причина в том, что деление полиномов обычно даёт дробь, а не полином. Приходится делать несколько шагов, в том числе разлагая композицию, или искать обходные пути. Естественный обходной путь состоит в разложении обеих частей равенства $f(x) = g(x)h(x)$ с использованием неопределённых коэффициентов.

Пример. Ищем коэффициенты разложения

$$\operatorname{tg} x \approx a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + \dots$$

необходимой нам точности, умножая его на известное разложение

$$\cos x \approx 1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 + \dots$$

той же точности, приводя подобные,

$$\cos x \operatorname{tg} x \approx a_0 + a_1 x + (?_2)x^2 + (?_3)x^3 + (?_4)x^4 + (?_5)x^5 + \dots,$$

а затем приравнивая коэффициент при каждой степени x соответствующему коэффициенту в известном разложении

$$\sin x \approx x - \tfrac{1}{6}x^3 + \tfrac{1}{120}x^5 + \dots$$

той же точности. Из полученной системы линейных уравнений

$$
\begin{aligned}
0 &= a_0, & 1 &= a_1, \\
0 &= a_2 - \tfrac{1}{2}a_0, & -\tfrac{1}{6} &= a_3 - \tfrac{1}{2}a_1, \\
0 &= a_4 - \tfrac{1}{2}a_2 + \tfrac{1}{24}a_0, & \tfrac{1}{120} &= a_5 - \tfrac{1}{2}a_3 + \tfrac{1}{24}a_1, \\
&\dots & &\dots
\end{aligned}
$$

находим неизвестные коэффициенты один за другим, идя от низших степеней к высшим. Из выписанных здесь строк получим

$$\operatorname{tg} x = x + \tfrac{1}{3}x^3 + \tfrac{2}{15}x^5 + o(x^5).$$

Упражнение. *Разложите* $\operatorname{tg} x$ *с точностью до* x^7*.*

Видно, что в разложении тангенса чётные коэффициенты нулевые. Это всегда так для нечётных функций, $f(-x) = -f(x)$. Наоборот, в разложении чётных функций, $f(-x) = f(x)$, присутствуют только чётные коэффициенты. Мы уже видели этот эффект на примерах синуса и косинуса, а также в двух примерах этого раздела с логарифмами. Сделанные заранее, подобные наблюдения могут заметно сократить писанину.

Упражнение. *Установите, что* $\ln\bigl(x+\sqrt{1+x^2}\bigr)$ *является нечётной функцией.*

Упражнение. *Разложите до* x^7 *функцию* $\sec x = 1/\cos x$ *по методу неопределённых коэффициентов.*

Упражнение. *Разложите до* x^5 *функцию* $\sec x = 1/\cos x$*, представляя её как композицию функций* $(1+x)^{-1}$ *и* $\cos x - 1$*.*

Обратные и неявные функции. Подход к разложению частных обобщается на различные уравнения, содержащие неизвестную функцию. Фактически речь идёт о разложении неявных функций.

Пример. Функция $y(x)$ задана уравнением $xy^2+y+1 = 0$ вблизи точки $(0,-1)$. Разложим её до x^3.

Подставляя разложение

$$y(x) = -1 + a_1x + a_2x^2 + a_3x^3 + o(x^3)$$

в уравнение, запишем

$$x \cdot \bigl(-1 + a_1x + a_2x^2 + a_3x^3 + o(x^3)\bigr)^2 + \bigl(-1 + a_1x + a_2x^2 + a_3x^3 + o(x^3)\bigr) + 1 = 0.$$

Раскроем скобки и приведём подобные в низших степенях:

$$(a_1+1)x + (a_2-2a_1)x^2 + (a_3-2a_2+a_1^2)x^3 + o(x^3) = 0.$$

Отсюда последовательно найдём коэффициенты и получим

$$y(x) = -1 - x - 2x^2 - 5x^3 + o(x^3).$$

Пример. Функция $y(x)$ задана уравнением $x = y\cos y$ вблизи точки $(0,0)$. Разложим её до x^3.

Заметим прежде всего, что функция $x(y) = y\cos y$ нечётна. Значит, искомая обратная функция $y(x)$ также нечётна, а потому все чётные коэффициенты её разложения нулевые:

$$y(x) = a_1x + a_3x^3 + o(x^3).$$

Подставим это в разложение уравнения по степеням y:

$$\begin{aligned} x &= y - \tfrac{1}{2}y^3 + o(y^3) \\ &= \big(a_1x + a_3x^3 + o(x^3)\big) - \tfrac{1}{2}\big(a_1x + a_3x^3 + o(x^3)\big)^3 + o(x^3) \\ &= \big(a_1x + a_3x^3 + o(x^3)\big) - \tfrac{1}{2}\big(a_1x + o(x)\big)^3 + o(x^3) \\ &= a_1x + (a_3 - \tfrac{1}{2}a_1^3)x^3 + o(x^3) \end{aligned}$$

Упрощение выделенных слагаемых допустимо, поскольку после возведения всей скобки в куб их вклад будет только в x^5 и выше, которые в задаче не требуется удерживать. Получим

$$y(x) = x + \tfrac{1}{2}x^3 + o(x^3).$$

лекция 9
01.10.15

Хотя отыскание слагаемых разложения по методу неопределённых коэффициентов движется пошагово, все необходимые коэффициенты включаются в рассмотрение сразу. Любимый физиками метод последовательных приближений организован иначе: вычисление каждого коэффициента состоит из нескольких действий, цепочка которых повторяется до тех пор, пока не будет достигнута требуемая точность. Суть метода та же, запись иная, а значит, чуть иной и ход мыслей.

Пример. Функция $y(x)$ задана уравнением $x = y\cos y$ вблизи точки $(0,0)$. Разложим её до x^5.

В нулевом приближении берём $y \approx y_0 = 0$.

В первом приближении берём $y \approx y_1 = y_0 + \delta_0$, где $\delta_0 \ll 1$. Подставляем в уравнение, опуская индексы переменных:

$$x \approx \delta\cos\delta \approx \delta - \tfrac{1}{2}\delta^3 \approx \delta.$$

Пренебрегаем δ^3 как малым более высокого порядка. Поэтому $y_1 = x$.

Во втором приближении берём $y \approx y_2 = y_1 + \delta_1$, где $\delta_1 \ll \delta_0 = x$. Подставляем в уравнение, опуская индексы переменных:

$$\begin{aligned} x \approx (x+\delta)\cos(x+\delta) &\approx (x+\delta) - \tfrac{1}{2}(x+\delta)^3 \\ &\approx x + \delta - \tfrac{1}{2}(x^3 + 3x^2\delta + 3x\delta^2 + \delta^3) \\ &\approx x - \tfrac{1}{2}x^3 + \delta. \end{aligned}$$

Пренебрегаем малыми по сравнению с δ и получаем $y_2 = x + \tfrac{1}{2}x^3$.

В третьем приближении берём $y \approx y_3 = y_2 + \delta_2$, где $\delta_2 \ll \delta_1 \sim x^3$. Подставляем в уравнение, опуская индексы переменных:

$$\begin{aligned} x &\approx (x + \tfrac{1}{2}x^3 + \delta) - \tfrac{1}{2}(x + \tfrac{1}{2}x^3 + \delta)^3 + \tfrac{1}{24}(x + \tfrac{1}{2}x^3 + \delta)^5 \\ &\approx (x + \tfrac{1}{2}x^3 + \delta) - \tfrac{1}{2}(x^3 + \tfrac{3}{2}x^5 + \tfrac{3}{4}x^7 + \tfrac{1}{8}x^9) + \tfrac{1}{24}x^5 \\ &\approx x - (\tfrac{3}{4} - \tfrac{1}{24})x^5 + \delta. \end{aligned}$$

Пренебрегаем малыми и получаем $y_3 = x + \frac{1}{2}x^3 + \frac{17}{24}x^5$.

До такого приближения на практике дело доходит нечасто.

Упражнение. *Разложите до x^5 функцию $y = \operatorname{tg} x$ как обратную к $x = \operatorname{arctg} y$ обоими методами.*

Решения дифференциальных уравнений. Метод неопределённых коэффициентов отлично работает для отыскания решений в виде разложений. При этом мы минуем необходимость решать уравнение аналитически, что особенно важно, ибо аналитическое решение известно только для простых типов уравнений. Обоснование такого подхода даётся в курсе дифференциальных уравнений.

