.

КУРС МАТЕМАТИЧЕСКОГО АНАЛИЗА

Семестр 1

для экономического факультета СПбГУ

А.Е.Барабанов

Санкт–Петербург
2002

Элементы теории множеств

Множество — это основное математическое понятие, обозначающее собрание, набор, совокупность, коллекцию каких–либо объектов, вещей. Эти объекты называются *элементами* множества, и говорят, что они множеству принадлежат.

Важнейшее и единственное ограничение в понятии множества состоит в том, что состав множества не может быть расплывчатым. А именно, для любого объекта a и любого множества M возможны только две ситуации: объект a принадлежит множеству M (что обозначается $a \in M$) или объект a не принадлежит множеству M (что обозначается $a \notin M$).

Наиболее употребительны два способа задания множества: явное указание элементов или через свойства элементов. Если некоторый набор объектов a_1, a_2, ..., a_n объединяется в множество, то такое множество обозначается фигурными скобками:

$$A = \{\, a_1, a_2, \ldots, a_n \,\}.$$

Предположим, что P — некоторое свойство объектов. Если это свойство выполнено для объекта x, то пишем $P(x)$. Тогда запись

$$A = \{\, x \,:\, P(x) \,\} \qquad \text{или} \qquad A = \{\, x \,|\, P(x) \,\}$$

обозначает множество A, состоящее из *всех* объектов x, которые обладают свойством P.

В математических утверждениях полезным бывает множество, которое не содержит ни одного элемента. Оно называется *пустым множеством* и обозначается $\emptyset$. Например, множество всех вещественных корней функции $f(x) = x^2 + 1$ является пустым множеством.

Пусть A и B — два множества. Множество B называется *подмножеством* множества A, если любой элемент множества B принадлежит множеству A. Обозначение: $B \subset A$.

Из определения следует, что любое множество A обладает следующими свойствами: $A \subset A$ и $\emptyset \subset A$.

Подмножество A — это множество, состоящее только из элементов множества A, но, быть может, не из всех, а только из тех, которые обладают некоторым дополнительным свойством P, т.е. $\{x \in A \,:\, P(x)\}$.

Два множества называются равными, если каждый элемент первого множества принадлежит второму и наоборот. Другими словами, $A = B$ означает, что $A \subset B$ и $B \subset A$. Например,

$$\{x \,:\, (x^2 - 3)(x^2 - 1) = 0\} = \{1, -1, \sqrt{3}, -\sqrt{3}\}.$$

Высказывания и логическая символика.

Высказыванием называется предложение, о котором можно сказать, что оно может быть либо истинным, либо ложным. Вещь сама по себе не является высказыванием, а утверждение о том, что эта вещь обладает конкретным свойством, есть высказывание.

Высказывания можно соединять логическими операциями для образования новых высказываний. Пусть α и β — некоторые высказывания.

Утверждение о том, что верно и α, и β, называется конъюнкцией этих высказываний и обозначается $\alpha \& \beta$ или $\alpha \wedge \beta$.

Утверждение о том, что верно α или β (возможно, верны сразу оба утверждения), называется дизъюнкцией этих высказываний и обозначается $\alpha \vee \beta$.

Утверждение "если верно α, то верно и β"называется импликацией и обозначается $\alpha \Rightarrow \beta$.

Два утверждения называются равносильными, если они могут быть выполнены или не выполнены только одновременно. Утверждение о том, что высказывания α и β равносильны, называется эквивалентностью и обозначается $\alpha \Leftrightarrow \beta$. Легко доказать, что утверждение $\alpha \Leftrightarrow \beta$ равносильно утверждению $(\alpha \Rightarrow \beta)\&(\beta \Rightarrow \alpha)$.

Утверждение о том, что высказывание α неверно, называется отрицанием α и обычно обозначается $\neg\alpha$ или $\bar{\alpha}$.

Пусть P есть некоторое свойство объектов (например, быть положительным числом). Для каждого конкретного объекта a выражение $P(a)$ есть высказывание, т.е. может быть верным (в примере: если a — число и $a > 0$) или неверным. Утверждение о том, что *существует* объект x, обладающий свойством P, есть новое высказывание. Оно обозначается $\exists x : P(x)$ и читается: "существует x такой, что $P(x)$". Символ $\exists$ называется *квантором существования.*

Утверждение о том, что *любой* объект x обладает свойством P, также является высказыванием. Оно обозначается $\forall x\ P(x)$ и читается: "для любого x верно $P(x)$". Символ $\forall$ называется *квантором всеобщности* и читается "для любого".

Утверждение о том, что существует и только один объект x, обладающий свойством P, обозначается $\exists!\, x : P(x)$. Читается: "существует единственный x такой, что $P(x)$".

Операции над множествами.

Пусть A и B — два множества. *Объединением* множеств A и B называется множество, состоящее из всех объектов, принадлежащих A или B:

$$A \cup B = \{x \mid x \in A \vee x \in B\}.$$

Пересечением множеств A и B называется множество, состоящее из всех объектов, принадлежащих как A, так и B:

$$A \cap B = \{x \mid x \in A \ \&\ x \in B\}.$$

Разностью множеств A и B называется множество, состоящее из всех элементов A, которые не принадлежат B:

$$A \backslash B = \{x \mid x \in A \ \&\ x \notin B\}.$$

Прямым произведением множеств A и B называется множество всех упорядоченных пар (a, b), состоящих из элементов этих множеств:

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Пример: пусть $A = B = \mathsf{R}$ — вещественная прямая. Тогда $\mathsf{R} \times \mathsf{R} = \mathsf{R}^2$ — плоскость.

Упражнение. Пусть A и B — произвольные множества. Доказать следующие равенства множеств:

$A \backslash B = A \backslash (A \cap B)$,

$A \cup A = A$,

$A \backslash A = \emptyset$,

$A \cap A = A$,

$(A \backslash B) \cap (B \backslash A) = \emptyset$.

Числовые множества.

Некоторые множества, элементами которых являются числа, имеют стандартное обозначение. N — множество всех натуральных чисел, оно содержит только целые положительные числа. Иногда в множество N включается 0. R — множество всех вещественных чисел. Z — множество всех целых чисел (положительных, отрицательных и нуля). Q — множество всех рациональных чисел, т.е. чисел вида $\frac{m}{n}$, где m и n — целые и $n \neq 0$. $\mathcal{I}$ — множество всех иррациональных чисел, т.е. вещественных чисел, не являющихся рациональными, $\mathcal{I} = \mathsf{R} \backslash \mathsf{Q}$. Из определений множеств непосредственно следует, что $\mathsf{N} \subset \mathsf{Z} \subset \mathsf{Q} \subset \mathsf{R}$.

В множестве вещественных чисел R определены операции сложения $+$, умножения $\cdot$, а также отношение порядка $<$. Отношение $\leq$ означает "$<$"или -". Эти операции обладают следующими свойствами, называемыми также *аксиомами вещественных чисел*. Все математические утверждения о вещественных числах могут быть логически выведены только из этих аксиом, без каких–либо ссылок на обычную интуицию. Сами же аксиомы независимы: ни одна из них не может быть выведена из всех остальных. Аксиомы удобно разбивать на группы.

I. *Сложение.*

1. $\forall x, y \in \mathsf{R} \quad x + y = y + x$ — коммутативность сложения.

2. $\forall x, y, z \in \mathsf{R} \quad (x + y) + z = x + (y + z)$ — ассоциативность сложения.

3. $\exists a_0 \in \mathsf{R} : \forall x \in \mathsf{R} \quad x + a_0 = x$ — существование нуля. Этот элемент a_0 множества R называется *нулем* и обозначается 0.

4. $\forall x \in \mathsf{R}\ \exists y \in \mathsf{R} : x + y = 0$ — существование противоположного элемента по отношению к сложению. Этот противоположный элемент y обозначается $-x$.

II. *Умножение.*

1. $\forall x, y \in \mathsf{R} \quad x \cdot y = y \cdot x$ — коммутативность умножения.

2. $\forall x, y, z \in \mathsf{R} \quad (x \cdot y) \cdot z = x \cdot (y \cdot z)$ — ассоциативность умножения.

3. $\exists a_1 \in \mathsf{R} : \forall x \in \mathsf{R} \quad x \cdot a_1 = x$ — существование единицы. Этот элемент a_1 множества R называется *единицей* и обозначается 1.

4. $\forall x \in \mathsf{R}\ (x \neq 0 \Rightarrow \exists y \in \mathsf{R} : x \cdot y = 1)$ — существование обратного элемента по отношению к умножению. Этот элемент y обозначается x^{-1} или $\frac{1}{x}$.

5. $\forall x, y, z \in \mathsf{R} \quad (x + y) \cdot z = (x \cdot z) + (y \cdot z)$ — дистрибутивность умножения относительно сложения.

Вещественные числа, которые могут быть получены сложением элемента 1 с самим собой конечное число раз, называются натуральными. N — множество таких чисел.

Множество целых чисел и множество рациональных чисел определяются через множество N:

$$\begin{aligned}
\mathsf{Z} &= \{x \mid \exists n \in \mathsf{N} : (x = n \vee x = (-n))\}, \\
\mathsf{Q} &= \{x \mid \exists m, n \in \mathsf{Z} : x = m \cdot n^{-1}\}.
\end{aligned}$$

III. *Аксиома индукции.*

Если подмножество $X \subset \mathsf{N}$ обладает двумя свойствами:

1) $1 \in \mathsf{X}$,

2) $\forall n \;\; n \in X \Rightarrow n + 1 \in X$,

то $X = \mathsf{N}$.

Первое свойство называется *базой индукции*, а второе —*индукционным переходом* или *индукционным шагом.*

IV. *Неравенства.*

1. $\forall x, y \in \mathsf{R}$ или $x < y$, или $y < x$, или $x = y$.
2. $\forall x, y, z \in \mathsf{R}$ $((x < y \;\&\; y < z) \Rightarrow x < z)$ — транзитивность неравенства.
3. $\forall x, y, z \in \mathsf{R}$ $(x < y \Rightarrow x+z < y+z)$ — согласованность сложения и неравенства.
4. $\forall x, y, z \in \mathsf{R}$ $((x < y \;\&\; z > 0) \Rightarrow x \cdot z < y \cdot z)$ — согласованность умножения и неравенства.
5. $\forall x \in \mathsf{R}\, \exists n \in N : x < n$ — аксиома Архимеда.

Абсолютной величиной вещественного числа x называется само это число x, если $x \geq 0$, или число $-x$, если $x < 0$.

Последняя аксиома — о плотности вещественных чисел — формулируется в терминах свойств промежутков.

Промежутки.

Пусть a и b — вещественные числа и $a \leq b$. Множество $[a, b] = \{x \in \mathsf{R} \,:\, a \leq x \leq b\}$ называется *замкнутым промежутком* или *отрезком* с концами a и b.

Множество $(a, b) = \{x \in \mathsf{R} \,:\, a < x < b\}$ называется *открытым промежутком* или *интервалом* с концами a и b.

Множества $(a, b] = \{x \in \mathsf{R} \,:\, a < x \leq b\}$ и $[a, b) = \{x \in \mathsf{R} \,:\, a \leq x < b\}$ называются *полуоткрытыми промежутками* с концами a и b.

Множества $(a, +\infty) = \{x \in \mathsf{R} \,:\, a < x\}$, $[a, +\infty) = \{x \in \mathsf{R} \,:\, a \leq x\}$, $(-\infty, a) = \{x \in \mathsf{R} \,:\, x < a\}$ и $(-\infty, a] = \{x \in \mathsf{R} \,:\, x \leq a\}$ называются *полубесконечными промежутками.*

Все перечисленные множества называются также *промежутками.* Длиной промежутка с концами a и b называется число $b - a$.

V. *Аксиома плотности вещественных чисел.*

Пусть $[a_n, b_n]$, $n = 1, 2, \ldots$ — последовательность отрезков, и пусть каждый следующий отрезок вложен в предыдущий: $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ для любого $n \in \mathsf{N}$. Тогда существует вещественное число, принадлежащее всем этим промежуткам:

$$\exists c \in \mathsf{R} \,:\, \forall n \in \mathsf{N} \quad c \in [a_n, b_n].$$

Определение *граней и ограниченных множеств.*

Пусть X — подмножество R и $a \in \mathsf{R}$. Число a называется *верхней гранью* множества X, если $\forall x \in X\; x \leq a$. Соответственно, число a называется *нижней гранью* множества X, если $\forall x \in X\; a \leq x$.

Множество X называется *ограниченным сверху*, если у него существует верхняя грань.

Множество X называется *ограниченным снизу*, если у него существует нижняя грань.

Множество X называется *ограниченным*, если оно ограничено сверху и ограничено снизу. $\square$

Теорема *о точных верхних и нижних гранях.*

Пусть непустое множество $X \subset \mathsf{R}$ ограничено сверху. Тогда среди всех его верхних граней существует наименьшая.

Аналогично, среди всех нижних граней непустого ограниченного снизу множества существует наибольшая нижняя грань.

Наименьшая верхняя грань множества X называется *точной верхней гранью* множества X и обозначается $\sup X$. Наибольшая нижняя грань множества X называется *точной нижней гранью* множества X и обозначается $\inf X$.

Если $X = \emptyset$ — пустое множество, то по определению, $\sup \emptyset = -\infty$, $\inf \emptyset = +\infty$.

Доказательство. Пусть множество X ограничено сверху. Это значит, что существует такое число $b_1 \in \mathsf{R}$, что $x \leq b_1$ при всех $x \in X$. Множество X непустое, выберем в нем произвольный элемент $a_1 \in X$. Получился отрезок $[a_1, b_1]$, у которого верхний конец b_1 есть верхняя грань X, а нижний конец принадлежит X. Далее будем строить последовательность вложенных отрезков $[a_n, b_n]$ рекуррентно.

Пусть $n \in \mathsf{N}$ и отрезок $[a_n, b_n]$ обладает свойствами: b_n есть верхняя грань X и $[a_n, b_n] \cap X \neq \emptyset$. Построим отрезок $[a_{n+1}, b_{n+1}]$, обладающий такими же свойствами.

Разделим отрезок $[a_n, b_n]$ пополам точкой c_n. Если c_n есть верхняя грань X, то определим $[a_{n+1}, b_{n+1}] = [a_n, c_n]$, а если нет — то $[a_{n+1}, b_{n+1}] = [c_n, b_n]$. Очевидно, что в обоих случаях b_{n+1} есть верхняя грань X.

Кроме того, в первом случае $[a_{n+1}, b_{n+1}] \cap X = [a_n, c_n] \cap X$, так как на промежутке $(c_n, b_n]$ нет чисел из множества X по определению верхней грани c_n. Во втором случае на промежутке $[c_n, b_n]$ есть некоторые точки множества X, так как число c_n не есть верхняя грань для X. Таким образом, в обоих случаях $[a_{n+1}, b_{n+1}] \cap X \neq \emptyset$.

Рассмотрим множество $M \subset \mathsf{N}$ индексов n, для которых построен отрезок $[a_n, b_n]$, обладающий двумя свойствами: b_n есть верхняя грань X и $[a_n, c_n] \cap X \neq \emptyset$. Множество M удовлетворяет аксиоме индукции. Поэтому $M = \mathsf{N}$.

Таким образом, построена последовательность вложенных отрезков $[a_n, b_n]$, $n = 1, 2, \ldots$. По аксиоме плотности вещественных чисел она имеет общую точку, которую обозначим c. Докажем, что это число c есть наименьшая верхняя грань X.

Предположим, от противного, что существует такое число $x \in X$, что $x > c$. Применим аксиому Архимеда к вещественному положительному числу $z = (b_1 - a_1)/(x - c)$. Существует такое натуральное число m, что $z < m$. Но тогда и $z < 2^m$, так как $m < 2^m$, что легко доказать по индукции. Число m удовлетворяет неравенству $(b_1 - a_1)/2^m < x - c$.

Длины отрезков $[a_n, b_n]$ уменьшаются в 2 раза при увеличении n на 1. Поэтому $b_m - a_m = (b_1 - a_1)/2^m$ и

$$b_m - c \leq b_m - a_m = (b_1 - a_1)/2^m < x - c.$$

Из полученного неравенства $b_m - c < x - c$ следует $b_m < x$, что невозможно, так как b_m — верхняя грань для X и $x \in X$. Противоречие доказывает, что число c является верхней гранью X.

Предположим, от противного, что существует другая верхняя грань d множества X и $d < c$. По аксиоме Архимеда существует такое число m, что $(b_1 - a_1)/2^m < c - d$. Промежуток $[a_m, b_m]$ содержит как число c, так и некоторое число $x \in X$, так как $[a_m, b_m] \cap X \neq \emptyset$. Поэтому

$$c - x \leq b_m - a_m = (b_1 - a_1)/2^m < c - d.$$

Из неравенства $c - x < c - d$ следует, что $d < x$, что невозможно для верхней грани d множества X. Противоречие доказывает, что не существует верхней грани множества X, которая была бы меньше, чем c. Таким образом, число c есть наименьшая верхняя грань множества X.

Существование наибольшей нижней грани ограниченного снизу непустого множества X доказывается аналогично. $\square$

Примененный в доказательстве теоремы метод построения вложенных отрезков половинной длины, обладающих заданным набором свойств, называется *дихотомией.*

Функции и отображения

Пусть X и Y — два множества. Пусть каждому элементу x множества X поставлен в соответствие ровно один элемент множества Y. Тогда говорят, что определено отображение из X в Y. Обозначим это отображение f. Тот элемент, в который преобразуется элемент x, обозначается $f(x)$ и называется образом элемента x при отображении f или значением отображения f на элементе x.

Таким образом, отображение задается тремя объектами: множеством аргументов X, множеством значений Y и правилом, по которому каждый элемент $x \in X$ преобразуется в ровно один элемент $y \in Y$. Эта тройка записывается следующим образом: $f : X \to Y$. Синонимами отображения являются термины: функция, оператор, преобразование.

Определение *основных понятий, связанных с отображениями.*

В каждом пункте будем предполагать, что задано некоторое отображение $f : X \to Y$.

1. Сужение отображения. Пусть X_0 —подмножество X. Тогда можно определить отображение $g : X_0 \to Y$ правилом $g(x) = f(x)$, если $x \in X_0$. Отображение g называется сужением f на X_0 и обозначается $f|_{X_0}$.

2. Образ множества. Пусть $X_0 \subset X$. Тогда множество значений отображения f на элементах X_0 называется образом множества X_0 при отображении f и обозначается $f(X_0)$:

$$f(X_0) = \{y \in Y \mid \exists x \in X : y = f(x)\}.$$

Пример: $f : \mathsf{R} \to \mathsf{R}$, $f(x) = x^2$, $X_0 = [-2, 3)$. Тогда $f(X_0) = [0, 9)$.

3. Прообраз множества. Пусть $Y_0 \subset Y$. Тогда множество всех элементов X, которые преобразуются в элементы множества Y_0, называется прообразом Y_0 при отображении f и обозначается $f^{-1}(Y_0)$:

$$f^{-1}(Y_0) = \{x \in X \mid f(x) \in Y_0\}.$$

Пример: $f : \mathsf{R} \to \mathsf{R}$, $f(x) = x^2$, $Y_0 = [-20, 1)$. Тогда $f^{-1}(Y_0) = (-1, 1)$.

4. Отображение на множество. Отображение f множества X в множество Y называется отображением *на* множество Y, если $f(X) = Y$. Другими словами, при отображении на Y уравнение $y = f(x)$ при любом $y \in Y$ разрешимо относительно хотя бы одного элемента $x \in X$.

Пример: отображение $y = x^2$, $x \in \mathsf{R}$ не есть отображение на Y, если $Y = \mathsf{R}$, однако оно будет отображением на Y, если $Y = [0, \infty)$.

5. Взаимно однозначное отображение. Отображение f называется взаимно однозначным, если в каждый элемент $y \in Y$ отображается один или ни одного элемента $x \in X$.

6. Биекция (или биективное отображение или взаимно однозначное соответствие). Если f взаимно однозначно и является отображением на Y, то оно называется биекцией. В этом случае каждому элементу множества X соответствует один и только один элемент множества Y, и наоборот.

С понятием биекции связано понятие равномощности множеств. Два множества A и B называются *равномощными*, если существует биективное отображение A на B.

Множество C называется *счетным*, если оно равномощно множеству натуральных чисел N. Множество D называется *континуальным* (или мощность D есть *континуум*), если оно равномощно множеству вещественных чисел R. Множества N и R не равномощны. Любой непустой интервал (a, b) имеет мощность континуум.

7. Обратное отображение. Пусть f — биекция. Тогда для каждого элемента $y \in Y$ существует и ровно один элемент $x \in X$, для которого $y = f(x)$. Поэтому можно определить отображение $g : Y \to X$, которое преобразует $y \in Y$ в решение x уравнения $y = f(x)$. Отображение g называется обратным по отношению к f и обозначается $g = f^{-1}$.

8. Композиция отображений (или сложное отображение или суперпозиция отображений). Пусть $f : X \to Y$, $Y \subset Y_1$ и $g : Y_1 \to Z$. Тогда можно определить отображений $h : X \to Z$ правилом $h(x) = g(f(x))$. Отображение h называется композицией f и g. Оно обозначается $h = g \circ f$.

Пределы

Предел числовой последовательности

Определение *последовательности.* Последовательностью называется функция натурального аргумента. □

Пусть a — последовательность. Это значит, что $a : \mathsf{N} \to Y$, где Y — множество значений. По определению отображения, для каждого натурального числа n задано число $y = a(n)$. По установившейся традиции это значение обозначается не $a(n)$, а a_n, и называется членом последовательности. Сама последовательность a обозначается не как отображение, а в виде совокупности: $a = (a_n)_{n=1}^{\infty}$.

В дальнейшем будем рассматривать только последовательности чисел, т.е. $Y = \mathsf{R}$, т.е. a_n — число при каждом n.

Определение *ограниченной последовательности.* Последовательность $a = (a_n)_{n=1}^{\infty}$ называется *ограниченной сверху* (соответственно, *снизу*), если существует такое число M, что $a_n \le M$ (соответственно, $a_n \ge M$) при всех $n \in \mathsf{N}$.

Последовательность называется *ограниченной*, если она ограничена снизу и сверху. □

Определение *подпоследовательности.* Пусть $a = (a_n)_{n=1}^{\infty}$ — последовательность вещественных чисел, а $k = (k_i)_{i=1}^{\infty}$ — последовательность натуральных чисел и $k_{i+1} > k_i$ при всех натуральных i. Определим для каждого натурального i число $b_i = a_{k_i}$. Тогда последовательность $b = (b_i)_{i=1}^{\infty}$ называется подпоследовательностью последовательности a с индексами k. Она обозначается $b = (a_{k_i})_{i=1}^{\infty}$. □

Определение *предела последовательности.* Число A называется *пределом* последовательности $a = (a_n)_{n=1}^{\infty}$, если

$$\forall \varepsilon > 0 \ \exists N \in \mathsf{N} \ : \ \forall n > N \qquad |a_n - A| < \varepsilon.$$

При выполнении этого условия предел последовательности обозначается

$$A = \lim_{n\to\infty} a_n,$$

а также $a_n \to A$ при $n \to \infty$. □

Последовательность называется *бесконечно малой*, если она имеет предел, равный нулю.

Определение *бесконечного предела последовательности.* Последовательность $a = (a_n)_{n=1}^{\infty}$ имеет *бесконечный предел* (или предел этой последовательности равен ∞), если

$$\forall M > 0 \ \exists N \in \mathsf{N} \ : \ \forall n > N \qquad |a_n| > M.$$

Говорят, что предел равен $+\infty$, если последнее неравенство заменено на $a_n > M$, и что предел равен $-\infty$, если на $a_n < -M$. □

Последовательность называется бесконечно большой, если она имеет бесконечный предел.

Определение *последовательности, отделенной от нуля.* Последовательность $a = (a_n)_{n=1}^{\infty}$ называется *отделенной от нуля* (или *отделимой от нуля*), если существует такое положительное число $C > 0$, что $|a_n| \ge C$ при всех $n \in \mathsf{N}$.

Теорема *об ограниченности и отделенности от нуля последовательностей, имеющих предел.*

1. Пусть последовательность $(a_n)_{n=1}^{\infty}$ имеет предел. Тогда она ограничена.

2. Пусть последовательность $(a_n)_{n=1}^{\infty}$ имеет бесконечный предел. Тогда она не ограничена.

3. Пусть последовательность $(a_n)_{n=1}^{\infty}$ имеет предел $A \neq 0$ и среди ее членов нет нулей. Тогда она отделена от нуля.

Доказательство.

1. Для $\varepsilon = 1$ существует такое число N, что при всех $n > N$ выполнено неравенство $|a_n - A| < 1$. Из последнего неравенства следует, что $|a_n| < |A| + 1$ при $n > N$. Пусть M — наибольшее из чисел $|a_1|$, $|a_2|$, ..., $|a_N|$, $|A| + 1$. Тогда, очевидно, $|a_n| \leq M$ при всех n.

2. Для любого числа M все числа $|a_n|$ больше M, начиная с некоторого номера N. Поэтому условие ограниченности, очевидно, не выполнено.

3. Для $\varepsilon = |A|/2 > 0$ существует такое число N, что при всех $n > N$ выполнено неравенство $|a_n - A| < |A|/2$. Из последнего неравенства следует, что $|a_n| \geq |A| - |A|/2 = |A|/2$ при $n > N$. Пусть M — наименьшее из чисел $|a_1|$, $|a_2|$, ..., $|a_N|$, $|A|/2$. Оно положительно. Тогда, очевидно, $|a_n| \geq M$ при всех n. $\square$

Теорема *об ограниченности и отделенности от нуля обратных величин.*

Пусть в последовательности $a = (a_n)_{n=1}^{\infty}$ нет нулевых членов. Определим последовательность обратных величин: $b = (1/a_n)_{n=1}^{\infty}$.

1. Если последовательность a ограничена, то последовательность b отделена от нуля.

2. Если a отделена от нуля, то b — ограниченная последовательность.

3. Если $a_n \to 0$, то b имеет бесконечный предел.

4. Если $a_n \to \infty$, то b стремится к нулю.

Доказательство.

1. Существует такое число M, что $|a_n| \leq M$ при всех n. Определим $C = 1/M > 0$. Тогда $1/|a_n| \geq 1/M = C$.

2. Существует такое число $C > 0$, что $|a_n| \geq C$ при всех n. Определим $M = 1/C$. Тогда $1/|a_n| \leq 1/C = M$.

3. Пусть число $M > 0$ — произвольное. Определим $\varepsilon = 1/M$. Тогда существует такой номер N, что $|a_n| < \varepsilon = 1/M$ при всех $n > N$. Это значит, что $1/|a_n| > M$ при тех же n. Поэтому b имеет бесконечный предел.

4. Пусть число $\varepsilon > 0$ — произвольное. Определим $M = 1/\varepsilon$. Тогда существует такой номер N, что $|a_n| > M = 1/\varepsilon$ при всех $n > N$. Это значит, что $1/|a_n| < \varepsilon$ при тех же n. Поэтому число 0 является пределом последовательности b. $\square$

Теорема *о произведении на ограниченную последовательность.*

Рассмотрим две последовательности: $a = (a_n)_{n=1}^{\infty}$ и $b = (b_n)_{n=1}^{\infty}$. Пусть $c = (c_n)_{n=1}^{\infty}$ — последовательность из произведений, $c_n = a_n b_n$ при всех n. Последовательность c будем также обозначать $c = (a_n b_n)_{n=1}^{\infty}$.

1. Пусть $a_n \to 0$ и b ограничена. Тогда $a_n b_n \to 0$.

2. Пусть $a_n \to \infty$ и b отделена от нуля. Тогда $a_n b_n \to \infty$.

Доказательство.

1. Пусть $|b_n| \leq B$ при всех n. Выберем произвольное положительное число $\varepsilon > 0$. Требуется доказать, что существует такое число N, что $|a_n b_n| < \varepsilon$ при $n > N$.

По определению предела из $a_n \to 0$ следует, что для числа $\varepsilon_1 = \varepsilon/B$ существует такое число N, что $|a_n| < \varepsilon_1$ при всех $n > N$. Однако, отсюда следует, что $|a_n b_n| < B\varepsilon_1 = \varepsilon$ при $n > N$.

2. По условию существует такое число $C > 0$, что $|b_n| \geq C$ при всех n. Выберем произвольное положительное число $M > 0$. Требуется доказать, что существует такое число N, что $|a_n b_n| \geq M$ при $n > N$.

По определению бесконечного предела из $a_n \to \infty$ следует, что для числа $M_1 = M/C$ существует такое число N, что $|a_n| \geq M_1$ при всех $n > N$. Однако, отсюда следует, что $|a_n b_n| \geq CM_1 = M$ при $n > N$. □

Теорема *о сумме бесконечно малых и бесконечно больших.*

Рассмотрим две последовательности: $a = (a_n)_{n=1}^{\infty}$ и $b = (b_n)_{n=1}^{\infty}$. Пусть $c = (c_n)_{n=1}^{\infty}$ — последовательность из сумм, $c_n = a_n + b_n$ при всех n. Последовательность c будем также обозначать $c = (a_n + b_n)_{n=1}^{\infty}$.

1. Если $a_n \to 0$ и $b_n \to 0$, то $a_n + b_n \to 0$.

2. Если $a_n \to \infty$ и $b_n \to \infty$ и при каждом n знаки a_n и b_n совпадают, то $a_n + b_n \to \infty$.

Доказательство.

1. Пусть $\varepsilon > 0$. Для числа $\varepsilon_1 = \varepsilon/2$ существует такое число N_a, что $|a_n| < \varepsilon_1$ при всех $n > N_a$. Для этого же числа существует такое число N_b, что $|b_n| < \varepsilon_1$ при всех $n > N_a$. Определим N как наибольшее из чисел N_a и N_b.

Если $n > N$, то $n > N_a$ и $n > N_b$, и поэтому $|a_n| < \varepsilon/2$ и $|b_n| < \varepsilon/2$. Следовательно, $|a_n + b_n| < \varepsilon$.

2. Если знаки a_n и b_n совпадают, то $|a_n + b_n| \geq |a_n|$. Поэтому $a_n + b_n \to \infty$. □

Теорема *о пределе суммы, произведения и частного.*

Пусть последовательности $a = (a_n)_{n=1}^{\infty}$ и $b = (b_n)_{n=1}^{\infty}$ имеют (конечные) пределы, соответственно, A и B.

1. Последовательность $(a_n + b_n)_{n=1}^{\infty}$ имеет предел, равный $A + B$.

2. Последовательность $(a_n b_n)_{n=1}^{\infty}$ имеет предел, равный AB.

3. Если $B \neq 0$ и среди чисел b_n нет нулей, то последовательность $(a_n/b_n)_{n=1}^{\infty}$ имеет предел, равный A/B.

Доказательство.

Рассмотрим последовательности с общими членами $\alpha_n = a_n - A$ и $\beta_n = b_n - B$. Из определения предела следует, что $\alpha_n \to 0$ и $\beta_n \to 0$ при $n \to \infty$.

1. По теореме о сумме бесконечно малых: $\alpha_n + \beta_n \to 0$. Подставим выражение для α_n и β_n: $a_n + b_n - A - B \to 0$. Из определения предела следует, что $a_n + b_n \to A + B$ при $n \to \infty$.

2. В выражении $a_n b_n - AB = A\beta_n + B\alpha_n + \alpha_n\beta_n$ каждое слагаемое стремится к нулю по теоремам о произведении на ограниченную последовательность. По теореме о сумме бесконечно малых эта сумма стремится к нулю.

3. По теореме об отделенности от нуля последовательности, имеющей ненулевой предел, последовательность b_n отделена от нуля. По теореме об ограниченности обратных величин последовательность $1/b_n$ ограничена. Далее,

$$\frac{1}{b_n} - \frac{1}{B} = \frac{1}{b_n}\frac{1}{B}(B - b_n) = -\frac{1}{b_n}\frac{1}{B}\beta_n \to 0$$

по теореме о произведении бесконечно малой на ограниченную. Следовательно, последовательность $1/b_n$ сходится к $1/B$. По только что доказанному утверждению о пределе произведения

$$\frac{a_n}{b_n} = \frac{1}{b_n}a_n \to \frac{1}{B}A$$

при $n \to \infty$. $\square$

Теорема *о пределе сжатой последовательности.*

Пусть последовательности $a = (a_n)_{n=1}^{\infty}$, $b = (b_n)_{n=1}^{\infty}$ и $c = (c_n)_{n=1}^{\infty}$ обладают следующим свойством: при каждом n либо $a_n \leq b_n \leq c_n$, либо $c_n \leq b_n \leq a_n$. Пусть последовательности a и c имеют одинаковый предел A. Тогда последовательность b имеет предел и этот предел равен A.

Доказательство.

Пусть $\varepsilon > 0$. Требуется найти такой номер N, что $|b_n - A| < \varepsilon$ при всех $n > N$. Последовательности a и c имеют предел, поэтому для заданного ε существуют такие числа N_a и N_c, что $|a_n - A| < \varepsilon$ при $n > N_a$, и $|c_n - A| < \varepsilon$ при $n > N_c$. Оба неравенства выполнены, если $n > N_a$ и $n > N_c$, т.е. при $n > \max\{N_a, N_c\}$.

Определим N как максимальное из чисел N_a и N_c. Тогда при $n > N$ одновременно a_n и c_n находятся в промежутке $(A - \varepsilon, A + \varepsilon)$. Но числа b_n находятся между a_n и c_n. Поэтому b_n также находится в этом промежутке, т.е. $|b_n - A| < \varepsilon$. $\square$

Определение *монотонной последовательности.*

Последовательность $a = (a_n)_{n=1}^{\infty}$ называется *возрастающей*, если $a_n \leq a_{n+1}$ при всех натуральных n. Последовательность называется *строго возрастающей*, если $a_n < a_{n+1}$ при всех n.

Аналогично определяется *убывающая* последовательность ($a_n \geq a_{n+1}$) и строго убывающая последовательность ($a_n > a_{n+1}$).

Последовательность называется *монотонной*, если она возрастающая или убывающая.

Теорема *о пределе монотонной последовательности.*

Возрастающая ограниченная сверху последовательность имеет (конечный) предел. Возрастающая неограниченная сверху последовательность стремится к $+\infty$.

Аналогично, убывающая последовательность либо ограничена снизу и тогда имеет (конечный) предел, либо неограничена и тогда стремится к $-\infty$.

Доказательство.

Пусть последовательность $a = (a_n)_{n=1}^{\infty}$ возрастающая. Пусть последовательность ограничена, т.е. множество ее значений $S = \{a_n | n \in \mathsf{N}\}$ ограничено. У ограниченного множества существует точная верхняя грань. Обозначим ее A и докажем, что число A есть предел последовательности a.

Пусть $\varepsilon > 0$. Число $A - \varepsilon$ не является верхней гранью множества S, так как оно меньше, чем A. Поэтому существует такое число N, то $a_N > A - \varepsilon$. Однако, при $n > N$ числа a_n могут находиться только в промежутке $(A - \varepsilon, A]$ и, следовательно, $|a_n - A| < \varepsilon$. Это условие совпадает с определением предела.

Пусть последовательность a неограничена. Докажем, что ее предел равен $+\infty$. Пусть $M > 0$. Существует такое число N, что $a_N > M$. Но тогда и $a_n > M$ при всех $n > N$, что совпадает с определением бесконечного предела, равного $+\infty$. $\square$

Предел функции

Одним из центральных понятий математического анализа является окрестность точки. Окрестностью точки a называется любое множество, содержащее открытый интервал, в который входит a. Например, интервал $(a-\delta, a+\delta)$ является окрестностью точки a (при $\delta > 0$). Любое множество, содержащее хотя бы какую-нибудь окрестность точки, является также окрестностью этой точки. Важное свойство окрестностей состоит в том, что пересечение любых двух окрестностей одной и той же точки есть окрестность этой точки.

Проколотой окрестностью точки a называется множество, полученное удалением самой точки a из окрестности точки a. Пересечение конечного числа проколотых окрестностей точки есть также проколотая окрестность этой точки.

В дальнейшем будем предполагать, что функция f задана в некоторой проколотой окрестности точки a.

Определение *предела функции в точке.* Число A называется *пределом функции f в точке a*, если

$$\forall \varepsilon > 0\ \exists \delta > 0\ :\ \text{если } |x-a| < \delta \text{ и } x \neq a, \text{ то } |f(x)-A| < \varepsilon.$$

Обозначение: $f(x) \to A$ при $x \to a$ или

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

Другими словами, число A является пределом функции f в точке a, если для любого числа $\varepsilon > 0$ неравенство $|f(x) - A| < \varepsilon$ выполнено для всех x из некоторой проколотой окрестности точки a.

Определение *бесконечного предела функции в точке.* Функция f имеет *бесконечный предел в точке a*, если

$$\forall M \in \mathsf{R}\ \exists \delta > 0\ :\ \text{если } |x-a| < \delta \text{ и } x \neq a, \text{ то } |f(x)| > M.$$

Обозначение: $f(x) \to \infty$ при $x \to a$ или

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

□

Аналогично определяются бесконечные пределы со знаком, т.е. $\lim_{x\to a} f(x) = +\infty$ (в определении $f(x) > M$) и $\lim_{x\to a} f(x) = -\infty$ (в определении $f(x) < M$).

По определению, окрестностью точки $+\infty$ называется любое множество, содержащее какой-нибудь луч $(c, +\infty)$.

Определение *предела функции в бесконечности.* Пусть функция f определена по крайней мере на некотором луче $(c, +\infty)$. Число A называется *пределом функции f в $+\infty$*, если

$$\forall \varepsilon > 0\ \exists K \in \mathsf{R}\ :\ \text{если } x > K, \text{ то } |f(x)-A| < \varepsilon.$$

Обозначение: $f(x) \to A$ при $x \to +\infty$ или

$$\lim_{x\to +\infty} f(x) = A.$$

□

Аналогично определяются предел при $x \to -\infty$. Функция f должна быть определена по крайней мере на луче $(-\infty, c)$ и в определении предела $x > K$ следует заменить на $x < K$.

Предел $\lim_{x\to\infty} f(x)$ определяется для функции f, заданной по крайней мере на некоторых лучах, направленных в $+\infty$ и в $-\infty$, т.е. на множестве $(-\infty, c) \cup (d, +\infty)$ с некоторыми числами c и d. В определении следует подставить $|x| > K$.

Определение *односторонних пределов.*

Пусть $a \in \mathsf{R}$ и область определения функции f содержит некоторый интервал (a, b) с левым концом в точке a. Число A называется *пределом справа функции f в точке a*, если

$$\forall\varepsilon > 0\ \exists\delta > 0: \text{ если } a < x < a + \delta, \text{ то } |f(x) - A| < \varepsilon.$$

Обозначение: $\lim_{x\to a+0} f(x) = A$. □

Аналогично определяется *предел слева функции f в точке a*. В определении участвуют те значения x, для которых $a - \delta < x < a$. Обозначение: $\lim_{x\to a-0} f(x) = A$.

Теорема *о связи предела и односторонних пределов.*

Для того, чтобы функция f имела предел в точке $a \in \mathsf{R}$, необходимо и достаточно, чтобы она имела оба односторонних предела в этой точке, которые были бы равны.

Доказательство.

1. Необходимость. Пусть функция имеет предел A в точке a. Из определения предела следует, что число A является также пределом слева и пределом справа функции f в точке A.

2. Достаточность. Пусть $\varepsilon > 0$. По определению предела слева существует такое число $\delta_+ > 0$, что $|f(x) - A| < \varepsilon$ при всех x в интервале $a < x < a + \delta_+$. Аналогично, по определению предела слева существует такое число $\delta_- > 0$, что $|f(x) - A| < \varepsilon$ при всех x в интервале $a - \delta_- < x < a$. Определим $\delta = \min\{\delta_+, \delta_-\}$. Очевидно, что если $|x - a| < \delta$ и $x \neq a$, то либо $a < x < a + \delta_+$, либо $a - \delta_- < x < a$. В обоих случаях $|f(x) - A| < \varepsilon$, что означает, что число A есть предел f в точке a. □

Говорят, что некоторое свойство выполняется *вблизи* точки a, если существует такая проколотая окрестность точки a, в которой оно выполнено.

Определение *ограниченности и отделимости от нуля функции вблизи точки.* Функция f называется *ограниченной вблизи точки a*, если существует такая проколотая окрестность $\mathcal{O}$ точки a, что сужение f на $\mathcal{O}$ ограничено (т. е. существует такое число C, что $|f(x)| \leq C$ для всех $x \in \mathcal{O}$).