Пример. Найдём решения уравнения $y'(x) = y(x)$ в виде разложения

$$y(x) \approx a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + \ldots$$

любой необходимой точности. Дифференцируя почленно, получим

$$y'(x) \approx a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4 + \ldots.$$

Приравнивая коэффициенты при одинаковых степенях x в этих двух разложениях, находим последовательно

$$a_1 = a_0, \quad a_2 = a_1/2, \quad a_3 = a_2/3, \quad a_4 = a_3/4, \quad \ldots.$$

Все коэффициенты выражаются через a_0. Поэтому

$$y(x) \approx a_0\bigl(1 + x + \tfrac{1}{2!}x^2 + \tfrac{1}{3!}x^3 + \tfrac{1}{4!}x^4 + \tfrac{1}{5!}x^5 + \ldots\bigr).$$

Тогда $a_0 = y(0)$, а в скобках мы видим разложение экспоненты. Как вы помните, точным решением является $y(x) = y(0)e^x$.

Пример. Найдём решения уравнения $y' = -2xy$ в виде разложения. Тут разложение правой части есть

$$-2xy \approx -2a_0x - 2a_1x^2 - 2a_2x^3 - 2a_3x^4 - 2a_4x^5 + \ldots.$$

Сравнение коэффициентов с разложением y' даёт $a_1 = 0$, а далее каждый a_n выражается через a_{n-2}. Значит, все нечётные коэффициенты равны нулю. Находим чётные:

$$a_2 = -a_0, \quad a_4 = -a_2/2, \quad a_6 = -a_4/3, \quad \ldots.$$

Отсюда находим

$$y(x) \approx a_0\bigl(1 - x^2 + \tfrac{1}{2!}x^4 - \tfrac{1}{3!}x^6 + \dots\bigr).$$

Здесь опять экспонента: точным решением является $y(x) = y(0)e^{-x^2}$.

Пример. Вернёмся к тангенсу и вспомним, что $(\operatorname{tg} x)' = 1 + \operatorname{tg}^2 x$. Оказывается, что один из быстрейших способов получить разложение $\operatorname{tg} x$ — решить дифференциальное уравнение $y' = 1 + y^2$ с начальным условием $y(0) = 0$. Сразу видно, что $y'(0) = 1$. Также считаем известной нечётность искомой функции. Тогда

$$\begin{aligned} 1 + y^2 &= 1 + \bigl(x + a_3x^3 + a_5x^5 + a_7x^7 + o(x^7)\bigr)^2 \\ &= 1 + x^2 + 2a_3x^4 + (2a_5 + a_3^2)x^6 + (2a_7 + a_3a_5)x^8 + o(x^8) \end{aligned}$$

сравниваем почленно с

$$y' = 1 + 3a_3x^2 + 5a_5x^4 + 7a_7x^6 + 9a_9x^8 + o(x^8)$$

и находим последовательно

$$a_3 = {}^1\!/_3, \quad a_5 = {}^2\!/_{15}, \quad a_7 = {}^{17}\!/_{315}, \quad a_9 = {}^{62}\!/_{2835}.$$

Общая формула для всех коэффициентов в точном бесконечном разложении функции $\operatorname{tg} x$ весьма любопытна, но выходит за тесные рамки нашего курса. В частности, она содержит знаменитые числа Бернулли, связанные с такими разными областями математики, как разностное исчисление, комбинаторика и теория (простых) чисел.

Как показывает последний пример, нелинейные уравнения можно решать методом неопределённых коэффициентов наравне с линейными, пока рекурсии для последовательного вычисления коэффициентов пробиваемы. Однако более хитрая зависимость уравнения от неизвестной функции резко усложняет теорию и поведение решений.

Пример. Найдём решения уравнения $y'' = -y$ в виде разложения. Сперва разложим y'', дифференцируя разложение y':

$$y''(x) \approx 1 \cdot 2a_2 + 2 \cdot 3a_3x + 3 \cdot 4a_4x^2 + 4 \cdot 5a_5x^3 + \dots.$$

Сравниваем с разложением $y(x)$:

$$\begin{aligned} &a_2 = -a_0/(1 \cdot 2), \quad a_4 = -a_2/(3 \cdot 4), \quad a_6 = -a_4/(5 \cdot 6), \quad \dots, \\ &a_3 = -a_1/(2 \cdot 3), \quad a_5 = -a_3/(4 \cdot 5), \quad a_7 = -a_5/(6 \cdot 7), \quad \dots. \end{aligned}$$

В итоге выражаем все чётные коэффициенты через a_0, все нечётные через a_1, а в дробях опять же собираются факториалы. Получаем

$$y(x) \approx a_0\bigl(1 - \tfrac{1}{2!}x^2 + \tfrac{1}{4!}x^4 \mp \dots\bigr) + a_1\bigl(x - \tfrac{1}{3!}x^3 + \tfrac{1}{5!}x^5 \mp \dots\bigr).$$

В скобках видим разложения косинуса и синуса. Как вы помните, точным решением является $y(x) = y(0)\cos x + y'(0)\sin x$.

Теперь мы готовы рассмотреть более сложный пример: уравнение, не решаемое в элементарных функциях. Оно возникло в астрономии и оптике, а позже его решения, функции Эйри, стали ключевыми в некоторых важных вопросах квантовой механики (здесь вам придётся подождать до третьего курса). Многие специальные функции вводят, изучают и применяют именно по той причине, что они удовлетворяют похожим уравнениям, возникшим в прикладных задачах.

Пример (Уравнение Эйри). Найдём решения уравнения $y'' = xy$ в виде разложения. Сравнение разложений обеих частей уравнения даёт

$$\begin{aligned} &a_3 = a_0/(2\cdot 3), \quad a_6 = a_3/(5\cdot 6), \quad a_9 = a_6/(8\cdot 9), \quad \dots, \\ &a_4 = a_1/(3\cdot 4), \quad a_7 = a_4/(6\cdot 7), \quad a_{10} = a_7/(9\cdot 10), \quad \dots. \end{aligned}$$

Кроме того, $a_2 = 0$ и далее зануляются a_5, a_8, …. Поэтому решения уравнения раскладываются как

$$a_0\bigl(1 + \tfrac{1}{3!}x^3 + \tfrac{1\cdot 4}{6!}x^6 + \dots\bigr) + a_1\bigl(x + \tfrac{2}{4!}x^4 + \tfrac{2\cdot 5}{7!}x^7 + \dots\bigr).$$

Выбор значений a_0 и a_1, дающих две стандартные функции Эйри, нам сейчас не важен. Главное то, что принципиально эти специальные функции ничем не хуже хорошо знакомых тригонометрических.

Глава 4. ТЕОРЕМЫ О СРЕДНЕМ И ПРИЛОЖЕНИЯ

4.1. Теоремы о среднем

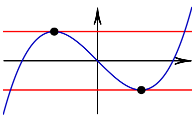
лекция 9+ 01.10.15

Введём сперва одно из самых характерных понятий анализа.

Определение. **Окрестностью** некоторой точки называют любой содержащий её интервал, особенно с целью подчеркнуть его малость.

Фразы «*в некоторой окрестности*» и (очень часто в этих лекциях) «*вблизи*» точки p означают: «*можно найти такую окрестность точки p, в которой*».

Локальные экстремумы. Точкой **локального максимума** функции f на промежутке Δ называют точку $p \in \Delta$, вблизи которой для всех x выполнено $f(x) \leqslant f(p)$. Противоположным неравенством $f(x) \geqslant f(p)$ аналогично определяется точка **локального минимума**. Как и глобальные, локальные максимумы и минимумы совместно называют **локальными экстремумами**. Строгими неравенствами аналогично определяются точки строгого локального экстремума.



Fermat 1637?

Теорема (Ферма)**.** *Если во внутренней точке локального экстремума функции её производная существует, то она равна нулю.*

Доказательство. Физически очевидно, что в точке максимального удаления скорость удаления зануляется. Геометрически столь же очевидно, что касательная к графику в точке максимума должна быть горизонтальна. Аналогично и для минимума.

Более формально, вблизи максимума наклон хорды удовлетворяет

$$\frac{f(x) - f(p)}{x - p} \geqslant 0 \qquad \text{при } x < p$$

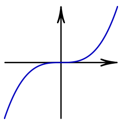
и противоположному неравенству при $x > p$. Поскольку предел этой дроби при $x \to p$ существует, он обязан быть равен нулю. $\square$

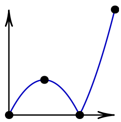
Следствие. *Экстремум непрерывной функции f на отрезке достигается в точке одного из трёх типов:*

- *в граничной точке;*
- *в точке, где производная не существует;*
- *в точке, где производная равна нулю.*

Это даёт нам способ поиска подозрительных на экстремум точек. Получив их список, нужно лишь сравнить значения функции в них и тогда мы найдём глобальные экстремумы. Для определения локальных экстремумов требуются дополнительные действия, поскольку таковым является не каждая подозрительная точка.

Пример. Функция $f(x) = x^3$ на отрезке $[-1, 1]$ имеет подозрительную точку $x = 0$, а экстремума в ней нет.