Функция f называется *отделимой от нуля вблизи точки a*, если существует такая проколотая окрестность $\mathcal{O}$ точки a, что сужение f на $\mathcal{O}$ отделено от нуля (т. е. существует такое число $C > 0$, что $|f(x)| \geq C$ для всех $x \in \mathcal{O}$). □

Теорема *об ограниченности и отделимости от нуля функции, имеющей предел.*

Пусть функция f имеет предел A в точке a.

1. Если A — число, то f ограничена вблизи точки a.
2. Если A — бесконечность (в т. ч. со знаком), то f не ограничена вблизи точки a.
3. Если A — число и $A \neq 0$, то f отделена от нуля вблизи точки a.

Доказательство.

1. Определим $\varepsilon = 1$. Тогда существует такая проколотая окрестность $\mathcal{O}$ точки a, что $|f(x) - A| < 1$ при всех $x \in \mathcal{O}$. Но тогда в этой окрестности $|f(x)| < 1 + |A|$, поэтому на $\mathcal{O}$ функция f ограничена.

2. От противного. Пусть на некоторой проколотой окрестности $\mathcal{O}$ функция f ограничена, т.е. $|f(x)| \leq C$ при некотором $C > 0$. По определению бесконечного предела существует другая проколотая окрестность $\mathcal{O}_1$ точки a, на которой $|f(x)| > C$. Но пересечение $\mathcal{O}_2 = \mathcal{O} \cap \mathcal{O}_1$ непусто. Для любого $x \in \mathcal{O}_2$ одновременно $|f(x)| \leq C$ и $|f(x)| > C$. Противоречие показывает, что предположение об ограниченности функции f неверно.

3. Определим $\varepsilon = |A|/2 > 0$. Тогда существует такая проколотая окрестность $\mathcal{O}$ точки a, что $|f(x) - A| < |A|/2$ при всех $x \in \mathcal{O}$. Но тогда в этой окрестности $|f(x)| > |A|/2$, поэтому на $\mathcal{O}$ функция f отделена от нуля. □

Теорема *о сжатой функции.*

Пусть $a \in \mathsf{R}$ и функции f, g, h заданы по крайней мере в проколотой окрестности точки a. Пусть

1. $f(x) \leq g(x) \leq h(x)$ или $h(x) \leq g(x) \leq f(x)$ для каждой точки x из этой окрестности.

2. Существуют пределы f и h в точке a и они равны A.

Тогда функция g также имеет предел в точке a, и этот предел равен A.

Доказательство.

Пусть $\varepsilon > 0$. Из условий

$$\lim_{x \to a} f(x) = A, \qquad \lim_{x \to a} h(x) = A$$

следует, что существуют такие числа $\delta_f, \delta_h > 0$, что если $0 \neq |x - a| < \delta_f$, то $|f(x) - A| < \varepsilon$, а если $0 \neq |x - a| < \delta_h$, то $|h(x) - A| < \varepsilon$.

Определим $\delta = \min\{\delta_f, \delta_h\}$. Требуется доказать, что если $x \neq a$, $|x - a| < \delta$, то $|g(x) - A| < \varepsilon$.

Если $f(x) \leq g(x) \leq h(x)$, то

$$-\varepsilon < f(x) - A \leq g(x) - A \leq h(x) - A < \varepsilon.$$

Следовательно, $|g(x) - A| < \varepsilon$. Аналогично для случая $h(x) \leq g(x) \leq f(x)$. □

Теорема *о пределе композиции функций.*

Пусть $h = g \circ f$ — композиция отображений f и g. Пусть функция f имеет предел в точке a, $\lim_{x \to a} f(x) = A$, а функция g имеет предел в точке A, $\lim_{y \to A} g(y) = B$. При этом a, A и B могут быть конечными или бесконечными, в т. ч. $+\infty$ и $-\infty$. Если предел A есть число и если f принимает значение A в любой проколотой окрестности точки a, то дополнительно предполагается, что $g(A) = B$.

Тогда функция h имеет предел в точке a, и он равен B. □

Доказательство.

Рассмотрим только случай, когда a, A, B есть числа (не бесконечности). Пусть $\varepsilon_h > 0$ – произвольное число. Существует такое число $\delta_g > 0$, что если $0 \neq |y - A| < \delta_g$, то $|g(y) - B| < \varepsilon_h$. Подставим в определение предела для функции f в точке a вместо ε число δ_g. Тогда существует такое число $\delta_f > 0$, что если $0 \neq |x - a| < \delta_f$, то $|f(x) - A| < \delta_g$. Определим $\delta_h = \delta_f$.

Пусть $0 \neq |x - a| < \delta_h$. Тогда $|f(x) - A| < \delta_g$. Обозначим $y = f(x)$. Если $y \neq A$, то $|g(y) - B| < \varepsilon_h$, т.е. $|h(x) - B| < \varepsilon_h$. Если же $y = A$, то $g(y) = B$ и, следовательно, $|h(x) - B| = 0$. □

Теорема *о секвенциальном пределе функции.*

Пусть $a \in \mathsf{R}$ и функция f задана по крайней мере в проколотой окрестности точки a. Тогда следующие утверждения равносильны:

1. f имеет предел в точке a, равный A.

2. Для любой последовательности $(x_n)_{n=1}^{\infty}$, не содержащей точку a и находящейся в области определения функции f, если $x_n \to a$, то $f(x_n) \to A$.

Доказательство.

1⇒2 следует из определений предела функции и последовательности.

2⇒1. Докажем от противного. Отрицание существования предела, равного A, записывается следующим логическим утверждением:

$$\exists\varepsilon > 0 : \forall\delta > 0\ \exists x \in (a-\delta, a+\delta) \setminus \{a\} : |f(x) - A| \geq \varepsilon.$$

Фиксируем это число ε. Выберем в качестве δ число $1/n$. Соответствующее число x обозначим x_n. Так поступим для всех натуральных n. Получим последовательность $(x_n)_{n=1}^{\infty}$.

По построению, $|x_n - a| < 1/n$. Следовательно, $x_n \to a$. Однако, $|f(x_n) - A| \geq \varepsilon$ при всех n. Это противоречит условию 2. □

Теорема *о пределе суммы, произведения, частного.*

Пусть $a \in \mathsf{R}$ и функции f, g заданы по крайней мере в проколотой окрестности точки a. Пусть f, g имеют пределы в точке a. Тогда $f+g$, fg имеют пределы в точке a и

$$\begin{aligned} \lim_{x\to a} f(x)g(x) &= \lim_{x\to a} f(x) \lim_{x\to a} g(x), \\ \lim_{x\to a}(f(x)+g(x)) &= \lim_{x\to a} f(x) + \lim_{x\to a} g(x). \end{aligned}$$

Если к тому же $\lim_{x\to a} g(x) \neq 0$, то существует предел фукнции f/g в точке a и

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

Доказательство.

Следует из теоремы о секвенциальном пределе и соответствующих теорем о пределе последовательностей. □

Замечание *о многомерных пространствах.*

Пусть $a \in \mathsf{R}^m$, $X \subset \mathsf{R}^m$, задана функция $f : X \to \mathsf{R}^n$ и некоторый проколотый шар $D = \{x \in \mathsf{R}^m : |x-a| < \gamma,\ x \neq a\}$ с центром в точке a содержится в X.

Вектор $A \in \mathsf{R}^n$ называется пределом функции f в точке a, если

$$\forall\varepsilon > 0\ \exists\delta > 0 : \text{ если } 0 < |x-a| < \delta, \text{ то } |f(x) - A| < \varepsilon.$$

□

Для отображений из R^m в R^n справедлива теорема о пределе сложной функции.

Одним из основных приемов вычисления пределов является замена переменной. По существу это применение теоремы о пределе композиции отображений. Пусть требуется вычислить предел

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

Пусть g – некоторая функция, $x = g(y)$, и известно, что $g(y) \to a$ при $y \to b$. Если формально подставить в предел вместо x значение $g(y)$, а вместо $x \to a$ подставить $y \to b$, то новый предел по теореме о пределе композиции будет равен исходному.

Точнее, если предел функции f в точке a существует, то

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

и последний предел существует.

Пример. В следующем разделе будет доказано, что $\sin x/x \to 1$ при $x \to 0$. Пусть требуется вычислить предел функции $\sin(2x-2)/(x^2-1)$ при $x \to 1$. Сделаем замену переменной: $y = 2x - 2$. Тогда $y \to 0$ при $x \to 1$. Поэтому

$$\lim_{x\to 1}\frac{\sin(2x-2)}{x^2-1} = \lim_{x\to 1}\frac{\sin(2x-2)}{2x-2}\cdot\frac{2x-2}{x^2-1} = \lim_{y\to 0}\frac{\sin y}{y}\cdot\lim_{x\to 1}\frac{2}{x+1}.$$

Оба последних предела существуют, и только поэтому предел произведения был заменен на произведение пределов. В итоге получается число 1.

Вычисление пределов. Раскрытие неопределенностей

При вычислении предела

$$\lim_{x\to a}\frac{f(x)}{g(x)}$$

можно отдельно вычислить пределы функций f и g в точке a и разделить один на другой, но только в том случае, когда выполнены два условия: 1) каждый из двух пределов существует и 2) предел знаменателя не равен нулю.

Если хотя бы одно из этих условий не выполнено, то предел отношения, тем не менее, может существовать, а правило его вычисления называется раскрытием неопределенности.

Например, при вычислении предела

$$\lim_{x\to 0}\frac{x^2}{x+x^2}$$

нельзя отдельно подставить $x = 0$ в числитель и знаменатель, так как предел знаменателя окажется равным нулю. Оба предела равны нулю, и поэтому такой предел называется неопределенностью типа $\frac{0}{0}$.

Для раскрытия этой неопределенности достаточно сократить дробь на x, и затем уже вычислить отношение предела числителя к пределу знаменателя. Получается число $0/1 = 0$.

Отношение тех же функций при $x \to \infty$ также является неопределенностью, которая имеет тип $\frac{\infty}{\infty}$. Однако, ее можно раскрыть, и предел равен 1.

1. Предел отношения многочленов при $x \to \infty$.

Пусть $P(x)$ и $Q(x)$ — два многочлена степеней n и m:

$$\begin{aligned} P(x) &= p_0 + p_1x + \ldots + p_nx^n, \qquad p_n \neq 0,\\ Q(x) &= q_0 + q_1x + \ldots + q_mx^m, \qquad q_m \neq 0. \end{aligned}$$

Докажем, что

$$\lim_{x\to\infty}\frac{P(x)}{Q(x)} = \begin{cases} 0, & \text{если } n < m,\\ \infty, & \text{если } n > m,\\ \frac{p_n}{q_m}, & \text{если } n = m. \end{cases}$$

Пусть $n < m$. Разделим числитель и знаменатель на x^n. Получится

$$\lim_{x\to\infty}\frac{P(x)}{Q(x)} = \lim_{x\to\infty}\frac{P(x)}{Q(x)} = \frac{\frac{p_0}{x^n} + \frac{p_1}{x^{n-1}} + \ldots + p_n}{\frac{q_0}{x^n} + \frac{p_1}{x^{n-1}} + \ldots + q_mx^{m-n}}.$$

Числитель имеет предел, равный p_n, а знаменатель стремится к бесконечности, так как содержит слагаемое q_mx^{m-n} и $m - n > 0$. Поэтому предел дроби равен нулю.

Аналогично доказывается, что при $n > m$ после деления на x^m знаменатель имеет предел, равный q_m, а числитель стремится к бесконечности из–за слагаемого $p_n x^{n-m}$.

При $n = m$ после деления на x^n предел числителя равен p_n, а предел знаменателя равен q_m, поэтому предел дроби равен p_n/q_m.

2. Предел отношения многочленов при $x \to a$.

Пусть a — некоторое число и $P(x)$, $Q(x)$ — два многочлена степеней n и m:

$$\begin{aligned} P(x) &= p_0 + p_1 x + \ldots + p_n x^n, \qquad p_n \neq 0, \\ Q(x) &= q_0 + q_1 x + \ldots + q_m x^m, \qquad q_m \neq 0, \end{aligned}$$

Покажем, что предел их отношения в точке a существует (конечный или бесконечный) и легко вычисляется. Оба многочлена имеют конечные пределы при $x \to a$, равные $P(a)$ и $Q(a)$, соответственно.

Если $Q(a) \neq 0$, то по теореме о пределе отношения функций

$$\lim_{x\to a} \frac{P(x)}{Q(x)} = \frac{P(a)}{Q(a)}$$

Пусть $Q(a) = 0$. Если при этом $P(a) \neq 0$, то функция $P(x)$ отделена от нуля вблизи точки a и $1/Q(x)$ — бесконечно большая в точке a. Следовательно, $P(x)/Q(x)$ — бесконечно большая в точке a и предел равен бесконечности.

Наиболее важным является случай $P(a) = 0$ и $Q(a) = 0$. Получается неопределенность типа $\frac{0}{0}$. Она легко сводится к одному из разобранных случаев следующим способом.

По теореме Безу если многочлен имеет корень в точке $x = a$, то он делится нацело на $(x - a)$. Поэтому $P(x) = (x - a)P_1(x)$, $Q(x) = (x - a)Q_1(x)$, где $P_1(x)$ и $Q_1(x)$ – многочлены. На $(x - a)$ можно сократить, и тогда

$$\lim_{x\to a} \frac{P(x)}{Q(x)} = \lim_{x\to a} \frac{P_1(x)}{Q_1(x)},$$

причем степени $P_1(x)$ и $Q_1(x)$ уменьшились на 1 по сравнению со степенями $P(x)$ и $Q(x)$. Получился снова предел отношения многочленов. Если $P_1(a) = 0$ и $Q_1(a) = 0$, то эти многочлены снова можно разложить на множители и сократить $(x - a)$. При каждом сокращении степени полиномов в числителе и в знаменателе уменьшаются на 1. Поэтому таких сокращений может быть только конечное число, и, наконец, значение в точке $x = a$ числителя или знаменателя будет ненулевым. А тогда предел вычисляется так, как указано для предыдущих случаев.

3. Первый замечательный предел.

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Пусть $0 < x < \pi/2$. Точку на единичной окружности, соответствующую длине дуги x, обозначим M. Пусть O — начало координат и A — начало единичной окружности, т.е. точка 1 на оси абсцисс. Вертикальная линия $x = 1$ является осью тангенсов. Обозначим точку пересечения прямой OM с осью тангенсов через B. Тогда длина отрезка AB равна $\operatorname{tg} x$.

Треугольник OAM является частью сектора OAM, который в свою очередь является частью треугольника OAB. В треугольнике OAM основание равно 1, а высота —

$\sin x$, поэтому площадь есть $(\sin x)/2$. В секторе OAB радиус равен 1, а угол раствора — x. Поэтому площадь сектора есть $x/2$. Наконец, площадь треугольника OAB есть $(\operatorname{tg} x)/2$. Напишем неравенство площадей этих фигур:

$$\sin x < x < \operatorname{tg} x.$$

Разделим эти неравенства на положительное число $\sin x$ и перейдем к обратным величинам:

$$1 > \frac{\sin x}{x} > \cos x.$$

При $x \to 0$ левое и правое выражения в этом неравенстве стремятся к 1. По теореме о сжатой функции

$$\lim_{x\to 0+0} \frac{\sin x}{x} = 1.$$

Если же $-\pi/2 < x < 0$, то определим $y = -x$. Рассмотрим сложное отображение h:

$$x \to y \to \frac{\sin y}{y},$$

так что

$$h(x) = \frac{\sin y}{y} = \frac{\sin(-x)}{-x} = \frac{\sin x}{x}, \qquad -\frac{\pi}{2} < x < 0.$$

Поскольку $y > 0$, то $\sin y/y \to 1$ при $y \to 0+0$. По теореме о пределе композиции функций существует предел функции h при $x \to 0-0$, равный композиции пределов. А именно, при $x \to 0-0$: $y \to 0+0$ и, следовательно, $h(x) \to 1$, т.е.

$$\lim_{x\to 0-0} \frac{\sin x}{x} = 1.$$

Таким образом, оба односторонних предела, как слева, так и справа, равны 1. Поэтому предел при $x \to 0$ существует и равен 1, что завершает доказательство.

4. Второй замечательный предел.

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Рассмотрим последовательность $a_n = (1 + 1/n)^n$. Возведение в степень n равносильно $n-1$ умножению, при выполнении которых можно раскрывать скобки. Получающаяся сумма известна как бином Ньютона. В общем виде результат записывается формулой

$$(x+y)^n = x^n + C_n^1 x y^{n-1} + C_n^2 x^2 y^{n-2} + \cdots + C_n^{n-1} x^{n-1} y + y^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k},$$

в которой числа C_n^k называются числами сочетаний из n по k и равны

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{1\cdot 2\cdot \ldots \cdot k}, \qquad k = 0, 1, \ldots, n.$$

Эти числа сочетаний обладают свойством симметрии: $C_n^k = C_n^{n-k}$. Все они целые. Полезно помнить первые значения: $C_n^0 = 1$, $C_n^1 = n$, $C_n^2 = n(n-1)/2$.

В данном случае при $x = 1$, $y = 1/n$ получится

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= 1 + n\frac{1}{n} + \frac{n(n-1)}{1\cdot 2}\frac{1}{n^2} + \frac{n(n-1)(n-3)}{1\cdot 2\cdot 3}\frac{1}{n^3} + \cdots \\ &\quad + \frac{n(n-1)(n-3)\cdots 1}{1\cdot 2\cdot 3\cdot \ldots \cdot n}\frac{1}{n^n} \\ &= 2 + \frac{1-\frac{1}{n}}{1\cdot 2} + \frac{(1-\frac{1}{n})(1-\frac{1}{n-1})}{1\cdot 2\cdot 3} + \ldots + \frac{(1-\frac{1}{n})(1-\frac{1}{n-1})\cdots(1-\frac{n-1}{n})}{1\cdot 2\cdot \ldots \cdot n}.\end{aligned}$$

Каждое слагаемое в последней сумме монотонно зависит от n: оно увеличивается при увеличении n. Кроме того, при увеличении n на 1 добавляется одно новое положительное слагаемое. По этой причине

$$a_{n+1} > a_n \qquad \forall n \in \mathsf{N}.$$

Это значит, что последовательность a_n возрастает. Докажем, что она ограничена.

Увеличим каждое слагаемое в последней сумме, заменив в числителе каждый множитель $1 - k/n$ на 1, а в знаменателе каждый множитель k, кроме 1, на 2:

$$a_n < 2 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} = 2 + \frac{\frac{1}{2}-\frac{1}{2^n}}{1-\frac{1}{2}} < 2 + \frac{\frac{1}{2}}{1-\frac{1}{2}} = 3.$$

Следовательно, возрастающая последовательность $(a_n)_{n=1}^{\infty}$ ограничена сверху. По теореме о пределе монотонной последовательности она имеет конечный предел.

Предел последовательности $(1 + 1/n)^n$ обозначается буквой e в честь Л.Эйлера. Число e играет большую роль в математическом анализе. Логарифмы с основанием e называются натуральными логарифмами и обозначаются ln.

Докажем, что число e является также пределом функции $(1 + 1/x)^x$ при $x \to \infty$. Сначала найдем предел этой функции при $x \to +\infty$, а затем при $x \to -\infty$.

Пусть $x > 0$. Целую часть x обозначим n. Тогда $n \le x < n + 1$ и

$$\left(1+\frac{1}{n+1}\right)^n < \left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{n}\right)^{n+1}.$$

Определим две новые функции f и h равенствами

$$\begin{aligned} f(x) &= \left(1+\frac{1}{n+1}\right)^n \qquad \text{при } n \le x < n+1, \\ h(x) &= \left(1+\frac{1}{n}\right)^{n+1} \qquad \text{при } n \le x < n+1. \end{aligned}$$

Очевидно, что при всех $x \ge 1$: $f(x) \le (1 + 1/x)^x \le h(x)$.

По теореме о пределе произведения и частного находим

$$\begin{aligned} \lim_{x\to+\infty} f(x) &= \lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^n = \lim_{n\to\infty}\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{1+\frac{1}{n+1}} = \frac{e}{1} = e, \\ \lim_{x\to+\infty} h(x) &= \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right) = e\cdot 1 = e. \end{aligned}$$

По теореме о сжатой последовательности функция $g(x) = (1+1/x)^x$ имеет предел при $x \to +\infty$, равный e, т. е.

$$\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e.$$

Вычислим этот же предел при $x \to -\infty$. Пусть $x < 0$. Обозначим $y = -x$ и проведем тождественные преобразования:

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

Если $x \to -\infty$, то $y \to +\infty$. Если $y \to +\infty$, то

$$\left(1+\frac{1}{y-1}\right)^{y-1}\left(1+\frac{1}{y-1}\right) \to e \cdot 1 = e.$$

По теореме о пределе композиции функций отсюда следует, что

$$\lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^{x} = e.$$

Пределы при $x \to +\infty$ и при $x \to -\infty$ равны e, поэтому предел при $x \to \infty$ существует и равен e.

Второй замечательный предел также записывают в виде

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e.$$

Порядок бесконечно малых

Бесконечно малой в точке a называется функция, для которой $\lim_{x\to a} f(x) = 0$. При этом предполагается, что f задана по крайней мере в некоторой проколотой окрестности точки a. В дальнейшем будем предполагать, что a фиксировано (точка или бесконечность). Предполагается также, что если a – точка, то все бесконечно малые заданы в соответствующих проколотых окрестностях, а если a – бесконечность (возможно, со знаком), то функции заданы по крайней мере на соответствующих полуосях ($(c, +\infty)$ и/или $(-\infty, d)$).

Кроме того, будем считать, что все бесконечно малые из данного раздела отличны от нуля в некоторой проколотой окрестности предельной точки. Если это конечная точка a, то $f(x) \neq 0$ при $x \in (a-\alpha, a+\alpha) \setminus \{a\}$. Отметим, что в самой точке a функция f может обращаться в ноль.

Определение *порядка бесконечно малых.*

Пусть f и g – две бесконечно малые.

1. Говорят, что f есть бесконечно малая более высокого порядка, чем g, если

$$\lim_{x\to a}\frac{f(x)}{g(x)} = 0.$$

Обозначение: $f = o(g)$ при $x \to a$.

2. f имеет порядок не выше, чем g, если функция $f(x)/g(x)$ ограничена вблизи a. Обозначение: $f = O(g)$ при $x \to a$.

3. f и g называются бесконечно малыми одного порядка, если $f = O(g)$ и $g = O(f)$. Равносильное свойство: функция $f(x)/g(x)$ ограничена и отделима от нуля вблизи a.

4. f и g называются эквивалентными, если

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

Обозначение эквивалентности: $f \sim g$. □

Замечание. Если $f \sim g$, то и $g \sim f$, т.е. $\lim_{x\to a} g(x)/f(x) = 1$. Действительно, функция $\alpha(x) = f(x)/g(x)$ определена вблизи a и имеет предел 1. Следовательно, по теореме о пределе частного, функция $1/\alpha(x) = g(x)/f(x)$ тоже имеет предел 1.

Теорема. *Свойства эквивалентных бесконечно малых.*
Пусть f и g – бесконечно малые. Следующие условия равносильны:
1. $f \sim g$.
2. $f - g$ есть бесконечно малая более высокого порядка, чем g.
Доказательство.
Следует из равносильности условий:

$$\lim_{x\to a} \frac{f(x)}{g(x)} = 1, \qquad \lim_{x\to a} \frac{f(x) - g(x)}{g(x)} = 0.$$

□

Упражнения. *1.* К теореме можно добавить еще одно равносильное условие: $f-g$ есть бесконечно малая более высокого порядка, чем f.
2. Если $f \sim g$ и $g \sim h$, то $f \sim h$ (свойство транзитивности).

Теорема *о замене эквивалентных бесконечно малых в пределе частного.*
Пусть $f \sim f_1$, $g \sim g_1$ и существует конечный или бесконечный предел

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

Тогда существует и равен A предел

$$\lim_{x\to a} \frac{f_1(x)}{g_1(x)} = A.$$

Доказательство.
Определим функции $\phi(x) = f_1(x)/f(x)$, $\psi(x) = g_1(x)/g(x)$ и $h(x) = f(x)/g(x)$. Каждая из этих функций имеет предел при $x \to a$. По теореме о пределе произведения и частного существует и равен A предел

$$A = \lim_{x\to a} h(x) \frac{\phi(x)}{\psi(x)} = \lim_{x\to a} \frac{f_1(x)}{g_1(x)}.$$

□

Бесконечно большие

1. Функция f называется *бесконечно большой при* $x \to a$, если $\lim_{x\to a} = \infty$.
2. Бесконечно большие f и g называются *эквивалентными*, если

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

3. Говорят, что f есть *бесконечно большая более высокого порядка, чем g, при* $x \to a$, если

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

Для бесконечно больших функций справедливы утверждения, аналогичные двум последним теоремам.

Теорема. *Свойства бесконечно больших.*

1. Пусть f и g – бесконечно большие при $x \to a$. Условие $f \sim g$ равносильно условию f имеет более высокий порядок, чем $f - g$.

2. В пределе частного бесконечно большие функции можно заменять на эквивалентные.

Доказательство. (Упражнение.)

Таблица эквивалентных бесконечно малых

При $x \to 0$ следующие бесконечно малые функции эквивалентны:

$$\begin{array}{rcll}
\log_a(1+x) & \sim & x\dfrac{1}{\ln a}, & a > 0, \\
a^x - 1 & \sim & x\ln a, & a > 0, \\
(1+x)^\alpha - 1 & \sim & \alpha x, & \alpha > 0, \\
\sin x & \sim & x, & \\
1 - \cos x & \sim & \dfrac{x^2}{2}, & \\
\operatorname{tg} x & \sim & x, & \\
\arcsin x & \sim & x, & \\
\operatorname{arctg} x & \sim & x. &
\end{array}$$

При доказательстве этих эквивалентностей будет использовано свойство непрерывности функций $\sin x$, $\operatorname{tg} x$ и $\ln x$, которое обсуждается в следующем разделе. Докажем по очереди каждую из представленных эквивалентностей по определению этого свойства.

1. Воспользуемся тождеством $(\log_a y)\ln a = \ln y$.

$$\lim_{x\to 0}\frac{\log_a(1+x)\ln a}{x} = \lim_{x\to 0}\frac{\ln(1+x)}{x} = \lim_{x\to 0}\ln\left((1+x)^{\frac{1}{x}}\right) = \ln e = 1.$$

2. Сделаем замену переменной: $x = \log_a(y+1)$, или $y = a^x - 1$, и воспользуемся предыдущей эквивалентностью:

$$\lim_{x\to 0}\frac{x\ln a}{a^x - 1} = \ln a \lim_{y\to 0}\frac{\log_a(1+y)}{y} = 1.$$

3. Сделаем замену переменной: $y = (1+x)^\alpha - 1$ и воспользуемся первой эквивалентностью при $a = e$:

$$1 = \lim_{y\to 0}\frac{\ln(y+1)}{y} = \lim_{x\to 0}\frac{\ln((1+x)^\alpha - 1 + 1)}{(1+x)^\alpha - 1} = \lim_{x\to 0}\frac{\alpha\ln(1+x)}{(1+x)^\alpha - 1} = \lim_{x\to 0}\frac{\alpha x}{(1+x)^\alpha - 1}.$$

В последнем равенстве бесконечно малая $\ln(1+x)$ была заменена на эквивалентную бесконечно малую функцию x.

4. Это первый замечательный предел.

5. Сделаем замену переменной: $x = 2y$.

$$\lim_{x\to 0}\frac{2(1-\cos x)}{x^2} = \lim_{x\to 0}\frac{4\sin^2\frac{x}{2}}{x^2} = \lim_{y\to 0}\frac{4\sin^2 y}{4y^2} = \left(\lim_{y\to 0}\frac{\sin y}{y}\right)^2 = 1.$$

6. Поскольку $\cos x \to 1$ при $x \to 0$, то

$$\lim_{x\to 0}\frac{\operatorname{tg} x}{x} = \lim_{x\to 0}\frac{\sin x}{x}\frac{1}{\cos x} = 1.$$

7. Сделаем замену переменной: $x = \sin y$. Тогда

$$\lim_{x\to 0}\frac{\arcsin x}{x} = \lim_{y\to 0}\frac{y}{\sin y} = 1.$$

8. Сделаем замену переменной: $x = \operatorname{tg} y$. Тогда

$$\lim_{x\to 0}\frac{\operatorname{arctg} x}{x} = \lim_{y\to 0}\frac{y}{\operatorname{tg} y} = 1.$$

Дополнительные сведения о пределах

Теорема. *Предельный переход в неравенствах.*

1. Пусть последовательности $a = (a_n)_{n=1}^{\infty}$ и $b = (b_n)_{n=1}^{\infty}$ имеют пределы, соответственно, A и B (конечные или бесконечные). Пусть $a_n \le b_n$ при всех $n \in \mathsf{N}$. Тогда $A \le B$.

2. Пусть a — число или $-\infty$. Пусть функции $f(x)$ и $g(x)$ заданы на некотором промежутке (a, b) с левым концом в точке a и имеют правосторонние пределы при $x \to a+0$, равные, соответственно, A и B (конечные или бесконечные). Пусть $f(x) \le g(x)$ при всех $x \in (a, b)$. Тогда $A \le B$.

Доказательство. Оба утверждения доказываются по одинаковой схеме от противного. Докажем второе утверждение для случая, когда a, A и B — числа. Пусть $A > B$. Выберем положительное число $\varepsilon = (A - B)/2$. По определению правостороннего предела существуют такие числа $\delta_f > 0$ и $\delta_g > 0$, что если $a < x < \delta_f$, то $|f(x) - A| < \varepsilon$, а если $a < x < \delta_g$, то $|g(x) - B| < \varepsilon$.

Пусть δ — минимальное из чисел δ_f и δ_g. Если $a < x < \delta$, то выполнены сразу оба неравенства:

$$|f(x) - A| < \varepsilon, \qquad |g(x) - B| < \varepsilon.$$

Но тогда расстояние между числами A и B должно быть меньше 2ε, так как $A > B$ и

$$A = (A - f(x)) + f(x) \le |f(x) - A| + f(x) < \varepsilon + g(x) = \varepsilon + (g(x) - B) + B < 2\varepsilon + B.$$

Таким образом, $|A - B| < 2\varepsilon$, что противоречит выбору $\varepsilon = (A - B)/2$. $\square$

Доказательства первого утверждения, а также второго утверждения при $a = -\infty$ или при бесконечном пределе A или B остаются в качестве упражнения.

Теорема *о единственности предела.*

1. Пусть последовательность $(a_n)_{n=1}^{\infty}$ имеет предел A. Тогда этот предел единственный, т.е. никакое другое число или бесконечность не могут быть пределом той же последовательности.

2. Пусть функция f имеет предел A при $x \to a$. Тогда этот предел единственный.

Доказательство. Докажем только утверждение 1, так как доказательство утверждения 2 полностью аналогично. От противного, пусть A — число и B — другое число, также являющееся пределом последовательности $(a_n)_{n=1}^{\infty}$. Выберем $\varepsilon < |B - A|/2$. Тогда условия $|a_n - A| < \varepsilon$ и $|a_n - B| < \varepsilon$ не могут выполняться одновременно. Но по определению предела они должны выполняться для всех достаточно больших чисел n. $\square$

Непрерывные функции

Определение *непрерывной функции в точке.*

Пусть $a \in \mathsf{R}$, $\mathcal{O}$ — окрестность точки a и функция f определена по крайней мере на множестве $\mathcal{O}$.

Приращением функции f в точке a называется функция $\Delta_a f$, определенная в окрестности нуля равенством

$$\Delta_a f(h) = f(a+h) - f(a).$$

Аргумент h этой функции часто называют *приращением аргумента* функции f.

Функция f называется *непрерывной в точке a*, если существует и равен нулю предел $\Delta_a f$ в нуле:

$$\lim_{h \to 0} \Delta_a f(h) = 0.$$

□

Условие непрерывности функции в точке можно выразить через определение предела на языке ε–δ:

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \quad \text{если } |h| < \delta \text{ и } a+h \in \mathcal{O}, \text{ то } |f(a+h) - f(a)| < \varepsilon.$$

Отметим, что условие $a + h \in \mathcal{O}$ необходимо добавить потому, что функция f должна быть определена в точке $a + h$. Напротив, условие $h \neq 0$, присутствующее в общем определении предела, можно исключить, потому что при $h = 0$ неравенство $|f(a+h) - f(a)| < \varepsilon$ выполнено автоматически.

Если выполнить замену переменной $x = a + h$, то последнее определение записывается следующим образом:

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \quad \text{если } |x - a| < \delta \text{ и } x \in \mathcal{O}, \text{ то } |f(x) - f(a)| < \varepsilon.$$

По определению предела это значит, что

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

Последнее равенство является одной из наиболее популярных форм определения непрерывности функции f в точке a.

Существенным отличием определений предела функции f в точке a и непрерывности функции f в точке a является участие самого значения $f(a)$ в последнем определении. Напомним, что функция f может иметь предел в точке a, даже если она в этой точке не определена, например, при раскрытии неопределенностей.

Если вместо предела при $h \to 0$ подставить односторонний предел при $h \to 0$ справа или при $h \to 0$ слева, то получится свойство функции, названное односторонней непрерывностью слева или, соответственно, справа.

Функция f называется *непрерывной слева в точке a*, если

$$\lim_{h \to 0-0} \Delta_a f(h) = 0.$$

Функция f называется *непрерывной справа в точке a*, если

$$\lim_{h \to 0+0} \Delta_a f(h) = 0.$$

Лемма *о связи непрерывности и односторонней непрерывности.*

Для того, чтобы функция f была непрерывна в точке a, необходимо и достаточно, чтобы она была непрерывна слева в точке a и непрерывна справа в точке a.

Доказательство. Необходимость очевидна, так как функция, имеющая предел $f(a)$ в точке a имеет пределы слева и справа в точке a, равные $f(a)$.

Докажем достаточность. Выберем произвольное число $\varepsilon > 0$. По определению односторонней непрерывности существуют такие числа δ_- и δ_+, что если $-\delta_- < h < 0$, то $|f(a+h) - f(a)| < \varepsilon$, а если $0 < h < \delta_+$, то тоже $|f(a+h) - f(a)| < \varepsilon$.

Определим δ как наименьшее из чисел δ_- и δ_+. Тогда если $|h| < \delta$ и $h \neq 0$, то $|f(a+h) - f(a)| < \varepsilon$. Следовательно, предел функции $\Delta_a f$ в точке $h = 0$ существует и равен нулю. □

Классификация разрывов

Пусть функция f определена на промежутке P и a — внутренняя точка P. Если функция f не является непрерывной в точке a, то a называется *точкой разрыва* функции f. На графике точке разрыва соответствует разрыв линии графика, откуда и появилось называние таких точек.

Рассмотрим два односторонних предела функции f в точке a:

$$L = \lim_{x \to a-0} f(x), \qquad R = \lim_{x \to a+0} f(x).$$

Если хотя бы один из этих пределов не существует или бесконечный, то точка a называется *разрывом второго рода.*

Пример: функция $f(x) = 1/x$ при $x \neq 0$ и $f(0) = 1$.

Если пределы L и R оба существуют, а непрерывности тем не менее нет, то это может быть по двум причинам: либо $L \neq R$, либо $L = R$ и это значение отличается от $f(a)$. В последнем случае разрыв называется *устранимым*: действительно, если переопределить функцию f только в одной точке равенством $f(a) = L (= R)$, то функция станет непрерывной в точке a.

Пример: $f(x) = \frac{\sin x}{x}$ при $x \neq 0$ и $f(0) = 0$. Эту функцию можно сделать непрерывной, если переопределить $f(0) = 1$.

Если $L \neq R$, т.е. левосторонний и правосторонний пределы существуют, конечны и различны, то разрыв называется *скачком*, а число $R - L$ называется *величиной скачка.* Устранимый разрыв или скачок называются *разрывами первого рода.*

Пример скачка — функция знака числа: $f(x) = -1$ при $x < 0$, $f(x) = 1$ при $x > 0$ и $f(0) = 0$. Здесь величина скачка в точке $a = 0$ равна 2.

Теорема *о непрерывности суммы, произведения, частного.*

Пусть функции f и g заданы в окрестности точки a и непрерывны в точке a. Тогда функции $\alpha f(x) + \beta g(x)$, $f(x)g(x)$ непрерывны в точке a (где α, β – числа), а функция $f(x)/g(x)$ непрерывна в точке a, если $g(a) \neq 0$.

Доказательство. Заключение теоремы следует из соответствующих свойств пределов. Например,

$$\lim_{x \to a} f(x)g(x) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x) = f(a)g(a),$$

что означает непрерывность произведения. □

Теорема *о непрерывности композиции непрерывных функций.*

Пусть заданы функции $f : X \to Y$, $g : Y_1 \to Z$ и $Y \subseteq Y_1$. Определим сложную функцию $h = g \circ f$. Пусть X и Y_1 — промежутки и $a \in X$. Обозначим $b = f(a)$. Пусть функция f непрерывна в точке a, а функция g непрерывна в точке b. Тогда функция h непрерывна в точке a.

Доказательство. По теореме о пределе композиции функций сложная функция h имеет предел в точке a, и он равен

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

Поэтому функция h непрерывна в точке a. □

Теорема *об ограниченности функции вблизи точки непрерывности.*

Пусть функция f задана в окрестности $\mathcal{O}$ точки a и непрерывна в точке a. Тогда существует такая окрестность $\mathcal{O}_1$ точки a, что сужение f на $\mathcal{O}_1$ ограничено.

Доказательство. По теореме об ограниченности и отделимости от нуля функции, имеющей предел, существует такая проколотая окрестность $\mathcal{O}_2$ точки a, что f ограничена на $\mathcal{O}_2$. Существует такое число $M_2 > 0$, что $|f(x)| \le M_2$ для любого $x \in \mathcal{O}_2$.

Добавим к проколотой окрестности $\mathcal{O}_2$ точку a, так что множество $\mathcal{O}_1 = \mathcal{O}_2 \cup \{a\}$ является окрестностью точки a. Определим M как максимальное из чисел M_2 и $|f(a)|$. Тогда, очевидно, $|f(x)| \le M$ для любого $x \in \mathcal{O}_1$. □

Функции, непрерывные на промежутке

Определение.

Пусть функция f определена на промежутке Δ. Функция f называется *непрерывной на* Δ, если она непрерывна в каждой точке этого промежутка. □

Непрерывные функции, заданные на ограниченном замкнутом промежутке, обладают рядом важных свойств, рассмотренных в данном разделе. Напомним, что отрезком или замкнутым промежутком $[a, b]$ называется множество всех чисел x, таких, что $a \le x \le b$. Отрезок полностью задается своими концами — числами a и b, — и содержит эти концы.

Теорема *об ограниченности непрерывной функции на отрезке.*

Функция, непрерывная на отрезке, ограничена на этом отрезке.

Доказательство. Пусть функция f непрерывна на отрезке $[a, b]$. Докажем от противного ограниченность этой функции. Пусть f неограничена на $[a, b]$. Далее определим последовательности чисел $(a_n)_{n=0}^{\infty}$ и $(b_n)_{n=0}^{\infty}$ индуктивно. Положим $a_0 = a$ и $b_0 = b$. Очевидно, что f неограничена на $[a_0, b_0]$ по предположению.

Пусть при некотором целом n числа a_n и b_n определены, функция f неограничена на $[a_n, b_n]$, и определим следующую пару a_{n+1}, b_{n+1}. Разделим промежуток $[a_n, b_n]$ пополам точкой c_n и рассмотрим два отрезка: $[a_n, c_n]$ и $[c_n, b_n]$.

Если бы функция f была ограничена на отрезке $[a_n, c_n]$ (положительным числом M_1) и ограничена на отрезке $[c_n, b_n]$ (положительным числом M_2), то она была бы ограничена и на $[a_n, b_n]$ числом, равным наибольшему из чисел M_1 и M_2. Но f неограничена на $[a_n, b_n]$ по предположению. Следовательно, она неограничена хотя бы на одном из отрезков $[a_n, c_n]$ или $[c_n, b_n]$.

Если f неограничена на $[a_n, c_n]$, то положим $a_{n+1} = a_n$ и $b_{n+1} = c_n$, так что $[a_{n+1}, b_{n+1}] = [a_n, c_n]$ и на этом отрезке функция f неограничена. Если же f ограничена на $[a_n, c_n]$, то f должна быть неограничена на $[c_n, b_n]$, и поэтому положим $a_{n+1} = c_n$ и $b_{n+1} = b_n$. В обоих случаях функция f оказывается неограниченной на отрезке $[a_{n+1}, b_{n+1}]$.

Нетрудно видеть, что получается последовательность стягивающихся отрезков $[a_n, b_n]$, длины которых делятся пополам. При этом функция f неограничена на каждом из этих отрезков. Согласно аксиоме плотности вещественных чисел существует вещественное число c, принадлежащее каждому из упомянутых отрезков.