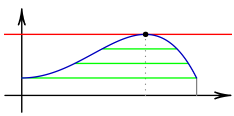
Пример. Функция $f(x) = |x^2 - 2x|$ на отрезке $[0, 3]$ имеет четыре локальных экстремума.

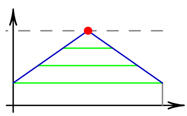
Дарбу?

Теорема Ролля. Следующие три теоремы также весьма очевидны геометрически, но притом имеют многочисленные и разнообразные приложения. Часто какой-нибудь «очевидный» шаг в доказательстве строго реализуется со ссылкой на теорему о среднем.

Функцию называют **дифференцируемой** на числовом множестве, если она имеет производную во всех его точках.

Теорема (Ролль). *Если функция f дифференцируема на отрезке $[a, b]$ и $f(a) = f(b)$, то найдётся такая точка $p \in (a, b)$, что $f'(p) = 0$.* Rolle 1690





На интуитивном уровне утверждение теоремы Ролля ясно из картинки: касательная параллельна хорде. Другая картинка показывает, что нельзя допускать отсутствия производной даже в одной точке.

Доказательство. Считаем очевидными два свойства:

(1) всякая дифференцируемая функция непрерывна;

(2) всякая непрерывная функция на отрезке имеет точки минимума и максимума.

Оба свойства правдоподобны и принимались без подозрений до формализации анализа. На самом деле, позже это будут две теоремы.

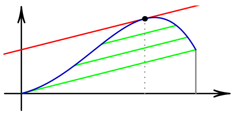
Итак, на отрезке имеются точки, где функция f принимает минимальное и максимальное значения. Если и минимум, и максимум находятся на краю, то f постоянна, а её производная всюду нулевая. Иначе имеется внутренний экстремум p, и по теореме Ферма $f'(p) = 0$. $\square$

Следствие. *Если функция f дифференцируема n раз на отрезке $[a, b]$ и принимает равные значения в $n+1$ его точке, то найдётся такая точка $p \in (a, b)$, что $f^{(n)}(p) = 0$.*

Упражнение. *Выведите следствие Ролля из теоремы Ролля.*

Теоремы Лагранжа и Коши. Столь раннее появление теоремы Ролля объяснимо тем, что сам Ролль изучал исключительно полиномы и действовал алгебраически. Её общность и значение в анализе осознали лишь во времена Лагранжа.

Центральной же в этом ряду теорем является следующая.



Lagrange 1797

Теорема (Лагранж). *Если функция f дифференцируема на отрезке $[a, b]$, то найдётся такая точка $p \in (a, b)$, что*

$$\frac{f(b) - f(a)}{b - a} = f'(p).$$

Это утверждение известно как теорема о среднем или как **формула конечных приращений**, особенно в форме

$$f(b) - f(a) = f'(p) \cdot (b - a).$$

Доказательство. Идея доказательства прозрачна из рисунка: дробь в левой части равна наклону отрезка, соединяющего концы графика. Обозначим её через k и рассмотрим дифференцируемую функцию $\varphi(x) = f(x) - kx$. Поскольку $\varphi(a) = \varphi(b)$, теорема Ролля даёт точку p, в которой $\varphi'(p) = 0$. Значит, $f'(p) = k$. $\square$

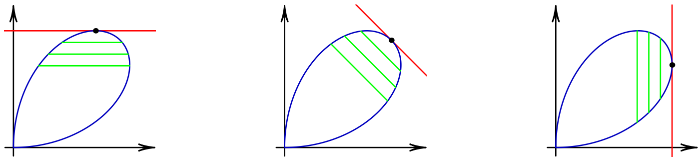
Cauchy 1821

Теорема (Коши). *Если функции f и g дифференцируемы на отрезке $[a, b]$ и всюду на нём $g'(x) \neq 0$, то $g(a) \neq g(b)$ и найдётся такая точка*

$p \in (a, b)$, *что*

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f'(p)}{g'(p)}.$$

Доказательство. Равенство $g(a) = g(b)$ противоречит теореме Ролля.



Частным случаем $g(x) = x$ является теорема Лагранжа, и её доказательство нетрудно обобщить. Обозначим дробь через k и рассмотрим дифференцируемую функцию $\varphi(x) = f(x) - kg(x)$. Поскольку опять $\varphi(a) = \varphi(b)$, теорема Ролля даёт точку p, в которой $\varphi'(p) = 0$, а тогда $f'(p)/g'(p) = k$. □

Отметим, что в приведённых формулировках дифференцируемость на всём отрезке предполагается только для краткости. Производные на концах отрезка не используются, поэтому в книгах формулировки этих теорем обычно предполагают, что функции непрерывны на отрезке и дифференцируемы внутри него. Это замечание относится и к следующим разделам, содержащим довольно простые и нудные приложения теоремы о среднем к исследованию функций. Нудные, ибо Лагранж, развивая их, стремился полностью избавиться от чертежей.

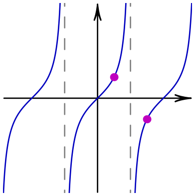
4.2. Монотонность и экстремумы

Участки монотонности. Функцию $f(x)$ называют **монотонной** на числовом множестве X, если для всех пар $x_1 < x_2$ из X сохраняется одно и то же неравенство между $f(x_1)$ и $f(x_2)$. Точнее, в зависимости от вида неравенства: лекция 10 05.10.15

- если $f(x_1) \geqslant f(x_2)$, то $f(x)$ **невозрастающая**;
- если $f(x_1) \leqslant f(x_2)$, то $f(x)$ **неубывающая**;
- если $f(x_1) > f(x_2)$, то $f(x)$ **убывающая**;
- если $f(x_1) < f(x_2)$, то $f(x)$ **возрастающая**.

Когда хотят объединить случаи убывания и возрастания, говорят о **строгой** монотонности. Во избежание путаницы лучше изучать монотонность только на промежутках.

Пример. Функция $f(x) = \operatorname{tg} x$ возрастает на каждом промежутке области определения, но при этом $\operatorname{tg}(\pi/4) > \operatorname{tg}(3\pi/4)$.



Упражнение. *Выпишите наибольшие промежутки монотонности каждой из основных элементарных функций.*

Далее вместо отрезка $[a, b]$ мы будем работать на промежутке Δ любого типа.

Следствие. *Если функция f дифференцируема на Δ, то*

- $f'(x) \geqslant 0 \Leftrightarrow f(x)$ *не убывает;*
- $f'(x) \leqslant 0 \Leftrightarrow f(x)$ *не возрастает;*
- $f'(x) = 0 \Leftrightarrow f(x)$ *постоянна;*
- $f'(x) > 0 \Rightarrow f(x)$ *возрастает;*
- $f'(x) < 0 \Rightarrow f(x)$ *убывает.*

Здесь все 8 половинок высказываний делаются на всём Δ.

Доказательство. Рассмотрим первый случай.

($\Rightarrow$) Для отрезка $[x_1, x_2] \subseteq \Delta$ по теореме Лагранжа получаем

$$f(x_2) - f(x_1) = f'(p) \cdot (x_2 - x_1),$$

поэтому $f'(p) \geqslant 0$ влечёт неубывание.

($\Leftarrow$) Для любой пары точек u, $x \in \Delta$ имеем

$$\frac{f(u) - f(x)}{u - x} \geqslant 0,$$

откуда предельный переход при $u \to x$ даёт $f'(x) \geqslant 0$.

Другие случаи аналогичны. Из-за того, что предельный переход превращает строгое неравенство в нестрогое, обратные импликации при возрастании и убывании неверны; для этих случаев легко найти контрпримеры. $\square$

Пример. Функция $f(x) = x^3$ всюду возрастает, хотя $f'(0) = 0$.

Условия локального экстремума. Стандартный подход к задаче поиска локальных экстремумов функции состоит из двух шагов:

(1) проверкой **необходимого условия** на всём промежутке определяем, где экстремумы возможны;

(2) проверкой **достаточного условия** в каждой из возможных точек определяем, какие из них действительно экстремумы.

Как необходимое условие локального экстремума обычно используют теорему Ферма о нуле производной. Два достаточных условия в терминах первых и вторых производных следуют из теоремы Лагранжа о среднем.

Следствие (Первое достаточное условие экстремума). *Вблизи каждой точки локального экстремума дифференцируемой функции её производная меняет знак, а вблизи прочих точек — сохраняет его:* рисунки?