По предположению о непрерывности функции f на $[a, b]$ функция f непрерывна в точке c. По теореме об ограниченности функции вблизи точки непрерывности существует такой открытый интервал (p, q), содержащий точку c, что сужение f на (p, q) ограничено, т.е. $|f(x)| \leq M$ для $p < x < q$. Поскольку интервал (p, q) не содержит концов, то $c \neq p$ и $c \neq q$.

Длины интервалов $[a_n, b_n]$ стремятся к нулю при $n \to \infty$. Поэтому начиная с некоторого числа N длина $b_n - a_n$ будет меньше, чем $c - p$ и меньше, чем $q - c$. Интервал (p, q) содержит точку c, поэтому $[a_n, b_n] \subseteq (p, q)$. Получается противоречие: функция f ограничена на интервале (p, q) и одновременно неограничена на его части $[a_n, b_n]$. Противоречие доказывает, что исходное предположение о неограниченности функции f на отрезке $[a, b]$ было неверным. $\square$

Замечания. *1.* Если промежуток, на котором функция непрерывна, не является замкнутым и ограниченным, то функция может не быть ограниченной на нем. Например, функция $f(x) = 1/x$ на промежутке $(0, 1]$.

2. Метод доказательства последней теоремы содержит построение вложенных отрезков половинной длины, каждый из которых обладает особым свойством, в данном случае, неограниченностью на нем функции f. Такой способ конструирования вложенных отрезков называется *дихотомией*.

Теорема Вейерштрасса *о достижении наибольшего и наименьшего значений непрерывной функции на отрезке.* [K. Weierstrass, 1860–е годы]

Функция f, непрерывная на отрезке $[a, b]$, достигает на нем своего наибольшего и наименьшего значений, т.е. существуют такие числа c и d на $[a, b]$, что

$$f(c) \leq f(x) \leq f(d) \qquad \forall x \in [a, b].$$

Доказательство. По предыдущей теореме функция f ограничена на $[a, b]$. Следовательно, множество всех значений функции f на $[a, b]$, или образ $Z = f([a, b])$ промежутка $[a, b]$ при отображении f, является ограниченным множеством. По теореме о верхних и нижних гранях множество Z имеет точную верхнюю грань M и точную нижнюю грань m.

Если существует такое число $c \in [a, b]$, что $f(c) = m$ и существует такое число $d \in [a, b]$, что $f(d) = M$, то теорема доказана. Докажем существование таких чисел c и d от противного.

Пусть для всех чисел $c \in [a, b]$: $f(c) > m$. Тогда можно определить функцию $g(x) = 1/(f(x) - m)$ при $x \in [a, b]$. Функция g непрерывна на $[a, b]$ как частное непрерывных функций, в которых знаменатель не обращается в нуль. По предыдущей теореме функция g ограничена на $[a, b]$, т.е. существует такое число $A > 0$, что $g(x) \leq A$ при всех $x \in [a, b]$. Но тогда эквивалентным преобразованием получаем, что $f(x) > m + 1/A$, а это значит, что число $m + 1/A$ является нижней гранью для множества Z. Это противоречит выбору m как наибольшей нижней грани множества Z. Противоречие показывает, что существует такое $c \in [a, b]$, что $f(c) = m$.

Аналогично доказывается, что значение M также достигается функцией f на отрезке $[a, b]$. $\square$

Теорема Больцано–Коши *об образе отрезка при непрерывном отображении.* [B.Bolzano, 1817, и A.L.Cauchy, 1821]

Пусть функция f определена и непрерывна на отрезке $[a, b]$. Минимальное значение функции f на $[a, b]$ обозначим m, а максимальное — M. Тогда для любого числа $y \in [m, M]$ существует такое число $c \in [a, b]$, что $f(c) = y$.

(Непрерывная функция на отрезке принимает все значения между минимальным и максимальным.)

Доказательство. Если $y = m$ или $y = M$, то существование такого числа c следует из теоремы Вейерштрасса: существуют такие числа d и D на $[a, b]$, что $f(d) = m$ и $f(D) = M$.

Пусть $m < y < M$. Применим метод дихотомии. Пусть $[a_0, b_0]$ — отрезок с концами d и D (т.е. либо $[d, D]$, либо $[D, d]$ в зависимости от того, какое число больше).

Одна итерация при построении последовательности вложенных отрезков состоит в следующем. Пусть отрезок $[a_n, b_n]$ уже построен и число y расположено между числами $f(a_n)$ и $f(b_n)$. Определим c_n как середину отрезка $[a_n, b_n]$.

Если число y принадлежит отрезку с концами $f(a_n)$ и $f(c_n)$, то определим $a_{n+1} = a_n$ и $b_{n+1} = c_n$. Если это не так, то число y должно принадлежать отрезку с концами $f(b_n)$ и $f(c_n)$. Поэтому определим $a_{n+1} = c_n$, $b_{n+1} = b_n$. В обоих случаях число y расположено на отрезке с концами $f(a_{n+1})$ и $f(b_{n+1})$.

По аксиоме плотности вещественных чисел вложенные отрезки $[a_n, b_n]$ имеют общую точку. Обозначим ее c. Докажем, что $f(c) = y$. Поскольку длины отрезков $[a_n, b_n]$ стремятся к нулю, то $a_n \to c$ и $b_n \to c$ при $n \to \infty$. В точке c функция f непрерывна, т.е. $f(x) \to f(c)$ при $x \to c$. По теореме о секвенциальном пределе: $f(a_n) \to f(c)$ и $f(b_n) \to f(c)$. Но число y расположено между числами $f(a_n)$ и $f(b_n)$. По теореме о пределе сжатой последовательности $y = f(c)$. $\square$

Следствие *о нуле непрерывной функции, принимающей значения разных знаков.*

Пусть функция f непрерывна на отрезке $[a, b]$ и значения $f(a)$ и $f(b)$ отличны от нуля и имеют разные знаки. Тогда между a и b существует корень функции f, т.е. существует такое число $c \in (a, b)$, что $f(c) = 0$.

Определение *равномерной непрерывности функции на множестве.*

Пусть функция f определена на множестве Δ. Функция f называется *равномерно непрерывной на множестве* Δ, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 : \ \forall x_1, x_2 \in \Delta \ \text{ если } |x_1 - x_2| < \delta, \text{ то } |f(x_1) - f(x_2)| < \varepsilon.$$

Теорема Кантора–Гейне *о равномерной непрерывности функции на отрезке.* [G.Cantor, 1932 г., и E.Heine, 1872 г.]

Если функция f непрерывна на отрезке $[a, b]$, то она равномерно непрерывна на этом отрезке.

(Без доказательства.)

Производная

Производная функции в точке

Пусть $k \in \mathsf{R}$. Функция $\ell(h) = kh$, заданная на множестве вещественных чисел R, называется линейной формой. Эта функция непрерывна во всех точках $h \in \mathsf{R}$. Кроме того, $\ell(0) = 0$, и поэтому ℓ – бесконечно малая в точке 0.

Определение *дифференцируемой функции в точке и дифференциала.*

Пусть P — промежуток, $a \in P$ и функция f задана на P. Пусть существует такая линейная форма ℓ, что функция $\Delta_a f - \ell$ является бесконечно малой в точке 0 порядка выше первого.

Тогда функция f называется дифференцируемой в точке a, а функция ℓ называется дифференциалом функции f в точке a. Обозначение дифференциала: $d_a f$. $\square$

Если функция ℓ является дифференциалом f в точке a, то по определению,

$$\lim_{h \to 0} \frac{\Delta_a f(h) - \ell(h)}{h} = 0.$$

Линейная форма ℓ определяется равенством $\ell(h) = kh$, где $k \in \mathsf{R}$. Число k называется угловым коэффициентом линейной формы ℓ.

Определение *производной в точке.*

Пусть функция f дифференцируема в точке a. Тогда угловой коэффициент дифференциала функции f в точке a называется производной функции f в точке a.

Обозначение: $k = f'(a)$ или $\frac{df}{dx}(a)$. $\square$

По определению дифференциала функция β, определяемая равенством

$$\beta(h) = \frac{\Delta_a f(h) - d_a f(h)}{h}$$

является бесконечно малой в нуле.

Выполним ряд тождественных преобразований:

$$0 = \lim_{h \to 0} \beta(h) = \lim_{h \to 0} \frac{\Delta_a f(h) - f'(a)h}{h} = \lim_{h \to 0} \frac{\Delta_a f(h)}{h} - f'(a).$$

Следовательно,

$$f'(a) = \lim_{h \to 0} \frac{\Delta_a f(h)}{h}.$$

Сделаем замену переменной: $h = x - a$ и подставим $\Delta_a f(h) = f(a + h) - f(a)$. Тогда получится равенство

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

которое часто принимается за определение производной функции f в точке a. Проведенные рассуждения доказывают, что существование данного предела равносильно дифференцируемости функции f в точке a, и поэтому такая замена определений корректна.

Замечание *о непрерывности и дифференцируемости функции в точке.*

Если функция f дифференцируема в точке a, то она непрерывна в точке a.

Действительно, из определение дифференцируемости следует, что

$$\Delta_a f(h) = f'(a)h + \beta(h)h.$$

Обе функции в правой части являются бесконечно малыми в нуле. Поэтому функция f непрерывна в точке a.

Таким образом, дифференцируемыми могут быть только непрерывные функции. Однако, существуют примеры функций, которые непрерывны в точке, но не дифференцируемы в ней. Дифференцируемость — более сильное свойство, чем непрерывность.

Согласно определению непрерывности, функция f непрерывна в точке a, если

$$f(x) = f(a) + \alpha(h),$$

где $\alpha(h)$ — бесконечно малая в нуле. Если же функция f еще и дифференцируема в точке a, то эта функция α допускает уточнение:

$$f(x) = f(a) + f'(a)h + \beta(h)h,$$

которое выражается в выделении главной линейной части $f'(a)h$.

Теорема *о единственности производной.*

Пусть P – промежуток, $a \in P$ и функция f дифференцируема в точке a. Тогда существует единственная линейная форма ℓ, такая, что функция $\Delta_a f - \ell$ бесконечно малая в нуле порядка выше первого.

Другими словами, если функция имеет дифференциал в точке, то только один.

Доказательство.

От противного. Предположим, что дифференциала два: $\ell_1(h) = k_1 h$ и $\ell_2(h) = k_2 h$. Это значит, что существуют такие функции β_1 и β_2, бесконечно малые в нуле порядка выше первого, что

$$\Delta_a f(h) = k_1 h + h\beta_1(h) = k_2 h + h\beta_2(h).$$

Из полученного равенства следует, что

$$k_2 - k_1 = \frac{\beta_2(h) - \beta_1(h)}{h}.$$

Правая часть имеет предел при $h \to 0$, равный нулю. Следовательно, $k_2 - k_1 = 0$, т.е. $k_1 = k_2$. $\square$

Функция $f(x) = kx + b$ дифференцируема в любой точке. Действительно, если $a \in \mathsf{R}$ — произвольное число, то $\Delta_a f(h) = kh$. Линейная форма $\ell(h) = kh$ является дифференциалом функции f, так как остаток $\beta(h) = \Delta_a f(h) - \ell(h)$ тождественно равен нулю, и это бесконечно малая, порядка выше первого.

По теореме единственности других дифференциалов у функции $f(x) = kx + b$ нет и $f'(x) = k$. В частности, производная константы $f(x) = b$ равна нулю, так как $k = 0$.

Геометрическая интерпретация производной и дифференциала.

Функция вида $\phi(x) = kx + b$ называется линейной. В частности, линейная функция является линейной формой, если $b = 0$.

График линейной функции есть по определению следующее множество точек на плоскости R^2:

$$\Gamma_\phi = \{(x, y) \in \mathsf{R}^2 \ : \ y = kx + b\}.$$

Это прямая на плоскости. Если α — угол наклона графика к оси абсцисс, то $\operatorname{tg}\alpha = k$.

Пусть функция f дифференцируема в точке a. Рассмотрим в одной системе координат два графика: график Γ_f функции f и график Γ_ϕ линейной функции $\phi(x) = f(a)+f'(a)(x-a)$. Покажем, что график Γ_ϕ теснее любых других прямых примыкает к графику Γ_f.

Точка M с координатами $(a, f(a))$ лежит на обоих графиках. Для произвольного числа $x \in P$ определим две точки: $M_x(x, f(x))$ лежит на графике функции f и $M'_x(x, \phi(x))$ лежит на графике линейной функции ϕ.

Если перенести начало координат в точку M, то прямая Γ_ϕ станет графиком дифференциала функции f в точке a. При $x \to a$ обе точки — M_x и M'_x — сходятся к точке M, поэтому расстояние между ними стремится к нулю. Это верно для любой линейной функции, график которой проходит через точку M. Однако, если $\phi(x) = f(a) + f'(a)(x - a)$, то расстояние

$$|M_x M'_x| = |f(x) - f(a) - f'(a)(x - a)|$$

стремится к нулю быстрее, чем $x - a$. Прямая, обладающая таким свойством, единственная, что следует из единственности дифференциала. Следовательно, погрешность приближенного равенства вблизи точки a

$$f(x) \approx f(a) + k(x - a)$$

наименьшая при $k = f'(a)$. Это значит, что в пучке прямых, проходящих через точку M, график ϕ теснее всего примыкает к графику функции f. Такая прямая единственная, и она называется *касательной* к графику функции f в точке графика $(a, f(a))$.

Таким образом, график линейной функции, проходящий через точку $(a, f(a))$ и имеющий угловой коэффициент, равный $f'(a)$, является касательной к графику функции f.

Производная арифметических операций и сложной функции

При вычислении производной основными являются правила дифференцирования сложной функции, а также простейших функций, составляющих таблицу производных. Далее в этом пункте будем предполагать, что задан промежуток P, точка a находится внутри этого промежутка и две функции f и g заданы по крайней мере на P.

Теорема *о производной суммы, произведения, частного.*

Если функции f и g дифференцируемы в точке a, то

1. Сумма $f + g$ дифференцируема в точке a и

$$(f + g)'(a) = f'(a) + g'(a).$$

2. Произведение fg дифференцируемо в точке a и

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

3. Если $g(x) \neq 0\ \forall x \in P$, то частное $\frac{f}{g}$ дифференцируемо в точке a и

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

Доказательство. По условию существуют такие бесконечно малые в нуле функции $\alpha(h)$ и $\beta(h)$, что

$$\begin{aligned} f(a+h) - f(a) &= f'(a)h + h\alpha(h), \\ g(a+h) - g(a) &= g'(a)h + h\beta(h). \end{aligned}$$

1. Функция $\gamma(h) = \alpha(h) + \beta(h)$ является бесконечно малой в нуле. Кроме того,

$$(f+g)(a+h) - (f+g)(a) = (f'(a) + g'(a))h + h\gamma(h).$$

Линейная форма $\ell(h) = (f'(a) + g'(a))h$ удовлетворяет всем условиям для дифференциала. Следовательно, функция $f + g$ является дифференцируемой, а число $f'(a) + g'(a)$ является ее производной в точке a.

2. Обозначим $p = fg$ и вычислим приращение функции p в точке a:

$$\begin{aligned} p(a+h) - p(a) &= f(a)[g(a+h) - g(a)] + g(a)[f(a+h) - f(a)] \\ &\quad +(f(a+h) - f(a))(g(a+h) - g(a)) \\ &= f(a)(g'(a)h + h\beta(h)) + g(a)(f'(a)h + h\alpha(h)) \\ &\quad +h^2(f'(a) + \alpha(h))(g'(a) + h\beta(h)) = (f(a)g'(a) + f'(a)g(a))h + h\gamma(h), \end{aligned}$$

где

$$\gamma(h) = f(a)\beta(h) + g(a)\alpha(h) + h(f'(a) + \alpha(h))(g'(a) + h\beta(h)).$$

Сумма и произведение бесконечно малых друг на друга и на константу является бесконечно малыми функциями. Поэтому функция $\gamma(h)$ является бесконечно малой в нуле. Но

$$p(a+h) - p(a) = (f(a)g'(a) + f'(a)g(a))h + h\gamma(h).$$

По определению, функция p является дифференцируемой в точке a и ее производная равна $f(a)g'(a) + f'(a)g(a)$.

3. Сначала рассмотрим функцию $p = 1/g$. Вычислим приращение частного:

$$p(a+h) - p(a) = \frac{g(a) - g(a+h)}{g(a+h)g(a)} = \frac{-(g'(a)h + h\beta(h))}{g(a)g(a+h)} = -\frac{g'(a)}{g(a)^2}h + h\gamma(h),$$

где

$$\gamma(h) = \frac{g'(a)(g(a+h) - g(a)) - \beta(h)g(a)}{g(a)^2 g(a+h)}.$$

Функция $\gamma(h)$ — бесконечно малая в нуле. Следовательно, функция $p = 1/g$ дифференцируема в точке a и $p'(a) = -g'(a)/g^2(a)$.

Функция $f(x)/g(x)$ может быть представлена как $f(x)p(x)$. По доказанному в п. 2 свойству производной произведения,

$$\left(\frac{f}{g}\right)'(a) = f'(a)\frac{1}{g(a)} - f(a)\frac{g'(a)}{g^2(a)} = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2},$$

что совпадает с утверждением теоремы. □

Следствия.

1. Вынесение множителя из–под знака производной. Если функция f дифференцируема в точке a и $\lambda \in \mathsf{R}$ — некоторое число, то $(\lambda \cdot f)'(a) = \lambda \cdot f'(a)$.

Для доказательства достаточно применить формулу производной произведения и подставить $(\lambda)' = 0$.

2. Производная многочлена. Многочленом (или полиномом) называется функция

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

в которой a_0, a_1, a_2, …, a_n — вещественные числа. Докажем, что

$$f'(x) = a_1 + 2a_2 x + \ldots + n a_n x^{n-1}.$$

Сначала рассмотрим частный случай $f(x) = x^k$. Докажем по индукции, что $(x^k)' = kx^{k-1}$. При $k = 1$ равенство $x' = 1$ следует из определения производной. Пусть равенство доказано для k, докажем его для $k+1$:

$$(x^{k+1})' = (x^k \cdot x)' = (x^k)'x + x^k x' = kx^{k-1}x + x^k = (k+1)x^k,$$

что доказывает индукционный шаг.

Рассмотрим произвольный многочлен $f(x)$. Для вычисления его производной остается воспользоваться теоремой о производной суммы, следствием о вынесении множителя из–под знака производной и формулой для производной от x^k.

Теорема *о производной сложной функции.*

Пусть функция f задана на промежутке P_f, функция g задана на промежутке P_g, $a \in P_f$ и $f(P_f) \subseteq P_g$. Пусть функция f дифференцируема в точке a, а функция g дифференцируема в точке $f(a)$. Тогда функция $g \circ f$ дифференцируема в точке a и

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

Доказательство. Пусть $b = f(a)$. Рассмотрим два случая: $f'(a){\neq}0$ и $f'(a){=}0$.

1. Пусть $f'(a) \neq 0$. Тогда существует предел

$$\lim_{h\to 0}(f'(a) + \alpha(h)) = 0.$$

По теореме об отделенности от нуля непрерывной функции существует такая окрестность нуля $\mathcal{O}$, что функция

$$r(h) = f'(a) + \alpha(h)$$

отлична от нуля в этой окрестности. Следовательно, и функция $\Delta_a f(h) = hr(h)$ отлична от нуля при всех $h \in \mathcal{O}$. В этой окрестности на величину $\Delta_a f(h) = f(a+h) - f(a)$ можно разделить:

$$\frac{\Delta_a(g\circ f)(h)}{h} = \frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} \cdot \frac{f(a+h) - f(a)}{h}.$$

Вычислим предел этой функции при $h \to 0$. Предел второго сомножителя существует и равен $f'(a)$ по условию дифференцируемости функции f.

При вычислении предела второго сомножителя воспользуемся заменой переменной: $y = f(a+h)$. Поскольку существует предел $f(a+h)$ при $h \to 0$ и существует предел

$$\lim_{y\to b}\frac{g(y) - g(b)}{y - b} = g'(b),$$

то по теореме о пределе композиции функций существует предел

$$\lim_{h\to 0}\frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} = g'(b).$$

В следующем равенстве предел произведения равен произведению пределов, потому что оба предела существуют и только что найдены:

$$\lim_{h\to 0}\frac{\Delta_a(g\circ f)(h)}{h} = \lim_{h\to 0}\frac{g(f(a+h)) - g(f(a))}{f(a+h) - f(a)} \cdot \frac{f(a+h) - f(a)}{h} = f'(a)g'(b),$$

что совпадает с заключением теоремы.

2. Пусть $f'(a) = 0$. Из определения дифференцируемости функций f и g следует, что существуют такие бесконечно малые в нуле функции $\alpha(h)$, $\beta(z)$, что

$$\begin{aligned} f(a+h) - b &= h\alpha(h), \\ g(b+z) - g(b) &= g'(b)z + z\beta(z). \end{aligned}$$

Подставим вместо z число $f(a+h) - b$:

$$g(f(a+h)) - g(b) = h\alpha(h)[g'(b) + \beta(h\alpha(h))].$$

В левой части стоит приращение функции $g \circ f$ в точке a. Поэтому

$$\lim_{h\to 0} \frac{\Delta_a(g\circ f)(h)}{h} = \lim_{h\to 0} \alpha(h)[g'(b) + \beta(h\alpha(h))] = 0.$$

В последнем равенстве при вычислении предела $\beta(h\alpha(h)$ применена теорема о пределе композиции функций. Итак, доказано, что существует и равна нулю производная функции $g \circ f$ в точке a. Но и произведение $f'(a) \cdot g'(b)$ равно нулю, что завершает доказательство теоремы. □

Неявная функция

Функция f может быть задана формулой $y = f(x)$ с конкретным правилом вычисления значения y по предъявленному аргументу x. Такой способ определения функции называется явным.

Однако, функцию можно задать и другим способом, включающим решение уравнения. Пусть задана функция $F(x, y)$ двух аргументов и P — некоторое множество аргументов x. Пусть выполнено следующее условие: при каждом фиксированном $x \in P$ уравнение

$$F(x, y) = 0$$

имеет единственное решение относительно y. Это решение зависит, естественно, от x.

Получается следующее правило, ставящее в соответствие каждому элементу x множества P число y путем решения уравнения $F(x, y) = 0$. Это правило вместе с областью P аргументов x и множеством R значений y образует отображение $f: P \to \mathsf{R}$.

Говорят, что отображение f задано неявно уравнением $F(x, y) = 0$. Иногда также функцию f (с некоторой языковой погрешностью) называют неявной функцией.

Предположим, что уравнение $F(x, y) = 0$ задает неявно функцию $y = f(x)$ на промежутке P. Тогда уравнение $F(x, f(x)) = 0$ выполнено при всех $x \in P$. Другими словами, функция $G(x) = F(x, f(x))$ тождественно равна нулю на промежутке P. Но тогда $G'(x) = 0$ как производная константы.

Пусть $x_0 \in P$ и функция f дифференцируема в точке x_0. Обозначим $y_0 = f(x_0)$, т.е. $F(x_0, y_0) = 0$. Пусть производная функции G в точке x_0 может быть записана в виде $G'(x_0) = q(x_0) + f'(x_0)r(x_0)$, причем $r(x_0) \neq 0$. Тогда $f'(x_0) = -q(x_0)/r(x_0)$.

Пример. Пусть функция $y = f(x)$ задана неявно уравнением $a^2x^2 + b^2y^2 = c^2$, где $a, b, c, > 0$, и условиями $x \in [0, c/a]$, $y \geq 0$. Вычислим производную этой функции в точке $x_0 \in (0, c/a)$. В данном случае

$$G(x) = a^2x^2 + b^2f(x)^2 - c^2, \qquad G'(x_0) = 2a^2x_0 + 2b^2f(x_0)f'(x_0).$$

Поэтому
$$f'(x_0) = -\frac{a^2x_0}{b^2y_0},$$
где y_0 определяется из условия $a^2x_0^2 + b^2y_0^2 = c^2$.

Производная неявной функции применяется, частности, при вычислении производной обратной функции, а также при параметрическом дифференцировании.

Теорема *о производной обратной функции.*

Пусть функция f задана на промежутке P и взаимно однозначно отображает P на промежуток Q. Обратную функцию обозначим g. Пусть $a{\in}P$, $b = f(a){\in}Q$, функция f дифференцируема в точке a и $f'(a) \neq 0$.

Тогда функция g дифференцируема в точке b и $g'(b) = 1/f'(a)$.

Схема доказательства. Обратная функция обладает по ее определению свойством $g(f(x)) = x$ при всех $x \in P$. Вычислим производную от обеих частей этого тождества. Если функция g дифференцируема в точке $b = f(a)$, то по правилу производной сложной функции $g'(b)f'(a) = 1$, откуда и следует заключение теоремы. Доказательство дифференцируемости обратной функции $g(y)$ будет дано позднее для более общей задачи с функциями нескольких переменных. □

Параметрическое дифференцирование.

Иногда функция $y = f(x)$ задается не явным указанием на способ расчета величины y по аргументу x, а опосредованно, параметрически.

Пусть две функции $M : T \to \mathsf{R}$ и $N : T \to \mathsf{R}$ заданы на одном и том же промежутке T. Значения функции M будем обозначать x, а функции N — y. Пусть функция M взаимно однозначно отображает промежуток T на промежуток P. Тогда существует и обратное к M отображение $S : P \to T$, т.е. $S(M(t)) = t$ при всех $t \in T$. Функция $f : P \to \mathsf{R}$, заданная как композиция $f = N \circ S$ называется заданной параметрически.

Сделанные предположения иногда формулируют следующим образом: заданы две функции $x = M(t)$ и $y = N(t)$ на промежутке T, причем функция M взаимно однозначна. Тогда y является функцией x, а эта функция $y = f(x)$ называется заданной параметрически.

Вычислим производную функции f, пользуясь формулой для производной обратной функции и производной сложной функции. Пусть $t_0 \in T$, $x_0 = M(t_0) \in P$ и $y_0 = N(t_0)$. Пусть $M'(t_0) \neq 0$. Тогда существует производная обратной функции S в точке x_0. Вычислим производную сложной функции $y = N(S(x))$:
$$f'(x_0) = S'(x_0) \cdot N'(S(x_0)) = \frac{N'(t_0)}{M'(t_0)}.$$
Последнюю формулу записывают также в форме
$$f'(x_0) = \frac{df}{dx}(x_0) = \frac{dN}{dt}(t_0) \Bigg/ \frac{dM}{dt}(t_0).$$

Техника дифференцирования.

Производные элементарных функций вычисляются как предел отношения приращения функции к приращению аргумента. Далее все производные вычисляются в точках x_0, находящихся внутри их областей определения.

1. Показательная функция. Пусть $a > 0$.

$$\lim_{h\to 0} \frac{a^{x_0+h} - a^{x_0}}{h} = a^{x_0} \lim_{h\to 0} \frac{a^h - 1}{h} = a^{x_0} \ln a.$$

В частности, $(e^x)' = e^x$.

2. Логарифмическая функция. Пусть $a > 0$.

$$\lim_{h\to 0} \frac{\log_a(x_0+h) - \log_a(x_0)}{h} = \lim_{h\to 0} \frac{\log_a \frac{x_0+h}{x_0}}{h} = \lim_{h\to 0} \frac{\log_a(1+\frac{h}{x_0})}{h} = \frac{1}{x_0 \ln a}.$$

В частности, $(\ln x)' = 1/x$.

3. Степенная функция. Пусть $\mu \in \mathsf{R}$ и $\mu \neq 0$.

$$\lim_{h\to 0} \frac{(x_0+h)^\mu - x_0^\mu}{h} = x_0^\mu \lim_{h\to 0} \frac{(1+\frac{h}{x_0})^\mu - 1}{h} = x_0^\mu \lim_{h\to 0} \frac{\mu\left(\frac{h}{x_0}\right)}{h} = \mu x_0^{\mu-1}.$$

В частности,

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

4. Тригонометрические функции.

4а. Синус:

$$\lim_{h\to 0} \frac{\sin(x_0+h) - \sin x_0}{h} = \lim_{h\to 0} \frac{2\sin(h/2)\cos(x_0+h/2)}{h} = \cos x_0.$$

4б. Косинус:

$$\lim_{h\to 0} \frac{\cos(x_0+h) - \cos x_0}{h} = \lim_{h\to 0} \frac{2\sin(-h/2)\sin(x_0+h/2)}{h} = -\sin x_0.$$

4в. Тангенс:

$$\lim_{h\to 0} \frac{\operatorname{tg}(x_0+h) - \operatorname{tg} x_0}{h} = \lim_{h\to 0} \frac{\sin(h)}{h\cos(x_0+h)\cos x_0} = \frac{1}{\cos^2 x_0}.$$

4г. Котангенс:

$$\lim_{h\to 0} \frac{\operatorname{ctg}(x_0+h) - \operatorname{ctg} x_0}{h} = \lim_{h\to 0} \frac{-\sin(h)}{h\sin(x_0+h)\sin x_0} = -\frac{1}{\sin^2 x_0}.$$

5. Обратные тригонометрические функции. Воспользуемся теоремой о производной обратной функции.

5а. Арксинус. Пусть $-1 < x_0 < 1$, $y_0 = \arcsin x_0 \in (-\pi/2, \pi/2)$. Тогда $x_0 = \sin y_0$. По правилу вычисления производной обратной функции

$$(\arcsin)'(x_0) = \frac{1}{\cos y_0} = \frac{1}{\sqrt{1-\sin^2 y_0}} = \frac{1}{\sqrt{1-x_0^2}}.$$

5б. Арккосинус. Пусть $-1 < x_0 < 1$, $y_0 = \arccos x_0 \in (0, \pi)$. Тогда $x_0 = \cos y_0$. По правилу вычисления производной обратной функции

$$(\arccos)'(x_0) = \frac{1}{-\sin y_0} = -\frac{1}{\sqrt{1-\cos^2 y_0}} = -\frac{1}{\sqrt{1-x_0^2}}.$$

5в. Арктангенс. Пусть $x_0 \in \mathsf{R}$, $y_0 = \operatorname{arctg} x_0 \in (-\pi/2, \pi/2)$. Тогда $x_0 = \operatorname{tg} y_0$. По правилу вычисления производной обратной функции

$$(\operatorname{arctg})'(x_0) = \cos^2 y_0 = \frac{1}{1 + \operatorname{tg}^2 y_0} = \frac{1}{1 + x_0^2}.$$

5г. Арккотангенс. Пусть $x_0 \in \mathsf{R}$, $y_0 = \operatorname{arcctg} x_0 \in (0, \pi)$. Тогда $x_0 = \operatorname{ctg} y_0$. По правилу вычисления производной обратной функции

$$(\operatorname{arcctg})'(x_0) = -\sin^2 y_0 = -\frac{1}{1 + \operatorname{ctg}^2 y_0} = -\frac{1}{1 + x_0^2}.$$

Полученные формулы дифференцирования элементарных функция, а также арифметических операций сведены в таблицу производных:

Таблица производных

1. $(u + v)' = u' + v'$.
2. $(uv)' = u'v + uv'$.
3. $\left(\dfrac{u}{v}\right)' = \dfrac{u'v - uv'}{v^2}$.
4. $(e^x)' = e^x, \qquad (a^x)' = a^x \ln a, \quad (a > 0)$.
5. $(\ln x)' = \dfrac{1}{x}, \qquad (\log_a x)' = \dfrac{1}{x \ln a}, \quad (a > 0, a \neq 1)$.
6. $(x^\alpha)' = \alpha x^{\alpha - 1}, \qquad \left(\dfrac{1}{x}\right)' = -\dfrac{1}{x^2}$.
7. $(\sin x)' = \cos x$.
8. $(\cos x)' = -\sin x$.
9. $(\operatorname{tg} x)' = \dfrac{1}{\cos^2 x}$.
10. $(\operatorname{ctg} x)' = -\dfrac{1}{\sin^2 x}$.
11. $(\arcsin x)' = \dfrac{1}{\sqrt{1 - x^2}}$.
12. $(\arccos x)' = -\dfrac{1}{\sqrt{1 - x^2}}$.
13. $(\operatorname{arctg} x)' = \dfrac{1}{1 + x^2}$.
14. $(\operatorname{arcctg} x)' = -\dfrac{1}{1 + x^2}$.

Производная от любой сложной функции, состоящей из элементарных функций, может быть рассчитана по этой таблице и по правилу вычисления производной композиции функций: $(g(f(x)))' = g'(f(x)) \cdot f'(x)$.

Иногда для вычисления производной полезно выполнить предварительное тождественное преобразование. Одно из них называется логарифмическим дифференцированием.

Логарифмическое дифференцирование. Этот прием основан на формуле

$$(\ln f(x))' = \frac{f'(x)}{f(x)},$$

которую можно также записать в виде

$$f'(x) = f(x)(\ln f(x))'.$$

Наиболее часто логарифмическое дифференцирование применяется при вычислении производной от функции вида $u(x)^{v(x)}$. Эта функция не является ни степенной, ни показательной, поэтому не подходит ни одно из правил таблицы производных. Однако, можно воспользоваться логарифмическим дифференцированием при $f(x) = u(x)^{v(x)}$:

$$\left(u(x)^{v(x)}\right)' = u(x)^{v(x)}(v(x)\ln u(x))' = u(x)^{v(x)}\left(v'(x)\ln u(x) + \frac{v(x)u'(x)}{u(x)}\right).$$

Свойства функции, дифференцируемой на промежутке

Определение *производной функции.*

Пусть функция f задана на промежутке P и дифференцируема в каждой точке этого промежутка. Тогда *производная функция от функции* f (или *производная функции* f) на промежутке P есть отображение, ставящее в соответствие каждой точке $x \in P$ значение $f'(x)$ производной функции f в точке x.

Множество всех функций, заданных и дифференцируемых в каждой точке промежутка P, у которых производная функция непрерывна, называется классом $C^1(P)$. Множество всех функций, заданных и непрерывных в каждой точке промежутка P, называется классом $C(P)$. Очевидно, что $C^1(P) \subset C(P)$. □

Теорема Ферма *о производной в точке экстремума.* [P. Fermat, 1629 г.]

Пусть P — промежуток, $a \in P$, функция f задана на P и дифференцируема в точке a. Пусть выполнены два условия:

1) функция f достигает в точке a максимального или минимального значения на P; и

2) точка a — внутренняя точка промежутка P.

Тогда $f'(a) = 0$.

Доказательство. От противного. Пусть $f'(a) \neq 0$. По определению производной

$$f(a+h) = f(a) + f'(a)h + h\beta(h),$$

причем $\beta(h) \to 0$ при $h \to 0$. Выберем $\varepsilon = |f'(a)|/2$. По определению предела существует такое число $\delta > 0$, что если $|h| < \delta$ и $a + h \in P$, то $|\beta(h)| < \varepsilon$.

Воспользуемся условием о том, что a — внутренняя точка P. Если $a-\delta$ или $a+\delta$ не принадлежит P, то число δ можно уменьшить так, чтобы оба числа — $a-\delta$ и $a+\delta$ принадлежали P. Тогда

$$-\delta < h < \delta \;\Rightarrow\; a + h \in P \text{ и } |\beta(h)| < |f'(a)|/2.$$

Пусть f достигает максимума в точке a на промежутке P. Выберем h в промежутке $(-\delta, \delta)$ того же знака, что и число $f'(a)$. Тогда

$$\begin{aligned} f(a+h) - f(a) &= hf'(a) + h\beta(h)) \geq |h|\,|f'(a)| - |h|\,|\beta(h)| \\ &= |h|(|f'(a)| - |\beta(h)|) \geq |h|\,|f'(a)|/2 > 0. \end{aligned}$$

Таким образом, $f(a+h) > f(a)$, что противоречит условию о максимуме f в точке a.

Если f достигает минимума в точке a, то число h следует выбрать обратного знака к знаку $f'(a)$, и снова получится противоречие. □

Теорема Ролля *о корне производной.* [M. Rolle, 1690 г.]

Пусть функция f задана на отрезке $[a,b]$, непрерывна в точках a и b и дифференцируема на открытом интервале (a,b).

Пусть $f(a)=f(b)$. Тогда существует такая точка $\xi\in(a,b)$, что $f'(\xi)=0$.

Доказательство. Функция f непрерывна в каждой точке интервала (a,b), так как она дифференцируема в этих точках. Следовательно, f непрерывна на отрезке $[a,b]$ и можно применить теоремы о функциях, непрерывных на отрезке.

По теореме Вейерштрасса f достигает наибольшего и наименьшего значений на $[a,b]$. Пусть наибольшее значение достигается в точке c, а наименьшее — в точке d. Если хотя бы одна из этих точек не совпадает с концами промежутка $[a,b]$, то она внутренняя, и по теореме Ферма производная в ней равна нулю, что доказывает данную теорему.

Пусть и c, и d совпадают с концами промежутка $[a,b]$. Но $f(a)=f(b)$, и поэтому наименьшее значение функции f на $[a,b]$ совпадает с наибольшим значением. Такое может быть только если функция f постоянна на $[a,b]$. Производная у константы равна нулю во всех точках интервала (a,b). □

Теорема Лагранжа *о конечных приращениях.* [J.L.Lagrange, 1760-е гг.]

Пусть функция f задана на отрезке $[a,b]$, непрерывна на концах a и b и дифференцируема на открытом интервале (a,b).

Тогда существует такая точка $\xi\in(a,b)$, что

$$f(b)-f(a)=f'(\xi)(b-a).$$

Доказательство. Определим функцию

$$g(x)=f(x)+(x-a)\frac{f(a)-f(b)}{b-a}.$$

Эта функция непрерывна в точках a и b, а также дифференцируема на интервале (a,b). Кроме того, $g(a)=g(b)=f(a)$. Поэтому можно применить теорему Ролля. Существует такая точка $\xi\in(a,b)$, что $g'(\xi)=0$. Из определения функции g следует, что

$$f'(\xi)-\frac{f(b)-f(a)}{b-a}=0,$$

что равносильно заключению теоремы. □

Теорема Коши *о среднем значении.* [A.L.Cauchy, 1810–е годы]

Пусть функции f и g заданы на отрезке $[a,b]$, непрерывны на концах a и b и дифференцируемы на открытом интервале (a,b). Пусть $g'(x)\neq 0$ на интервале (a,b).

Тогда существует такая точка $\xi\in(a,b)$, что

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}.$$

Доказательство. Докажем сначала, что формула из заключения теоремы имеет смысл, т.е. $g(a)\neq g(b)$. Если предположить, что $g(a)=g(b)$ то для функции g можно применить теорему Ролля, по которой $g'(\xi)=0$ в некоторой точке $\xi\in(a,b)$. Это противоречит предположению $g'(x)\neq 0$ на (a,b).

Определим на $[a, b]$ функцию

$$h(x) = f(x) + (g(x) - g(a))\,\frac{f(a) - f(b)}{g(b) - g(a)}.$$

Функция h удовлетворяет всем условиям теоремы Ролля, в частности, $h(a) = h(b) = f(a)$. По теореме Ролля существует такая точка $\xi \in (a, b)$, что $h'(\xi) = 0$. Из определения функции h следует, что

$$f'(\xi) - g'(\xi)\,\frac{f(b) - f(a)}{g(b) - g(a)} = 0,$$

что равносильно заключению теоремы. □

Теорема. *Правило Лопиталя раскрытия неопределенностей.* [G.F.l'Hospital]

Пусть $a \in \mathsf{R}$ и $\mathcal{O}$ – проколотая окрестность точки a. Пусть функции f и g дифференцируемы на $\mathcal{O}$, являются бесконечно малыми в точке a. Пусть производная $g'(x)$ не обращается в нуль на $\mathcal{O}$ и существует предел

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = A.$$

(конечный или бесконечный). Тогда существует и равен A предел

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

Аналогично, предел отношения бесконечно больших функций f и g при $x \to a$ (в том числе при $a = \pm\infty$ или при односторонних пределах) равен пределу отношения их производных при $x \to a$, если последний существует (конечный или бесконечный).

Доказательство. Докажем теорему в случае $a \in \mathsf{R}$ и $A \in \mathsf{R}$, т.е. когда точка a и предел A – конечны. Доопределим функции f и g в точке a нулем: $f(a) = g(a) = 0$. Значения функциq в точке a не влияют на пределы этих функций в точке a. Однако, после такого доопределения функции f и g становятся непрерывными в точке a.