- *если* $f'(x) < 0$ *слева и* $f'(x) > 0$ *справа от точки* $x = p$, *то в ней* $f(x)$ *имеет строгий локальный минимум;*
- *если* $f'(x) > 0$ *слева и* $f'(x) < 0$ *справа от точки* $x = p$, *то в ней* $f(x)$ *имеет строгий локальный максимум;*
- *если знаки* $f'(x)$ *слева и справа от точки* $x = p$ *одинаковые, то она не является строгим локальным экстремумом.*

Доказательство. Применим предыдущее следствие отдельно слева и справа от p, причём включая её в $\Delta = (a, p]$ и $\Delta = [p, b)$. При этом точки a и b должны выбираться так, чтобы на каждом из полуинтервалов $f'(x)$ сохраняла знак. □

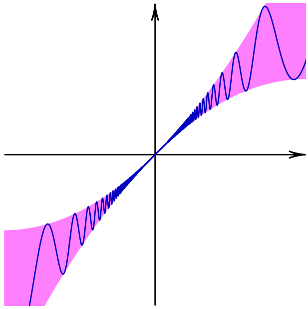
В первом достаточном условии экстремума не используется производная в само́й подозрительной точке, так что с тем же успехом можно предполагать функцию дифференцируемой на выколотой окрестности точки p и непрерывной в самой точке.

Перемены знака вблизи точки $x = p$ по одну сторону от неё могут происходить при любом выборе её окрестности (a, b). Тогда первое достаточное условие не работает и локального экстремума при $x = p$ может не быть.

Пример. Подобные контрпримеры часто строят на основе бешеных колебаний функции $\sin\frac{1}{x}$. В этом конкретном случае годится функция $f(x) = x + x^2\sin\frac{1}{x^2}$ с устранённым разрывом в точке $x = 0$. График колеблется в закрашенной области между параболами $y = x + x^2$ и $y = x - x^2$.

Следствие (Второе достаточное условие экстремума). *Для всякой функции* f *дважды дифференцируемой на интервале в каждой его точке с условием* $f'(p) = 0$ *выполнено:* рисунок?

- *если* $f''(p) \neq 0$, *то* p — *строгий локальный экстремум;*



- *если $f''(p) < 0$, то p — строгий локальный максимум;*
- *если $f''(p) > 0$, то p — строгий локальный минимум.*

Доказательство. При $f''(p) > 0$ найдётся окрестность, на которой

$$\frac{f'(x) - f'(p)}{x - p} > 0,$$

поэтому $f'(x) < 0$ слева от точки $x = p$ и $f'(x) > 0$ справа от неё. По предыдущему следствию там находится минимум.

Случай $f''(p) < 0$ совершенно аналогичен. □

Если не только $f'(p) = 0$, но и $f''(p) = 0$, то для выяснения, есть ли экстремум в точке $x = p$, нужно проверять высшие производные. Возможные варианты происходящего становятся понятны из разложения функции с удержанием низшей из производных, отличной от нуля в подозрительной точке:

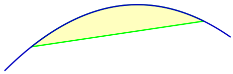
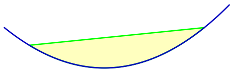
$$f(x) = f(p) + \tfrac{1}{k!} f^{(k)}(p)(x - p)^k + o\big(|x - p|^k\big).$$

Вблизи $x = p$ график такой функции очень похож на сдвинутый график степенной функции. Поэтому локальный экстремум есть тогда и только тогда, когда k чётно. Вскоре мы увидим, что существование всех удерживаемых производных вблизи точки $x = p$ гарантирует наличие используемого тут разложения.

рисунок?

4.3. Выпуклость и перегибы

Участки выпуклости. Множество на плоскости (или даже в пространстве) называют **выпуклым**, если отрезок, соединяющий любую пару его точек, целиком лежит в нём. Функцию называют **выпуклой вверх/вниз** на промежутке Δ, если любая дуга её графика на Δ проходит (нестрого) над/под стягивающей эту дугу хордой. При этом часть плоскости между любой дугой и стягивающей её хордой, включающая их самих, является выпуклым множеством.

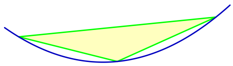
Ключевую роль в исследовании выпуклости играет наклон хорды графика как функция абсцисс её концов. Поэтому положим

$$k(u,v) = \frac{f(v)-f(u)}{v-u} = k(v,u).$$

Лемма. *Выпуклость вниз функции f на промежутке Δ равносильна выполнению неравенства $k(u,x) \leqslant k(x,v)$ для любых троек $u<x<v$ точек в Δ.*

Доказательство. Равенство $k(u,x) = k(x,v)$ реализуется только на прямолинейном участке графика.

Неравенства $k(u,x) < k(u,v) < k(x,v)$ справедливы в точности тогда, когда треугольник, образованный соответствующими хордами, лежит под хордой, соединяющей крайние точки. $\square$



Упражнение. *Определите строгую выпуклость функции и дайте соответствующую лемму.*

Теорема. *Если функция f дифференцируема на Δ, то*

- *$f'(x)$ не убывает $\Leftrightarrow$ $f(x)$ выпукла вниз;*
- *$f'(x)$ не возрастает $\Leftrightarrow$ $f(x)$ выпукла вверх;*
- *$f'(x)$ постоянна $\Leftrightarrow$ $f(x)$ линейна;*
- *$f'(x)$ возрастает $\Leftrightarrow$ $f(x)$ строго выпукла вниз;*
- *$f'(x)$ убывает $\Leftrightarrow$ $f(x)$ строго выпукла вверх.*

Здесь все 10 половинок высказываний делаются на всём Δ.

Доказательство. Рассмотрим первый случай. Возьмём тройку точек $x_1 < x < x_2$. Нам нужно неравенство между наклонами хорд:

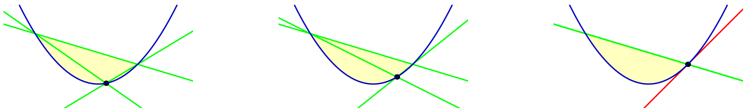
$$(*) \qquad k(x_1,x) \leqslant k(x,x_2).$$

($\Rightarrow$) По теореме Лагранжа найдутся такие точки p_i между x_i и x, что $k(x_i,x) = f'(p_i)$. Производная не убывает, так что $(*)$ выполнено и лемма даёт выпуклость вниз.

(⇐) Тут работает весьма красивый фокус. В неравенстве $(*)$ перейдём к пределам при $x \to x_1 + 0$ и $x \to x_2 - 0$:

$$f'(x_1) \leqslant k(x_1, x_2) \leqslant f'(x_2).$$

Строгие варианты доказываются через нестрогие с помощью дополнительных точек: берём $x_1 < z_1 < x < z_2 < x_2$. □



Функцию называют **дважды дифференцируемой** в точке, и соответственно на множестве, если она имеет там вторую производную. Говорят также о функциях, дифференцируемых трижды или же n раз.

Следствие. *Если функция f дважды дифференцируема на Δ, то*

- $f''(x) \geqslant 0 \Leftrightarrow f(x)$ *выпукла вниз;*
- $f''(x) \leqslant 0 \Leftrightarrow f(x)$ *выпукла вверх;*
- $f''(x) = 0 \Leftrightarrow f(x)$ *линейна;*
- $f''(x) > 0 \Rightarrow f(x)$ *строго выпукла вниз;*
- $f''(x) < 0 \Rightarrow f(x)$ *строго выпукла вверх.*

Здесь опять все 8 половинок высказываний делаются на всём Δ.

Точки перегиба. Подобно тому, как промежутки монотонности разделяются обычно точками экстремума, промежутки, где функция сохраняет направление выпуклости, разделяются обычно точками перегиба. Точку p называют точкой **перегиба** функции f, если $f'(p)$ существует и с разных сторон вблизи p направления выпуклости f разные. В точке перегиба график функции переходит на другую сторону от касательной. Простейшая функция с точкой перегиба — это $f(x) = x^3$.

рисунок?

Лемма. *Для всякой дважды дифференцируемой функции f:*

(1) *точка перегиба f является точкой локального экстремума f';*
(2) *точка строгого локального экстремума f' является точкой перегиба f.*

Доказательство. Упражнение. □

Следствие (Условия перегиба)**.** *Для всякой функции f дважды дифференцируемой в точке $x = p$:*

(1) *если p — точка перегиба, то $f''(p) = 0$;*

(2) *если* $f''(p) = 0$ *и существует* $f'''(p) \neq 0$, *то* p *— точка перегиба.*

Доказательство. (1) По лемме, $f'(x)$ имеет локальный экстремум в точке $x = p$. Далее по теореме Ферма $(f')'(p) = 0$.

(2) По лемме, указанные условия достаточны для строгого локального экстремума функции $f'(x)$ в точке $x = p$. □

Упражнение. *Выпишите в деталях условия перегиба по аналогии с достаточными условиями локального экстремума.*

Упражнение. *Сопоставьте знание полиномов Тейлора и исследование точек перегиба.*

Неравенство Йенсена. Из сохранения функциями направления выпуклости на промежутках можно вывести ряд классических и широко применяемых неравенств.