Пусть $x \in \mathcal{O}$. Применим теорему Коши о среднем значении к отрезку $[a, x]$, если $a < x$ или к отрезку $[x, a]$, если $a > x$:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)},$$

где число $\xi = \xi(x)$ находится между a и x. Очевидно, что $\xi(x) \to a$ при $x \to a$. Поэтому по теореме о пределе сложной функции существует предел

$$\lim_{x \to a} \frac{f'(\xi(x))}{g'(\xi(x))} = \lim_{x \to a} \frac{f(x)}{g(x)},$$

который совпадает с A. □

Замечания. *1.* Правило Лопиталя позволяет вычислять предел отношения бесконечно малых или бесконечно больших через предел отношения их производных, если последний существует. Если же предел отношения производных не существует, то предел отношения исходных функций может, тем не менее, существовать. Например, существует следующий предел:

$$\lim_{x \to 0} \frac{(\sin x)^2 \sin \frac{1}{x}}{x} = 0,$$

однако предел отношения производных

$$\lim_{x\to 0}\,(2\sin x\cos x\sin\frac{1}{x}-\frac{\sin^2 x\cos\frac{1}{x}}{x^2})$$

не существует.

Поэтому все равенства, которые ставятся между пределом отношения функций и пределом отношения их производных, являются условными: если в конце получится какой–либо предел, то они справедливы, а если не получится – то они могут быть неверными.

2. Иногда правило Лопиталя требуется применять несколько раз подряд. Если оказалось, что отношение производных тоже имеет вид неопределенности типа $\frac{0}{0}$ или $\frac{\infty}{\infty}$, то можно снова продифференцировать отдельно числитель и знаменатель. Например:

$$\lim_{x\to 0}\frac{\ln\cos x}{x^2}=\lim_{x\to 0}\frac{-\sin x}{2x\cos x}=\lim_{x\to 0}\frac{-\cos x}{2\cos x-2x\sin x}=-\frac{1}{2}.$$

Производные высших порядков

Пусть функция f дифференцируема на промежутке P. Производная функция f' задана на промежутке P. Она, в свою очередь, также может быть дифференцируемой в некоторых точках промежутка P. Если $a\in P$ и функция f' дифференцируема в точке a, то ее производная $(f')'$ называется второй производной функции f и обозначается $f''(a)$.

Если вторая производная функции f существует в каждой точке промежутка P, то отображение $x\to f''(x)$, заданное на P, называется второй производной функцией функции f.

Если функция f'' дифференцируема в каждой точке промежутка P, то ее производная называется третьей производной функции f на P. По индукции, если n–ая производная функции f дифференцируема на P, то ее производная называется $(n+1)$–ой производной функции f.

n–ая производная функции f называется также производной порядка n функции f и обозначается $f^{(n)}$ или $\frac{d^n f}{dx^n}$. Младшие производные обозначаются, как правило, штрихами: $f^{(1)}=f'$, $f^{(2)}=f''$, $f^{(3)}=f'''$.

Если у функции существует n–ая производная, то существуют и все производные порядка меньше n, что следует из определения. Множество всех функций, у которых существует и непрерывна n–ая производная на промежутке P, называется классом $C^n(P)$.

Формула Тейлора

Поставим следующую задачу: найти многочлен степени n, у которого заданы значения производных в точке a вплоть до порядка n.

Лемма *о многочлене с заданными производными в точке.*

Пусть $a\in\mathsf{R}$ и задан набор вещественных чисел $(c_k)_{k=0}^{n}$. Тогда существует и единственен многочлен P степени не выше n, у которого $P^{(k)}(a)=c_k$ при всех $k=0,1,\ldots,n$.

Этот многочлен может быть записан в форме

$$P(x) = \sum_{k=0}^{n} \frac{c_k}{k!}(x-a)^k = c_0 + \frac{c_1}{1!}(x-a) + \frac{c_2}{2!}(x-a)^2 + \cdots + \frac{c_n}{n!}(x-a)^n.$$

Доказательство. Проверим, что предъявленный многочлен P обладает требуемыми свойствами. Очевидно, что степень P не выше n. Проверим условия на производные в точке a.

Из определения следует, что $P(a) = c_0$. Продифференцируем многочлен P:

$$P'(x) = c_1 + \frac{c_2}{1!}(x-a) + \frac{c_3}{2!}(x-a)^2 + \cdots + \frac{c_n}{(n-1)!}(x-a)^{n-1}.$$

Следовательно, $P'(a) = c_1$.

Продифференцируем еще раз полученный многочлен:

$$P''(x) = c_2 + \frac{c_3}{1!}(x-a) + \frac{c_4}{2!}(x-a)^2 + \cdots + \frac{c_n}{(n-2)!}(x-a)^{n-2}.$$

Следовательно, $P''(a) = c_2$.

Далее применим метод индукции по n. Получим $P^{(k)}(a) = c_k$ при $0 \le k \le n$.

Докажем единственность многочлена P. Пусть Q — другой многочлен степени не выше n, у которого те же производные в точке a. Выполним следующее преобразование:

$$Q(x) = \sum_{k=0}^{n} Q_k x^k = \sum_{k=0}^{n} Q_k((x-a)+a)^k = \sum_{k=0}^{n} q_k (x-a)^k,$$

где коэффициенты q_k получаются после раскрытия скобок в биномах $(y+a)^j$ и замене y на $(x-a)$. Степени выше n не появляются.

Непосредственным дифференцированием проверяется, что $Q^{(k)}(a) = q_k k!$, откуда $q_k = k! c_k$ и, следовательно, $Q = P$. $\square$

Определение *многочлена Тейлора.* [B.Taylor, 1715]

Пусть функция f имеет n производных в точке a. Тогда многочлен

$$\mathcal{P}_{f,n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

называется многочленом Тейлора порядка n функции f в точке a.

Основное свойство многочлена Тейлора:

$$\mathcal{P}^{(k)}_{f,n,a}(a) = f^{(k)}(a), \qquad 0 \le k \le n,$$

и степень $\mathcal{P}_{f,n,a}$ не превосходит n. $\square$

Из определения непосредственно следует, что функция $f - \mathcal{P}_{f,n,a}$ имеет n производных в точке a, и все они равны нулю. Функция $R = f - \mathcal{P}_{f,n,a}$ называется остаточным членом формулы Тейлора, так как $f(x) = \mathcal{P}_{f,n,a}(x) + R(x)$. Остаточный член определяет точность приближения функции f своим многочленом Тейлора.

Лемма. *(Обобщенная теорема Ролля.)*

Пусть функция f имеет $n+1$ производную на промежутке $[a,b]$ и пусть $f(a) = f(b) = 0$, $f^{(k)}(a) = 0$ при всех $k = 0, 1, \ldots, n$. Тогда существует такая точка $\xi \in (a,b)$, что $f^{(n+1)}(\xi) = 0$.

Доказательство. Проведем доказательство индукцией по n. Если $n = 0$, то получается теорема Ролля: из $f(a) = f(b)$ следует, что $f'(\xi) = 0$ в некоторой точке $\xi \in (a, b)$.

Индукционный шаг. Пусть утверждение доказано для $n-1$, докажем его для n. По индукционному предположению существует такая точка $\eta \in (a, b)$, что $f^{(n)}(\eta) = 0$. Но по условию $f^{(n)}(a) = 0$. Применим теорему Ролля к функции $f^{(n)}(x)$ на промежутке $[a, \eta]$. Существует такая точка $\xi \in (a, \eta)$, что $(f^{(n)})'(\xi) = 0$. Но в этом случае $\xi \in (a, b)$ и $f^{(n+1)}(\xi) = 0$. □

Замечание. Утверждение леммы остается справедливым, если нулевые производные заданы не на левом, а на правом конце промежутка: $f(a) = f(b) = 0$ и $f^{(k)}(b) = 0$ при всех $k = 0, 1, \ldots, n$.

Теорема. *Формула Тейлора с остаточным членом в форме Лагранжа.*

Пусть функция f имеет $n + 1$ производную на промежутке P и $a \in P$. Тогда для любого $x \in P$ существует точка ξ между a и x, такая, что остаточный член формулы Тейлора равен

$$R(x) = f(x) - \mathcal{P}_{f,n,a}(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-a)^{n+1}.$$

В развернутой форме это равенство можно записать в виде

$$f(x) = f(a) + f'(a)\frac{(x-a)}{1!} + f''(a)\frac{(x-a)^2}{2!} + \cdots + f^{(n)}(a)\frac{(x-a)^n}{n!} + f^{(n+1)}(\xi)\frac{(x-a)^{n+1}}{(n+1)!}$$

(формула Тейлора).

Доказательство. Пусть для определенности $a < x$ (при $x < a$ доказательство аналогично). Подберем число c так, чтобы функция

$$g(y) = R(y) - c(y-a)^{n+1}, \qquad y \in P,$$

имела корень в точке $y = x$. Достаточно выбрать $c = R(x)/(x-a)^{n+1}$, и тогда $g(x) = 0$.

Проверим для функции g на промежутке $[a, x]$ все условия обобщенной теоремы Ролля. Функция g имеет $n + 1$ производную на $[a, x]$, так как ее имеет функция f, а многочлены $\mathcal{P}_{f,n,a}(y)$ и $c(y-a)^{n+1}$ имеют все производные во всех вещественных точках.

В точке $y = a$ первые n производных функции $R(y)$ равны нулю, так как это остаточный член формулы Тейлора. У многочлена $c(y-a)^{n+1}$ каждое дифференцирование понижает степень $(y-a)$ ровно на 1. Поэтому в точке $y = a$ все производные до порядка n равны нулю. Кроме того, $g(x) = g(a) = 0$.

По обобщенной теореме Ролля существует такое число $\xi \in (a, x)$, что $g^{(n+1)}(\xi) = 0$. Вычислим эту производную:

$$0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (\mathcal{P}_{f,n,a})^{(n+1)}(\xi) - (c(y-a)^{n+1})^{(n+1)}|_{y=\xi}.$$

Многочлен Тейлора $\mathcal{P}_{f,n,a}$ имеет степень не выше n по определению. Поэтому его производная порядка $n+1$ равна нулю. Производная от $(y-a)^{n+1}$ порядка $n+1$ есть постоянная величина, равная $(n+1)!$. Следовательно,

$$0 = f^{(n+1)}(\xi) - c(n+1)!\,, \quad \text{или} \quad c = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

После подстановки этого значения c в уравнение $g(x) = 0$ получится утверждение теоремы. □

Теорема. *Остаточный член формулы Тейлора в форме Пеано.*

Пусть функция f имеет n производных на промежутке P и $a \in P$. Тогда остаточный член формулы Тейлора есть бесконечно малая в точке a порядка выше n–го, т.е.

$$R(x) = f(x) - \mathcal{P}_{f,n,a}(x) = o((x-a)^n)$$

при $x \to a$. □

Без доказательства.

Остаточный член в форме Лагранжа равен $(x-a)^{n+1}f^{(n+1)}(\xi)/(n+1)!$, и это есть бесконечно малая в точке a порядка не менее $n+1$. В этом смысле оценка по Лагранжу является более сильным результатом, так как уточняет оценку по Пеано. Однако, оценка по Пеано справедлива и в том случае, когда функция f имеет ровно n производных в точке a, а $n+1$–ю производную не имеет, а оценка по Лагранжу в таком случае невозможна.

Вычисление многочленов Тейлора для некоторых элементарных функций

Формула Тейлора в точке $a = 0$ также называется формулой Маклорена:

$$f(x) = f(0) + f'(0)\frac{x}{1!} + f''(0)\frac{x^2}{2!} + \cdots + f^{(n)}(0)\frac{x^n}{n!} + R(x),$$

где $R(x)$ — остаточный член формулы Маклорена. По теореме Пеано $R(x)/x^n \to 0$ при $x \to 0$, а если функция f имеет $n+1$–ю производную в окрестности нуля, то в форме Лагранжа $R(x) = f^{(n+1)}(\xi)x^{n+1}/(n+1)!$, причем число $\xi = \xi(x)$ находится между 0 и x. Принято число ξ записывать в виде $\xi = \theta x$, где $0 < \theta < 1$.

Формула названа по имени шотландца К.Маклорена (C. Mac Laurin) и опубликована им в 1742 г. Конкретизируем формулу Маклорена для некоторых элементарных функций.

1. $f(x) = e^x$. Функция имеет производные всех порядков, равные самой функции: $f^{(n)}(x) = f(x) = e^x$ для любого $x \in \mathsf{R}$ и для любого натурального числа n. Поэтому $f^{(n)}(0) = 1$ и в форме Лагранжа

$$e^x = \sum_{k=0}^{n}\frac{x^k}{k!} + e^{\xi}\frac{x^{n+1}}{(n+1)!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots + \frac{x^n}{n!} + e^{\theta x}\frac{x^{n+1}}{(n+1)!},$$

где $0 < \theta < 1$.

2. $f(x) = \sin x$. Функция имеет производные всех порядков, равные $f^{(n)}(x) = \sin(x + \pi n/2)$ и, следовательно,

$$f^{(n)}(0) = \begin{cases} 0, & \text{если } n = 2m, \text{ т.е. } n \text{ – четное,} \\ (-1)^m, & \text{если } n = 2m+1, \text{ т.е. } n \text{ – нечетное.} \end{cases}$$

Поэтому формула Маклорена порядка $2n+1$ для синуса имеет вид

$$\sin x = \sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{(2k+1)!} + R(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \ldots + (-1)^n\frac{x^{2n+1}}{(2n+1)!} + R(x),$$

где остаточный член в форме Лагранжа есть

$$R(x) = (-1)^{n+1}\frac{x^{2n+3}}{(2n+3)!}\cos(\theta x), \qquad 0 < \theta < 1.$$

3. $f(x) = \cos x$. Функция имеет производные всех порядков, равные $f^{(n)}(x) = \cos(x + \pi n/2)$ и, следовательно,

$$f^{(n)}(0) = \begin{cases} (-1)^m, & \text{если } n = 2m, \text{ т.е. } n \text{ -- четное,} \\ 0, & \text{если } n = 2m+1, \text{ т.е. } n \text{ -- нечетное.} \end{cases}$$

Поэтому формула Маклорена порядка $2n$ для косинуса имеет вид

$$\cos x = \sum_{k=0}^{n}(-1)^k\frac{x^{2k}}{(2k)!} + R(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \ldots + (-1)^n\frac{x^{2n}}{(2n)!} + R(x),$$

где остаточный член в форме Лагранжа есть

$$R(x) = (-1)^{n+1}\frac{x^{2n+2}}{(2n+2)!}\cos(\theta x), \qquad 0 < \theta < 1.$$

4. $f(x) = (1+x)^\mu$ при $x \in (-1, +\infty)$. Показатель μ — произвольное вещественное число. Для этой степенной функции $f'(x) = \mu(1+x)^{\mu-1}$, $f''(x) = \mu(\mu-1)(1+x)^{\mu-2}$, и по индукции несложно доказать, что

$$f^{(n)}(x) = \mu(\mu-1)\cdots(\mu-n+1)(1+x)^{\mu-n}.$$

При $x = 0$ последний сомножитель равен 1, и поэтому

$$(1+x)^\mu = 1+\mu x+\frac{\mu(\mu-1)}{2}x^2+\frac{\mu(\mu-1)(\mu-2)}{6}x^3+\ldots+\frac{\mu(\mu-1)\cdots(\mu-n+1)}{n!}x^n+R(x).$$

Остаточный член в форме Лагранжа имеет вид

$$R(x) = \frac{\mu(\mu-1)\cdots(\mu-n)}{(n+1)!}x^{n+1}(1+\theta x)^{\mu-n-1}, \qquad 0 < \theta < 1.$$

Заметим, что при $\mu = -1$ многочлен Маклорена становится геометрической прогрессией:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + (-1)^{n+1}x^{n+1}\frac{1}{(1+\theta x)^{n+2}}, \qquad 0 < \theta < 1.$$

5. $f(x) = \ln(1+x)$ при $x \in (-1, +\infty)$. После вычисления первой производной $f'(x) = (1+x)^{-1}$ расчет следующих производных сводится к предыдущему случаю. Получается, что $f^{(n)}(x) = (-1)^{n-1}(n-1)!/(1+x)^n$, и поэтому

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^{n-1}\frac{x^n}{n} + R(x),$$

и по Лагранжу

$$R(x) = (-1)^n\frac{x^{n+1}}{(n+1)(1+\theta x)^{n+1}}, \qquad 0 < \theta < 1.$$

Исследование функций и построение графиков

Определение *монотонной функции на множестве.*

Пусть функция f определена на множестве P. Функция f называется *возрастающей* (*строго возрастающей*) на P, если для любых $x, y \in P$ из $x < y$ следует $f(x) \leq f(y)$ (соответственно, $f(x) < f(y)$).

Аналогично, функция f называется *убывающей* (*строго убывающей*) на P, если для любых $x, y \in P$ из $x < y$ следует $f(x) \geq f(y)$ (соответственно, $f(x) > f(y)$).

Функция f называется *монотонной на* P, если она убывающая на P или возрастающая на P. Функция f называется *строго монотонной на* P, если она строго убывающая на P или строго возрастающая на P. □

Теорема. *Критерий монотонности функции на промежутке.*

Пусть функция f определена и дифференцируема на промежутке P. Тогда следующие утверждения равносильны:

1. f возрастает на P.
2. $f'(x) \geq 0$ для любого $x \in P$.

Аналогично, функция f убывает на P тогда и только тогда, когда $f'(x) \leq 0$ при всех $x \in P$.

Доказательство. Необходимость. Пусть выполнено утверждение 1: f возрастает на P. Докажем от противного, что $f'(x) \geq 0$ для любого $x \in P$. Пусть $a \in P$ и $f'(a) < 0$.

Функция f дифференцируема в точке a по условию. Следовательно, если $a+h \in P$, то
$$f(a+h) = f(a) + h(f'(a) + \beta(h)),$$
причем $\beta(h) \to 0$ при $h \to 0$. Выберем $\varepsilon = |f'(a)|/2$. По определению предела существует такое число $\delta > 0$, что $|\beta(h)| < \varepsilon = |f'(a)|/2$, если $|h| < \delta$ и $a + h \in P$.

Если точка a — внутренняя точка промежутка P или его нижняя граница, то $a + h \in P$ при достаточно малых $h > 0$. Если к тому же $|h| < \delta$, то
$$f(a+h) \leq f(a) + h(-|f'(a)| + \varepsilon) \leq f(a) - h|f'(a)|/2 < f(a).$$
Условия $a + h > a$ и $f(a+h) < f(a)$ противоречат предположению о возрастании функции f на промежутке P.

Если точка a совпадает с верхним концом промежутка P, то надо выбрать $h<0$ при дополнительных условиях $a + h \in P$ и $|h|<\delta$. Тогда условия $a + h < a$ и
$$f(a+h) \geq f(a) + |h|(|f'(a)| - \varepsilon) \geq f(a) + |h|\,|f'(a)|/2 > f(a)$$
противоречат предположению о возрастании функции f на промежутке P. Противоречие доказывает необходимость в утверждении теоремы (т.е. $1 \Rightarrow 2$).

Докажем достаточность. Пусть $f'(x) \geq 0$ при всех $x \in P$. Выберем произвольные точки $x, y \in P$, причем $x < y$. Функция f удовлетворяет на промежутке $[x, y]$ всем условиям теоремы Лагранжа о конечных приращениях. Поэтому существует такая точка $\xi \in (x, y)$, что
$$f(y) - f(x) = f'(\xi)(y - x).$$
Очевидно, что $f(y) - f(x) \geq 0$, так как $f'(\xi) \geq 0$ и $y - x > 0$. Таким образом, если $x < y$, то $f(x) \leq f(y)$ при $x, y \in P$, что означает возрастание функции f на P.

Критерий убывания доказывается аналогично. □

Замечание *о строгой монотонности.* Если $f'(x) > 0$ на P, то функция f строго возрастает на P, что следует из доказательства теоремы. Однако, это условие не является необходимым: существуют строго возрастающие функции, у которых производная может принимать и нулевые значения. Например, функция $f(x) = x^3$ на промежутке $[-1, 1]$ является строго возрастающей, но $f'(0) = 0$.

Определение *экстремума функции на промежутке.*

Пусть функция f определена на промежутке P и $a \in P$. Число a называется *точкой локального максимума f на P* (или *f имеет в точке a локальный максимум*), если

$$\exists \delta > 0 \ : \ f(x) \leq f(a) \quad \forall x \in P \cap (a - \delta, a + \delta).$$

Число a называется *точкой локального минимума f на P*, если выполнено последнее условие с заменой знака $\leq$ на $\geq$.