Лемма. *Выпуклость внизу функции* f *на промежутке* Δ *равносильна выполнению для любых пар точек* $x_i \in \Delta$ *неравенства*

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

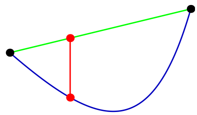
при всех $\alpha_i \geqslant 0$ *с условием* $\alpha_1 + \alpha_2 = 1$.

В этом условии легко исключить один из числовых параметров, однако приведённую симметричную формулировку проще обобщить. Такими параметрами параметризуются точки отрезка $[x_1, x_2]$.

Доказательство. Поместим в точки $(x_i, f(x_i))$ на плоскости массы α_i. Центр масс этой системы находится в точке

$$(x_0, y_0) = (\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 f(x_1) + \alpha_2 f(x_2)),$$

лежащей на хорде над точкой $(x_0, f(x_0))$ графика. □

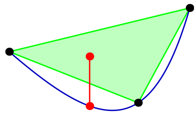


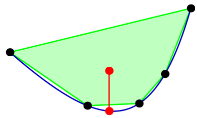
Теорема. *Для всякой функции* f *выпуклой внизу на промежутке* Δ *неравенство*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leqslant \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$$

верно для любых точек $x_i \in \Delta$ *и чисел* $\alpha_i \geqslant 0$ *с условием* $\sum \alpha_i = 1$.

Теорема сформулирована так, как её обычно применяют. Однако выполнение неравенства Йенсена, наоборот, влечёт выпуклость: мы ведь можем выбрать **веса** α_i так, чтобы ненулевыми были всего два, и попадём тогда в ситуацию, уже изученную в лемме.





Доказательство. Поместим в точки $(x_i, f(x_i))$ на плоскости массы α_i. Центр масс этой системы находится в точке

$$(x_0, y_0) = (\alpha_1 x_1 + \dots + \alpha_n x_n, \alpha_1 f(x_1) + \dots + \alpha_n f(x_n)),$$

которая лежит в выпуклом n-угольнике с вершинами в наших материальных точках. Он расположен между графиком функции и хордой, стягивающей две крайние точки, что и даёт искомое неравенство $f(x_0) \leqslant y_0$. Случай всего двух точек сделан в лемме. Если дальше всё ещё не понятно, рассмотрите случай трёх точек. □

Упражнение. *Проведите строгое математическое рассуждение по индукции на количество точек.*

Пример. Функция $\ln x$ строго выпукла вверх. Поэтому неравенство Йенсена для неё обращено в противоположную сторону:

$$\ln\left(\sum \alpha_k x_k\right) \geqslant \sum \alpha_k \ln x_k.$$

Поскольку функция e^x возрастает, отсюда получим, потенцируя,

$$\sum \alpha_k x_k \geqslant \prod x_k^{\alpha_k}.$$

Для n точек, взяв все $\alpha_k = 1/n$, получаем неравенство между средним арифметическим и средним геометрическим положительных чисел:

$$\tfrac{1}{n} \sum x_k \geqslant \left(\prod x_k\right)^{1/n}.$$

Упражнение. *Докажите, что как здесь, так и в неравенстве Йенсена для любой строго выпуклой функции, равенство достигается только при совпадении всех x_k.*

Упражнение. *Выведите неравенство Юнга:*

$$\frac{1}{p} + \frac{1}{q} = 1, \text{ где } p, q > 0, \quad \Rightarrow \quad ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} \text{ для всех } a, b \geqslant 0.$$

4.4. Доказательства формулы Тейлора

лекция 11
08.10.15

Формула Тейлора. Эта общая формула — одна из вершин элементарного анализа. В разных формах ей владели до Тейлора сами основатели предмета: и Ньютон, и Лейбниц в переписке с Бернулли. Способы её получения также были разными, а позже оформились различные доказательства. Даже сама формулировка сейчас даётся в нескольких формах. Познакомимся с тремя путями к этой вершине.

Вопроса о приближении функции полиномами Тейлора мы уже касались. Поскольку коэффициенты такого полинома выражаются через производные соответствующих порядков от исходной функции, его наличие требует дифференцируемости функции достаточное количество раз. На практике большинство функций дифференцируемы сколько угодно раз, так что проблемы с этим не возникает и особенно заботиться об этом условии не сто́ит. Полезно тем не менее отметить, что производная $f^{(k+1)}(a)$ в точке $x = a$ может существовать, лишь если производная $f^{(k)}(x)$ существует в некоторой её окрестности.

Теорема. *Если функция f дифференцируема n раз в точке $x = a$, то вблизи неё*

$$f(x) = \sum_{0 \leqslant k \leqslant n} \tfrac{1}{k!} f^{(k)}(a)(x-a)^k + o\big((x-a)^n\big).$$

Эта сумма без $o\big((x-a)^n\big)$ и есть полином Тейлора степени n функции f вблизи точки $x = a$. Обозначим его через $P_n(x)$. Отметим, что $P_n^{(k)}(a) = f^{(k)}(a)$ при $0 \leqslant k \leqslant n$, а производной нулевого порядка удобно считать саму функцию.

Во всех доказательствах этого раздела считаем для определённости, что $x > a$; случай $x < a$ изучается аналогично.

Интегрирование по частям. Для начала запишем

$$f(x) = f(a) + \int_a^x 1 \cdot f'(t)\,dt$$

и проинтегрируем по частям, подбирая части особо хитро: $u(t) = f'(t)$ и $v(t) = -(x-t)$. Внутри интеграла x является постоянной. Внеинтегральное слагаемое формулы интегрирования по частям превращается в слагаемое формулы Тейлора с первой производной: Bernoulli 169 D'Alembert 1

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

Воодушевившись сиим, второй раз проинтегрируем по частям, хитро беря теперь $u(t) = f''(t)$ и $v(t) = -(x-t)^2/2$. Получим

$$f(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2!} f''(a)(x-a)^2 + \tfrac{1}{2!} \int_a^x (x-t)^2 f'''(t)\,dt.$$

Внеинтегральные слагаемые вместе дают квадратичное приближение Тейлора. Повторяя эти действия, мы на каждом шаге будем выделять следующее слагаемое искомой формулы и после n шагов придём к

$$f(x) = P_n(x) + \tfrac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)\,dt.$$

Cauchy 1821

Подынтегральное выражение усложнилось, но точно вычислять интеграл мы и не собираемся, а лишь оценим его. Для этого заменим его на простенький, заведомо имеющий большее (абсолютное) значение, и увидим, насколько оно мало вблизи $x = a$. А именно, заменим функцию $f^{(n+1)}(t)$ на её максимальное по модулю значение на отрезке $[a, x]$ и вынесем эту постоянную из-под интеграла. В нём тогда останется

$$\int_a^x (x-t)^n\,dt = \tfrac{1}{n+1}(x-a)^{n+1}.$$

Невозможность уменьшения абсолютного значения интеграла при таком его упрощении обоснуем позже интегральной теоремой о взвешенном среднем. Оценка остатка в итоге получится такая же, как ниже в третьем доказательстве, а это точнее, чем в формулировке теоремы.

Теорема о среднем как шаг индукции. Обозначим через $r_n(x)$ остаток $f(x) - P_n(x)$ доказываемой формулы. В точке $x = a$ эта функция зануляется вместе со своими производными вплоть до $r_n^{(n)}(a)$. Требуемое её свойство, равносильное формуле Тейлора, полезно оформить отдельно. Строя его доказательство, мы значительно отступим в сторону от основной темы раздела.

Лемма. *Если в точке $x = a$ функция $r(x)$ имеет $k \geqslant 1$ производных и все они равны нулю, то $r(x) = o\big((x-a)^k\big)$.*

Утверждение зависит от натурального параметра k. Тем самым, в одной этой лемме собрана целая цепочка: при $k = 1$, при $k = 2$, и так далее. Часто в таких случаях прибегают к специальному способу доказательства, называемому **индукцией**. В нашем курсе он ещё не применялся, поэтому мы постараемся придти к нему, анализируя подробное пошаговое рассуждение. Знакомые с методом математической индукции могут ~~пока поспать~~ пропустить подготовительные подробности и сразу перейти к обработанной, индуктивной версии.

Сырое доказательство. Начнём двигаться по цепочке со случая $k = 1$. По определению производной видим, что

$$\frac{r(x)}{x-a} = \frac{r(x) - r(a)}{x-a} \to r'(a) = 0.$$

Это означает, что $r(x) = o(x-a)$. Случай $k = 1$ теперь доказан.

Перейдём к случаю $k = 2$. По формуле конечных приращений

$$r(x) - r(a) = r'(p) \cdot (x-a),$$

где $x > p > a$. Поскольку $r'(a) = 0$, функция $r'(x)$ удовлетворяет условиям случая $k = 1$ леммы, который уже доказан. Значит,

$$r'(p) = o(p-a) = o(x-a)$$

и следовательно,

$$r(x) = (x-a) \cdot o(x-a) = o\big((x-a)^2\big).$$

Случай $k = 2$ теперь доказан.

Перейдём к случаю $k = 3$. Формула Лагранжа не зависит от случая, а функция $r'(x)$ тут удовлетворяет условиям случая $k = 2$ леммы, который уже доказан. Разбор предыдущего случая можно повторить здесь дословно, подправляя только порядки бесконечно малых:

$$r'(x) = o\big((x-a)^2\big), \quad r(x) = o\big((x-a)^3\big).$$

Случай $k = 3$ теперь доказан.

Таким образом, пошагово мы можем добраться до любого необходимого значения k. Шаг от достигнутого значения к следующему выразим так: если $r(a) = 0$ и $r'(x) = o\big((x-a)^{k-1}\big)$, то $r(x) = o\big((x-a)^k\big)$ благодаря формуле Лагранжа. □

Как часть общего подхода проб и ошибок, такой пошаговый метод важен для отыскания фактов и доказательств. Однако затем новые, сырые доказательства такого типа перерабатывают и представляют в более эффективном индуктивном виде. Иногда это касается и решения задачек. Доказательство по индукции состоит из двух частей.

- **База индукции** — это проверка утверждения при наименьшем значении параметра, делающем утверждение верным. Обычно она проста, но верность не всегда наступает при наименьшем значении, делающем утверждение осмысленным.

- **Шаг индукции** — это проверка утверждения для произвольного значения параметра в предположении, что для всех меньших значений оно уже установлено.

Реже встречаются другие формы оформления индукции. Заслуживает упоминания нисходящая индукция: в ней утверждение со значением параметра k сводят к нему же, но со значением параметра $k-1$. По сути работает та же логика, но с меньшими словесными украшениями.

Приведём теперь переработанное доказательство леммы.

Доказательство. Проведём индукцию по k.

База индукции проверяется непосредственно по определениям производной и относительно бесконечно малой.

Шаг индукции опирается на формулу конечных приращений:

$$r(x) - r(a) = r'(p) \cdot (x-a).$$

Действительно, предположим, что лемма верна для функций с $k-1$ нулевой производной в точке $x=a$. Если $r(x)$ имеет k нулевых производных в точке $x=a$, то у $r'(x)$ их никак не меньше $k-1$. Отсюда по предположению индукции $r'(p) = o\big((x-a)^{k-1}\big)$. Подставляя в формулу Лагранжа и упрощая, получим $r(x) = o\big((x-a)^k\big)$. □

В данном случае можно избавиться от упоминания индукции, заменив её на применение формулы конечных приращений n раз. Так устроено, с некоторыми дополнительными изменениями, наше третье доказательство формулы Тейлора в следующем подразделе.

Попутно любопытно отметить подход Лагранжа к теории функций. Функции, происходящие из алгебраических действий, он выделил в качестве примитивных, а по-русски первообразных. Возможность разложения по степеням он (ошибочно) принимал за очевидную, а коэффициенты разложения, помимо факториалов, назвал производными функциями. Этим объясняется происхождение двух терминов.

Lagrange 1772

Кроме того, Лагранж впервые выделил формулу конечных приращений именно как частный случай формулы Тейлора первого порядка. Исходно его первым заинтересовала задача оценки остатка в формуле Тейлора. Как раз его ответ мы и получим ниже.

Доказательство Коши. Проведём рассуждение, требующее от $f(x)$ бо́льшей дифференцируемости, но чуть больше и дающее.

Cauchy 1821

Для вспомогательных функций

$$\varphi(x) = f(x) - P_n(x), \quad \psi(x) = (x-a)^{n+1}$$

проверим условия теоремы Коши: $\varphi(a) = \psi(a) = 0$ и $\psi(x) \neq 0$ при $x \neq a$. Таким же условиям удовлетворяют их производные $\varphi^{(k)}$ и $\psi^{(k)}$ вплоть до $k = n$. Можно применить теорему Коши $n+1$ раз, потребовав существования производной $f^{(n+1)}$ на интервале (a, x). На каждом шаге она выдаёт точку между a и x, причём точки приближаются к a:

$$\frac{\varphi(x)}{\psi(x)} = \frac{\varphi'(x_1)}{\psi'(x_1)} = \frac{\varphi''(x_2)}{\psi''(x_2)} = \cdots = \frac{\varphi^{(n)}(x_n)}{\psi^{(n)}(x_n)} = \frac{\varphi^{(n+1)}(p)}{\psi^{(n+1)}(p)}$$

для подходящих точек $x > x_1 > x_2 \cdots > x_n > p > a$. Забудем теперь промежуточные равенства, подставим определения φ и ψ в крайние, затем чуть перегруппируем и получим

$$f(x) - P_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(p) \cdot (x-a)^{n+1}.$$

Тут даже более точный вид остатка, чем требовалось.

Оформим полученный в доказательстве Коши бонус: оценку остатка в форме Лагранжа. Это лишь одна из многочисленных форм выражения остатка.

Следствие. *Если функция f дифференцируема $n+1$ раз в окрестности U точки $x = a$, то на этой окрестности верна оценка*

$$|f(x) - P_n(x)| \leqslant M \cdot |x-a|^{n+1},$$

где $M = \frac{1}{(n+1)!} \max\{|f^{(n+1)}(u)|$ для $u \in U\}$.

Для функции, дифференцируемой неограниченное количество раз, полиномы $P_n(x)$ представляют собой частичные суммы бесконечного степенного ряда — ряда Тейлора. Возникает проблема сходимости. Хотя при повышении степени приближение лучше, сопутствующая окрестность может становиться меньше. Вопрос об области сходимости таких рядов решается позже, но в этом курсе.

4.5. Раскрытие неопределённостей

В этом разделе мы изучим методы вычисления пределов функций. Когда функция задана сложным выражением и требуется найти её предел в конечной точке $x = a$, следует сперва определить, может ли функция быть там разрывной. Если разрыва нет, то предел находится простой подстановкой $x = a$. При разрыве из-за деления на нуль часто можно заключить, что искомый предел бесконечен или не существует.

Примеры. При $x \to 0$ предел функции $1/x$ не существует, а предел функции $1/x^2$ существует и равен $+\infty$.

Неопределённостями называют случаи, когда подстановка не даёт ответа. Они бывают нескольких типов; чаще всего встречается тип 0/0. Здесь нужно попытаться преобразовать выражение, чтобы устранить неопределённость.

Пример. Найдём предел гиперболического тангенса при $x \to +\infty$. Тут неопределённость ∞/∞, но мы избавляемся от неё простым делением числителя и знаменателя на e^x:

$$\operatorname{th} x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}}.$$

Теперь в пределе просто зануляются экспоненты, так что $\operatorname{th} x \to 1$ при $x \to +\infty$.

Пример. Найдём предел функции $\frac{\operatorname{tg} 2x}{\operatorname{tg} x}$ при $x \to 0$. Тут неопределённость 0/0. Преобразуем тангенс двойного угла и упрощаем дробь:

$$\frac{\operatorname{tg} 2x}{\operatorname{tg} x} = \frac{2 \operatorname{tg} x}{1 - \operatorname{tg}^2 x} \cdot \frac{1}{\operatorname{tg} x} = \frac{2}{1 - \operatorname{tg}^2 x} \to 2.$$

Вспоминая эквивалентность $\operatorname{tg} x \sim x$ или разложение $\operatorname{tg} x = x + o(x)$, мы найдём ответ быстрее.

Для более сложных случаев дифференциальное исчисление предоставляет два основных метода.

Применение формулы Тейлора. Возьмём $f(x) = x^2$. Уравнение $f(x) = 0$ имеет двойной корень в точке $x = 0$; при этом зануляется не только $f(x)$, но и $f'(x)$, тогда как $f''(x) = 2$ не зануляется.

Аналогично, при $x = 0$ у функции $f(x) = x^n$ зануляются все производные порядка ниже n, включая саму функцию как производную нулевого порядка, поэтому говорят о корне кратности n уравнения $f(x) = 0$ или нуле кратности n функции $f(x)$.

По той же логике говорят, что функция $f(x)$, дифференцируемая достаточное количество раз вблизи точки $x = a$, имеет в этой точке нуль **кратности** n, если $f^{(k)}(a) = 0$ при $0 \leqslant k < n$, но $f^{(n)}(a) \neq 0$. В таком случае её разложение Тейлора начинается с

$$f(x) = \tfrac{1}{n!} f^{(n)}(a)(x-a)^n + o\big((x-a)^n\big).$$

Интересно изучить поведение частного двух таких функций:

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{n!} f^{(n)}(a)(x-a)^n + o\big((x-a)^n\big)}{\frac{1}{m!} g^{(m)}(a)(x-a)^m + o\big((x-a)^m\big)}.$$

В случае $n > 0$ и $m > 0$ дробь $f(x)/g(x)$ при $x \to a$ представляет собой неопределённость $0/0$. Однако из приведённых разложений значение предела этой дроби вполне ясно:

- если $n > m$, то предел равен 0;
- если $n < m$, то предел, возможно односторонний, равен $\pm\infty$;
- если $n = m$, то предел конечен и равен $f^{(n)}(a)/g^{(n)}(a)$.

Пример. Найдём этим методом предел

$$\lim_{x\to 0} \frac{\operatorname{tg} x - x}{x - \sin x} = \lim_{x\to 0} \frac{x + \frac{1}{3}x^3 + o(x^3) - x}{x - x + \frac{1}{3!}x^3 + o(x^3)} = \dots = 2.$$

Не удержав кубические слагаемые, мы бы получили $o(x)/o(x)$, не дающее ответа. В таких случаях нужно брать более точные разложения.

Правила Бернулли — Лопиталя. Так называют внешне совсем иной способ действий для раскрытия неопределённостей $0/0$ и ∞/∞.

Теорема (Бернулли)**.** *Если для хороших функций $f(x) \to 0$ и $g(x) \to 0$ при $x \to a$, включая пределы на бесконечности и односторонние, то* Bernoulli 169

$$\lim_{x\to a} \frac{f(x)}{g(x)} \overset{*}{=} \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

как только предел в правой части существует.

Пример. Можно найти (замечательный) предел $\frac{\sin x}{x}$ при $x \to 0$, поскольку отношение производных равно $\cos x$ и стремится к 1.

Пример. В обратную сторону это правило не работает: $\frac{\sin x}{x} \to 0$ при $x \to +\infty$, но отношение производных не имеет предела.

Пример. Можно найти (замечательный) предел $\frac{\ln(1+x)}{x}$ при $x \to 0$, поскольку отношение производных равно $\frac{1}{1+x}$ и стремится к 1.

Доказательство. Утверждение очевидно, когда функции столь хороши — дифференцируемы в точке $x = a$, — что можно написать

$$\frac{f(x)}{g(x)} = \frac{f(a) + f'(a)(x-a) + o(x-a)}{g(a) + g'(a)(x-a) + o(x-a)} = \frac{f'(a) + o(1)}{g'(a) + o(1)}.$$

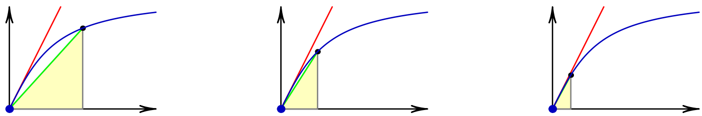
Этот случай покрывает большинство практических задач.

Когда числитель и знаменатель дифференцируемы лишь на некотором интервале (a, b), разложений нет, но работает иной, очень простой подход. Нарисуем плоскую линию, заданную параметрически парой функций $(g(x), f(x))$ вблизи $x = a$, где она попадает в начало координат. Если какая-то из функций не определена при $x = a$, то продолжим её нулевым значением, оставляя непрерывной.

Для каждой хорды линии найдётся параллельная ей касательная; выразим равенство их наклонов как

$$\frac{f'(p)}{g'(p)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)},$$

где $p \in [a, x]$. При $x \to a$ также имеем $p = p(x) \to a$, и если отношение производных имеет предел, то к нему же стремится отношение самих функций.



Присутствие тут дроби со знаменателем $g'(p)$, вкупе со ссылкой на теорему Коши о среднем, формально заставляет требовать, чтобы $g'(x) \neq 0$ вблизи $x = a$, и обычно это входит в формулировку теоремы. Однако на картинке вертикальные хорды и касательные допустимы, так что требование $g'(x) \neq 0$ фактически излишне при наличии предела отношения $f'(x)/g'(x)$. $\square$

Упражнение. *Доказательство не работает при $x \to \infty$. Почему? Сведите этот случай к разобранному.*

лекция 12
15.10.15

Пример. Найдём этим методом предел

$$\lim_{x\to 0} \frac{\operatorname{tg} x - x}{x - \sin x} \overset{*}{=} \lim_{x\to 0} \frac{\operatorname{tg}^2 x}{1 - \cos x} = \cdots = \lim_{x\to 0} \frac{1 + \cos x}{\cos^2 x} \to 2.$$

После применения правила мы тут избавляемся от неопределённости тригонометрическими преобразованиями и упрощением дроби. Можно продолжить иначе, привлекая известные эквивалентности $\operatorname{tg} x \sim x$ и $1 - \cos x \sim x^2/2$.

Упражнение. *Выведите формулу Тейлора из теоремы Бернулли. Указание: составьте подходящую неопределённость и раскройте её.*

Теорема. *Если для хороших функций $f(x) \to +\infty$ и $g(x) \to +\infty$ при $x \to a$, включая пределы на бесконечности и односторонние, то*

$$\lim_{x\to a} \frac{f(x)}{g(x)} \overset{*}{=} \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

как только предел в правой части существует.

Доказательство второго правила несколько сложнее, чем первого, поэтому мы отложим его до следующей главы.

Пример. Сравним при $x \to +\infty$ скорости роста функций $\ln x$ и x^α с любой постоянной $\alpha > 0$. Для этого найдём предел

$$\lim_{x\to+\infty} \frac{\ln x}{x^\alpha} \overset{*}{=} \lim_{x\to+\infty} \frac{x^{-1}}{\alpha x^{\alpha-1}} = 0.$$

Вывод: логарифм растёт медленнее любой степенной функции.

Нередко неопределённость не разрешается после однократного применения правила Бернулли — Лопиталя, и тогда его можно попробовать применить снова. Иногда перед тем полезно совершить над новой дробью чисто алгебраические манипуляции.

Пример. Сравним при $x \to +\infty$ скорости роста функций x^n и e^x. Применив правило дважды, увидим

$$\lim_{x\to+\infty} \frac{x^n}{e^x} \overset{*}{=} n \cdot \lim_{x\to+\infty} \frac{x^{n-1}}{e^x} \overset{*}{=} n(n-1) \cdot \lim_{x\to+\infty} \frac{x^{n-2}}{e^x}.$$

Применяя правило дальше, через n шагов придём к $e^{-x} \to 0$. Вывод: экспонента растёт быстрее любой степенной функции. Ничего по сути не изменится, если вместо e^x взять $e^{\beta x}$ с любой постоянной $\beta > 0$.

К этому заключению также можно придти, заменяя переменную $u = \ln x$ в предыдущем примере.

Прочие неопределённости. Все другие неопределённости сводят к разобранным типам $0/0$ и ∞/∞ простыми преобразованиями, чаще всего с дробями или при помощи логарифма, например:

$$f(x) \cdot g(x) = f(x)/g(x)^{-1}$$

сводит тип $0 \cdot \infty$ к типу $0/0$;

$$f(x) - g(x) = \frac{1}{f(x)^{-1}} - \frac{1}{g(x)^{-1}} = \frac{g(x)^{-1} - f(x)^{-1}}{f(x)^{-1} \cdot g(x)^{-1}}$$

сводит тип $\infty - \infty$ к типу $0/0$;

$$f(x)^{g(x)} = \exp\bigl(g(x) \ln f(x)\bigr)$$

сводит типы ∞^0, 0^0 и 1^∞ к типу $0 \cdot \infty$.

Пример. Раскроем неопределённость типа $0 \cdot \infty$:

$$\lim_{x\to0+0} x \ln x = \lim_{x\to0+0} \frac{\ln x}{x^{-1}} \overset{*}{=} \lim_{x\to0+0} \frac{x^{-1}}{-x^{-2}} = 0.$$

Пример. Раскроем неопределённость типа $\infty - \infty$:

$$\lim_{x\to 0+0}\left(\frac{1}{x} - \operatorname{ctg} x\right) = \lim_{x\to 0+0}\frac{\operatorname{tg} x - x}{x \operatorname{tg} x} \overset{*}{=} \cdots = 0.$$

Для степенных неопределённостей, когда сложная функция переписывается через экспоненциальную, предельный переход нужно внести в показатель экспоненты. Эта операция допустима ввиду непрерывности экспоненциальной функции.

Пример. Раскроем неопределённость типа 0^0:

$$\lim_{x\to 0+0} x^x = \lim_{x\to 0+0}\exp(x\ln x) \overset{!}{=} \exp\Big(\lim_{x\to 0+0} x\ln x\Big) = \exp(0) = 1.$$

Пример. Раскроем неопределённость типа ∞^0:

$$\lim_{x\to +\infty} x^{1/x} = \lim_{x\to +\infty}\exp\Big(\frac{\ln x}{x}\Big) \overset{!}{=} \exp\Big(\lim_{x\to +\infty}\frac{\ln x}{x}\Big) = \exp(0) = 1.$$

Пример. Раскроем неопределённость типа 1^∞:

$$\lim_{x\to 0+0}(1+x)^{1/x} = \lim_{x\to 0+0}\exp\Big(\frac{\ln(1+x)}{x}\Big) \overset{!}{=} \exp\Big(\lim_{x\to 0+0}\frac{\ln(1+x)}{x}\Big) = \exp(1) = e.$$

Предел этой функции при $x \to 0$ слева находится таким же путём и имеет то же значение.

Упражнение. *Предложите разумный способ определить функцию* $f(x) = 0^{0^x}$.

4.6. Интерполяционный полином

пропустим

Задача интерполяции. Приближённое восстановление неизвестной функции $y = f(x)$ по заданному набору её значений в некоторых точках $x_0 < x_1 < \ldots < x_n$ называют **интерполяцией**, если восстановление проводится на отрезке $[x_0, x_n]$ или меньшем, и **экстраполяцией**, если оно выходит за эти пределы.

Когда речь идёт об экспериментальных данных, часто можно что-то заранее предполагать об общей форме теоретического закона, выражаемого искомой функцией, и подгонять параметры под данные. Здесь мы рассмотрим иную задачу. При отсутствии явных предположений, или из других соображений, например, вычислительной эффективности, бывает необходимым найти полином $P(x)$, значения которого в указанных точках совпадают с экспериментальными: $P(x_i) = y_i$.

Имеет ли решение эта задача, и единственно ли оно? Ответ зависит от степени полинома. Полином $P_m(x)$ степени m содержит $m+1$ коэффициентов, которые должны удовлетворять системе из $n+1$ линейных уравнений $P(x_i) = y_i$. Из линейной алгебры известно, что при $m < n$ следует ожидать отсутствия решений, а при $m > n$ — их множественности. Практически интересен случай $m = n$.

Теорема. *Для любых значений $y_0, \ldots, y_n$ решение задачи интерполяции полиномом степени n существует и единственно.*

Это решение называют **интерполяционным полиномом**.

Доказательство. Утверждение следует из формулы Крамера; нужно лишь убедиться, что определитель рассматриваемой системы уравнений отличен от нуля. Он зависит только от x_i, а именно, это определитель Вандермонда

$$V(x_0, \ldots, x_n) = \begin{vmatrix} x_0^n & x_0^{n-1} & \ldots & x_0^2 & x_0 & 1 \\ x_1^n & x_1^{n-1} & \ldots & x_1^2 & x_1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_n^n & x_n^{n-1} & \ldots & x_n^2 & x_n & 1 \end{vmatrix}.$$

В линейной алгебре доказывается, что условие $V \neq 0$ равносильно тому, что все аргументы $x_0, \ldots, x_n$ различны. □

Интерполяционный полином записывают в разных формах; мы рассмотрим две наиболее важных. Явный же вид коэффициентов полинома $P_n(x)$, получаемый из формулы Крамера, не применяется ввиду трудоёмкости раскрытия больших определителей. Раскрыв их заранее в общем виде, в итоге получим интерполяционный полином в форме Лагранжа, найти который проще другим путём.

Решение Лагранжа. Возьмём для простоты $n = 2$ и начнём с частного случая, когда одно из значений y_i равно 1, а остальные равны 0; например, $y_0 = 1$ для определённости. По теореме Безу, решение $P_2(x)$ делится нацело на $x - x_1$ и на $x - x_2$. Поэтому $P_2(x) = a(x - x_1)(x - x_2)$, где коэффициент a находится из условия $P_2(x_0) = 1$. Значит,

$$P_2(x) = \frac{x - x_1}{x_0 - x_1} \cdot \frac{x - x_2}{x_0 - x_2}.$$

Для произвольных y_0, y_1, y_2 отсюда получим решение

$$P_2(x) = y_0 \cdot \frac{x - x_1}{x_0 - x_1} \cdot \frac{x - x_2}{x_0 - x_2} + y_1 \cdot \frac{x - x_0}{x_1 - x_0} \cdot \frac{x - x_2}{x_1 - x_2} + y_2 \cdot \frac{x - x_0}{x_2 - x_0} \cdot \frac{x - x_1}{x_0 - x_1}.$$

Обобщение на произвольное количество контрольных точек,

$$P_n(x) = \sum_{0 \leqslant i \leqslant n} \Bigl(y_i \prod_{j \neq i} \frac{x - x_j}{x_i - x_j} \Bigr),$$

называют интерполяционным полиномом в форме Лагранжа.

Решение Ньютона. Эта форма интерполяционного полинома гораздо старше и применима при весьма специальном выборе точек.

Познакомимся сперва с таблицей разностей Ньютона. Исходные значения записываем в столбик. Рядом во второй столбик записываем их разности. В каждый следующий столбик записываем разности величин из предыдущего столбика. Для четырёх значений y_0, y_1, y_2, y_3 получим таблицу разностей

$$\begin{array}{llll}
y_0 & & & \\
 & y_1 - y_0 & & \\
y_1 & & y_2 - 2y_1 + y_0 & \\
 & y_2 - y_1 & & y_3 - 3y_2 + 3y_1 - y_0. \\
y_2 & & y_3 - 2y_2 + y_1 & \\
 & y_3 - y_2 & & \\
y_3 & & &
\end{array}$$

Упражнение. *Определите последний элемент таблицы разностей, начинающейся со значений* $y_0, \ldots, y_n$.

Теперь перепишем разности как $\Delta y_i = y_{i+1} - y_i$, вторые разности как $\Delta^2 y_i = \Delta(\Delta y_i) = \Delta y_{i+1} - \Delta y_i$, и далее по правилу $\Delta^k y_i = \Delta(\Delta^{k-1} y_i)$. Будем также писать $(1+\Delta)y_i$ вместо $y_i + \Delta y_i = y_{i+1}$. Тогда вся таблица разностей выражается через y_0 и становится

$$\begin{array}{llll}
y_0 & & & \\
 & \Delta y_0 & & \\
(1+\Delta)y_0 & & \Delta^2 y_0 & \\
 & \Delta(1+\Delta)y_0 & & \Delta^3 y_0. \\
(1+\Delta)^2 y_0 & & \Delta^2(1+\Delta)y_0 & \\
 & \Delta(1+\Delta)^2 y_0 & & \\
(1+\Delta)^3 y_0 & & &
\end{array}$$

Для произвольного числа контрольных точек получим

$$y_n = (1+\Delta)^n y_0 = \sum_{0 \leqslant k \leqslant n} \binom{n}{k} \Delta^k y_0.$$

Перейдём, наконец, к специальному случаю Ньютона: разобьём отрезок $[x_0, x_n]$ на n равных частей точками $x_m = x_0 + m\Delta x$. Заменяя целочисленный параметр m на вещественный параметр t, представим произвольную точку x вблизи точки x_0, хотя не обязательно внутри отрезка, как $x(t) = x_0 + t\Delta x$. Смысл этой буквы Δ другой, нежели в таблице разностей.

Ньютон определил обобщённые биномиальные коэффициенты $\binom{t}{k}$ и с их помощью ввёл

$$y(x(t)) = \sum_{0 \leqslant k \leqslant n} \binom{t}{k} \Delta^k y_0 = \sum_{0 \leqslant k \leqslant n} \frac{1}{k!} t(t-1)\cdots(t-k+1)\, \Delta^k y_0.$$

Подставляя сюда

$$t = \frac{x - x_0}{\Delta x},$$

получим интерполяционный полином в форме Ньютона:

$$y(x) = \sum_{0 \leqslant k \leqslant n} \frac{1}{k!}(x-x_0)(x-x_1)\cdots(x-x_{k-1}) \frac{\Delta^k y_0}{(\Delta x)^k}.$$

Предельный переход к формуле Тейлора. Предположим, что приближаемая функция $y = f(x)$ непрерывно дифференцируема n раз на отрезке $[x_0, x]$. В формуле Ньютона перейдём к пределу при $\Delta x \to 0$, сохраняя постоянными значения x_0, x и n. Можно показать, что разностные отношения стремятся к производным,

$$\left(\frac{\Delta}{\Delta x}\right)^k y(x_0) \to \left(\frac{d}{dx}\right)^k y(x_0),$$

а формула Ньютона переходит в формулу Тейлора:

$$f(x) \approx \sum_{0 \leqslant k \leqslant n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k.$$

Включив в рассмотрение остаток, можно вывести и строгое равенство.

Именно через формулу Ньютона пришёл к своей формуле Тейлор, но сразу в виде бесконечного ряда, очень смело и легкомысленно. Впрочем, он не был первым.

столкновени

Содержание