Условие для *строгого* или *точного локального минимума* получаются заменой неравенства $\leq$ на $<$:

$$\exists \delta > 0 \ : \ f(x) < f(a) \quad \forall x \in [P \cap (a - \delta, a + \delta)] \backslash \{a\},$$

и аналогично для *строгого локального максимума.*

Если функция f имеет в точке a локальный минимум или локальный максимум, то говорят, что f имеет в точке *a экстремум.* □

Наибольшие и наименьшие значения функции f на промежутке P называются также *глобальными максимумами* и *глобальными минимумами*, соответственно.

Глобальный максимум является также и локальным, поэтому точка глобального максимума находится среди точек экстремума. Перебором всех экстремумов часто удается определить глобальные максимумы и минимумы.

По теореме Ферма внутренняя точка промежутка P не может быть точкой экстремума, если функция f в ней дифференцируема и $f'(a) \neq 0$. Оставшиеся точки имеют специальное название, и среди них следует искать экстремумы.

Определение *точки, подозрительной на экстремум.*

Пусть функция f определена на промежутке P. Точка $a \in P$ называется *точкой, подозрительной на экстремум*, если выполнено одно из следующих условий:

1. Точка a — крайняя точка промежутка P.
2. Функция f не является дифференцируемой в точке a (в том числе имеет бесконечную производную).
3. Функция f дифференцируема в точке a и $f'(a) = 0$. □

Если на промежутке P лишь в конечном множестве точек функция f не является дифференцируемой, а уравнение $f'(x) = 0$ имеет лишь конечное множество решений, то количество точек, подозрительных на экстремум, конечно. Перебором этих точек можно найти наибольшее и наименьшее значения функции f на P.

В точке, подозрительной на экстремум, функция может не иметь экстремума. Например, функция $f(x) = x^3$ не имеет экстремумов, но точка $x = 0$ является подозрительной на экстремум, так как $f'(0) = 0$. Для нахождения настоящих экстремумов можно воспользоваться следующим критерием.

Теорема. *Достаточное условие экстремума по производной первого порядка.*

Пусть функция f задана и непрерывна на промежутке P, точка a — внутренняя в P и функция f дифференцируема на $P \backslash \{a\}$.

Тогда если функция f' меняет знак в точке a, то f имеет экстремум в точке a, а если не меняет знака — то не имеет.

Точнее, если существует такое число $\delta > 0$, что

$$\begin{cases} f'(x) \leq 0 & \text{при} \quad a - \delta < x < a, \ x \in P, \\ f'(x) \geq 0 & \text{при} \quad a < x < a + \delta, \ x \in P, \end{cases}$$

то в точке a функция f имеет локальный минимум.

Соответственно, если существует такое число $\delta > 0$, что

$$\begin{cases} f'(x) \geq 0 & \text{при} \quad a - \delta < x < a, \ x \in P, \\ f'(x) \leq 0 & \text{при} \quad a < x < a + \delta, \ x \in P, \end{cases}$$

то в точке a функция f имеет локальный максимум.

Если же знаки $f'(x)$ одинаковы и не равны нулю при $x < a$ и при $x > a$ (при условии $x \in P$, $|x - a| < \delta$), то число a не является точкой экстремума функции f.

Доказательство. Рассмотрим первый случай и докажем, что в точке a достигается локальный минимум. Достаточно доказать, что $f(a) \leq f(x)$, если $a - \delta < x < a + \delta$ и $x \in P$.

На интервале $P \cap (a, a + \delta)$ функция f имеет неотрицательную производную. По критерию монотонности функция f возрастает на этом интервале. Следовательно, $f(a) \leq f(x)$ при $x \in P \cap (a, a + \delta)$.

Аналогично, из условия $f'(x) \leq 0$ при $x \in P \cap (a - \delta, a)$ следует убывание функции f на $P \cap (a - \delta, a)$. Следовательно, $f(x) \geq f(a)$ при $x \in P$ и $a - \delta < x < a$.

Таким образом, $f(x) \leq f(a)$ при $x > a$ и при $x < a$ вблизи точки a. По определению это значит, что в точке a функция f имеет локальный минимум.

Доказательство локального максимума в точке a при выполнении второй группы условий в теореме проводится полностью аналогично.

Наконец, предположим, что $f'(x) \neq 0$ и $f'(x)$ не меняет знак на $[P \cap (a - \delta, a + \delta)] \setminus \{a\}$. Тогда на последнем множестве функция f строго монотонна и поэтому не может иметь экстремума. □

В соответствии с предыдущей теоремой для проверки существования экстремума в точке a, подозрительной на экстремум, достаточно исследовать поведение производной функции вблизи a. Как показывает следующая теорема, вместо исследования поведения функции f' иногда можно вычислить только одно значение второй производной f'' в точке a.

Теорема. *Достаточное условие экстремума по производной второго порядка.*

Пусть функция f задана на промежутке P, $a \in P$ и функция f имеет вторую производную в точке a. Пусть $f'(a) = 0$. Тогда

если $f''(a) > 0$, то f имеет строгий локальный минимум в точке a,

если $f''(a) < 0$, то f имеет строгий локальный максимум в точке a,

если $f''(a) = 0$, то может иметь или не иметь экстремума в точке a.

Доказательство. Воспользуемся формулой Тейлора второго порядка с оценкой остаточного члена в форме Пеано:

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + (x - a)^2 \gamma(x - a)$$

при $x \in P$, где $\gamma(h) \to 0$ при $h \to 0$.

В данном случае $f'(a) = 0$ и второе слагаемое исчезает. Предположим, что $f''(a) \neq 0$. По определению предела для числа $\varepsilon = |f'(a)|/2$ существует такое число $\delta > 0$, что неравенство $|\gamma(h)| < |f'(a)|/2$ выполнено для всех h, у которых $|h| < \delta$ и $a + h \in P$.

Пусть $x \in P$ и $|x - a| < \delta$. Число

$$f(x) - f(a) = (x-a)^2[f''(a) + \gamma(x-a)]$$

имеет тот же знак, что и число $f''(a)$, так как в квадратных скобках второе слагаемое меньше половины первого по абсолютной величине.

Если $f''(a) > 0$, то получается, что $f(x) > f(a)$ и, следовательно, функция f имеет в точке a строгий локальный минимум. Если же $f''(a) < 0$, то по аналогичным причинам число a есть точка строгого локального максимума. □

Замечание *о критериях высших порядков.* Если $f'(a) = 0$ и $f''(a) = 0$, то для проверки наличия экстремума в точке a можно воспользоваться производными высших порядков, при условии их существования. Из формулы Тейлора следует, что если самая младшая из отличных от нуля производных f в точке a имеет четный порядок, что в точке a есть строгий экстремум, определяемый знаком этой производной: минимум, если она положительна и максимум, если отрицательна. Если же порядок самой младшей отличной от нуля производной нечетный, то в точке a экстремума нет.

Выпуклые функции

Определение *линейной комбинации и выпуклой комбинации.*

Пусть $X = (x_k)_{k=1}^n$ — конечный набор чисел или векторов, а $\Lambda = (\lambda_k)_{k=1}^n$ — конечный набор чисел. Тогда вектор

$$\bar{x} = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = \sum_{k=1}^n \lambda_k x_k$$

называется *линейной комбинацией* элементов $(x_k)_{k=1}^n$, с коэффициентами $(\lambda_k)_{k=1}^n$.

Если при этом все коэффициенты λ_k неотрицательны и в сумме составляют единицу,

$$\lambda_k \geq 0, \quad 1 \leq k \leq n, \qquad \sum_{k=1}^n \lambda_k = 1,$$

то линейная комбинация с такими коэффициентами называется также *выпуклой комбинацией* элементов $(x_k)_{k=1}^n$, с коэффициентами $(\lambda_k)_{k=1}^n$. □

Рассмотрим в качестве примера линейные и выпуклые комбинации векторов на плоскости R^2, в трехмерной пространстве R^3 или в произвольном n–мерном пространстве R^n. Точку в пространстве будем отождествлять с радиусом–вектором, выходящим из начала координат с концом в этой точке.

В пространстве R^3 множество всех линейных комбинаций двух заданных неколлинеарных векторов x_1 и x_2 образует плоскость, натянутую на эти два вектора. Действительно, к двум векторам x_1 и x_2 всегда можно добавить третий вектор $y \in \mathsf{R}^3$, который перпендикулярен x_1 и x_2. Множество всех векторов $x \in \mathsf{R}^3$, которые перпендикулярны y, образуют плоскость, в которой каждый вектор может быть разложен в линейную комбинацию векторов x_1 и x_2. Эта плоскость и есть множество всех линейных комбинаций x_1 и x_2.

Рассмотрим множество всех выпуклых комбинаций векторов x_1 и x_2 на плоскости. Поскольку коэффициенты обладают свойством $\lambda_1 + \lambda_2 = 1$, то можно ввести более удобное обозначение: $\alpha = \lambda_1$, и тогда $1 - \alpha = \lambda_2$, причем $0 \le \alpha \le 1$. Выпуклую комбинацию векторов x_1 и x_2 можно представить следующим образом:

$$\bar{x} = \lambda_1 x_1 + \lambda_2 x_2 = \alpha x_1 + (1-\alpha)x_2 = x_2 + \alpha(x_1 - x_2), \qquad 0 \le \alpha \le 1.$$

Нетрудно видеть, что при изменении α от 0 до 1 вектор $\bar{x}$ пробегает отрезок с концами x_1 и x_2.

Аналогично в произвольном n–мерном пространстве R^n выпуклая комбинация двух точек лежит внутри отрезка, их соединяющего. Множество всех выпуклых комбинаций двух точек есть отрезок, соединяющий эти точки.

Множество всех выпуклых комбинаций трех точек x_1, x_2, x_3 в пространстве R^3 есть треугольник с вершинами x_1, x_2 и x_3.

Определение *выпуклой функции.*

Пусть функция f задана на промежутке P. Функция f называется *выпуклой* на P, если для любых точек $x_1, x_2 \in P$ и для любого числа $\alpha \in [0,1]$

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha)f(x_2).$$

Функция f называется *вогнутой* на P, если в последнем неравенстве знак противоположный:

$$f(\alpha x_1 + (1-\alpha)x_2) \ge \alpha f(x_1) + (1-\alpha)f(x_2) \qquad \forall x_1, x_2 \in P,\ 0 \le \alpha \le 1.$$

Если неравенства строгие при $0 < \alpha < 1$, то функция f называется *строго выпуклой на P*, или, соответственно, *строго вогнутой на P*. □

Замечание *о терминологии.* К сожалению, в учебниках на русском языке, изданных в 50–е годы и ранее, термины "выпуклая функция"и "вогнутая функция"имеют прямо противоположное значение. В распространенном сборнике задач Г.Н.Бермана, написанном в конце 40–х годов, функции, называемые в настоящее время выпуклыми, названы вогнутыми, и наоборот.

Более надежным является также широко употребляемый термин *выпуклой вниз функции* – для выпуклой функции, и *выпуклой вверх функции* – для вогнутой функции.

Определение выпуклости функции имеет вид неравенства, связывающего две выпуклые комбинации, а именно: значение функции f в выпуклой комбинации аргументов не превосходит соответствующей выпуклой комбинации значений функции f.

Выпуклые комбинации, стоящие в левой части неравенства в аргументе функции f, пробегают отрезок $[x_1, x_2]$. Выпуклые комбинации, стоящие в правой части неравенства, пробегают отрезок $[f(x_1), f(x_2)]$. На графике функции f отрезок, соединяющий точки с координатами $(x_1, f(x_1))$ и $(x_2, f(x_2))$, называется хордой графика функции f. Часть графика функции f, соединяющая те же точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$, называется дугой графика.

Определение выпуклой функции можно дать и в следующих геометрических терминах: функций называется выпуклой, если всякая хорда ее графика лежит над соответствующей дугой графика.

Теорема *о касательных к графику выпуклой функции.*

Пусть функция f задана и дифференцируема на промежутке P. Для того, чтобы она была выпуклой, необходимо и достаточно, чтобы ее график лежал выше любой касательной к графику.

(График Γ лежит выше касательной $\mathcal{L}$, если для любого $x_0 \in P$ точка (x_0, y_f) пересечения вертикальной прямой $x = x_0$ с графиком Γ лежит выше точки (x_0, y_d) пересечения этой же прямой с касательной $\mathcal{L}$, т.е. $y_f \geq y_d$.)

Доказательство приведено ниже, поскольку ссылается на следующий критерий выпуклости функции.

Теорема. *Критерий выпуклости дважды дифференцируемой функции.*

Пусть функция f определена и имеет вторую производную на промежутке P. Тогда

$$f \text{ выпукла на } P \iff \forall x \in P\ f''(x) \geq 0.$$

Если же $f''(x) > 0$ при всех $x \in P$, то функция f строго выпукла на P.

Доказательство. Докажем *необходимость* (т.е. $\Rightarrow$). Пусть функция f выпуклая на P. Требуется доказать, что $f''(x) \geq 0$ при всех $x \in P$.

Выберем произвольную внутреннюю точку x промежутка P. При малых $h > 0$ числа $x - h$ и $x + h$ принадлежат P. Запишем формулу Тейлора в форме Пеано:

$$\begin{aligned} f(x+h) &= f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + h^2\gamma(h), \\ f(x-h) &= f(x) - f'(x)h + \frac{f''(x)}{2}h^2 + h^2\gamma(-h), \end{aligned}$$

причем $\gamma(h) \to 0$ при $h \to 0$. Сложим эти уравнения:

$$f(x+h) + f(x-h) = 2f(x) + h^2(f''(x) + \frac{\gamma(h) + \gamma(-h)}{2}).$$

По условию выпуклости функции f:

$$f(x) = f\left(\frac{1}{2}(x+h) + \frac{1}{2}(x-h)\right) \leq \frac{1}{2}f(x+h) + \frac{1}{2}f(x-h).$$

Следовательно, $f(x+h) + f(x-h) - 2f(x) \geq 0$ и

$$f''(x) + \frac{\gamma(h) + \gamma(-h)}{2} \geq 0$$

при всех h, для которых $x + h \in P$ и $x - h \in P$. При $h \to 0$ по теореме о предельном переходе в неравенствах получим, что $f''(x) \geq 0$.

Неравенство $f''(x) \geq 0$ доказано для всех внутренних точек промежутка P. Следовательно, функция f' возрастает на P. По критерию монотонности для функции f' значения f'' в крайних точках промежутка P также неотрицательны.

Докажем *достаточность* (т.е. $\Leftarrow$). Пусть $f''(x) \geq 0$ при всех $x \in P$. Докажем выпуклость функции f.

Пусть $x_1 < x_2$ и $x_1, x_2 \in P$, $0 < \alpha < 1$ и $x_\alpha = \alpha x_1 + (1-\alpha)x_2$. Требуется доказать, что

$$f(x_\alpha) \leq \alpha f(x_1) + (1-\alpha)f(x_2).$$

Запишем формулу Тейлора первого порядка для функции f в точке x_α:

$$f(x) = f(x_\alpha) + f'(x_\alpha)(x - x_\alpha) + \frac{f''(\xi)}{2}(x - x_\alpha)^2,$$

где точка $\xi = \xi(x)$ расположена между x_α и x. Поскольку вторая производная не меньше нуля в любой точке ξ, а $(x - x_\alpha)^2 \geq 0$, то

$$f(x) \geq f(x_\alpha) + f'(x_\alpha)(x - x_\alpha) \qquad \forall x \in P.$$

Подставим в эту формулу $x = x_1$, затем $x = x_2$ и сложим с коэффициентами α и $1 - \alpha$:

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(x_\alpha) + f'(x_\alpha)(\alpha x_1 + (1-\alpha)x_2 - x_\alpha) = f(x_\alpha),$$

что совпадает с неравенством выпуклости.

Если неравенство $f''(x) > 0$ строгое, то в последнем переходе неравенство также получается строгим, что указывает на строгую выпуклость функции f. □

Следствие *о монотонности первой производной у выпуклой функции.* У выпуклой функции f производная f' возрастает.

Для дважды дифференцируемых функций это утверждение следует из критерия монотонности функции и неравенства $(f')' \geq 0$. Утверждение верно также и для выпуклых функций, не имеющих вторых производных. □

Доказательство теоремы о касательных к графику выпуклой функции.

Докажем *достаточность*. Пусть график функции f лежит выше любой касательной. Докажем, что функция f — выпуклая.

Пусть $x_1, x_2 \in P$, $0 < \alpha < 1$ и $x_\alpha = \alpha x_1 + (1-\alpha)x_2$ — выпуклая комбинация точек x_1 и x_2. Требуется доказать, что

$$f(x_\alpha) \leq \alpha f(x_1) + (1-\alpha) f(x_2).$$

Введем обозначения: $y_1 = f(x_1)$, $y_2 = f(x_2)$, $y_\alpha = f(x_\alpha)$. Проведем касательную $\mathcal{L}$ к графику функции f в точке (x_α, y_α). Согласно предположению, точки графика (x_1, y_1) и (x_2, y_2) лежат выше касательной. Но тогда и весь отрезок на плоскости, соединяющий эти точки, лежит выше касательной. Следовательно, выше касательной лежит и точка

$$\alpha(x_1, y_1) + (1-\alpha)(x_2, y_2) = (x_\alpha, \alpha y_1 + (1-\alpha) y_2).$$

Касательная $\mathcal{L}$ проходит через точку (x_α, y_α), следовательно,

$$y_\alpha \leq \alpha y_1 + (1-\alpha) y_2,$$

что совпадает с неравенством из определения выпуклости функции f.

Докажем *необходимость*. Пусть функция f выпуклая на P. Пусть $x_1 \in P$, $y_1 = f(x_1)$ и $\mathcal{L}$ — касательная к графику функции f в точке (x_1, y_1). Требуется доказать, что для любой точки $x_2 \in P$ точка (x_2, y_d) пересечения касательной $\mathcal{L}$ с вертикальной прямой $x = x_2$ лежит ниже точки графика (x_2, y_2), т.е. $y_d \leq y_2$, где $y_2 = f(x_2)$.

Касательная $\mathcal{L}$ является графиком аффинной функции

$$y = y_1 + f'(x_1)(x - x_1).$$

При $x = x_2$ получаем ординату $y = y_d$ точки пересечения прямых $\mathcal{L}$ и $x = x_2$. Поэтому неравенство $y_2 \geq y_d$ можно записать также в виде

$$y_2 \geq y_1 + f'(x_1)(x_2 - x_1)$$

или

$$y_2 - y_1 \geq f'(x_1)(x_2 - x_1).$$

По теореме Лагранжа о конечных приращениях существует такое число ξ между x_1 и x_2, что

$$y_2 - y_1 = f'(\xi)(x_2 - x_1).$$

По следствию о монотонности первой производной для выпуклой функции f ее производная f' является возрастающей функцией.

Если $x_1 < x_2$, то $\xi > x_1$, и следовательно, $f'(\xi) \geq f'(x_1)$. Поэтому

$$y_2 - y_1 = f'(\xi)(x_2 - x_1) \geq f'(x_1)(x_2 - x_1).$$

Если же $x_2 < x_1$, то $\xi < x_1$ и $f'(\xi) \leq f'(x_1)$. Снова

$$y_2 - y_1 = f'(\xi)(x_2 - x_1) \geq f'(x_1)(x_2 - x_1),$$

так как в этом случае $x_2 - x_1 < 0$.

В обоих случаях неравенство $y_2 - y_1 \geq f'(x_1)(x_2 - x_1)$ доказано. □

Теорема. *Неравенство Иенсена.* [J.L.Jensen, 1906; O.Hölder, 1889].

Пусть f — выпуклая функция на промежутке P. Тогда для любого набора точек $x_1, x_2, \ldots, x_n$ в P образ любой выпуклой комбинации не превосходит соответствующей выпуклой комбинации образов:

$$f\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} \lambda_k f(x_k), \qquad 0 < \lambda_k < 1, \quad 1 \leq k \leq n, \qquad \sum_{k=1}^{n} \lambda_k = 1.$$

Если функция f строго выпукла на P и среди чисел x_k есть различные, то последнее неравенство строгое.

Доказательство. Индукция по n. При $n = 2$ получается определение выпуклой функции.

Докажем индукционный шаг. Пусть утверждение доказано для $n - 1$. Определим числа $\lambda^* = \sum_{k=1}^{n-1} \lambda_k$, $\mu_k = \lambda_k/\lambda^*$ при $1 \leq k \leq n-1$. Тогда сумма неотрицательных чисел μ_k равна 1, и поэтому точка

$$x^* = \sum_{k=1}^{n-1} \mu_k x_k$$

является выпуклой комбинацией $n - 1$ точки $(x_k)_{k=1}^{n-1}$. По индукционному предположению,

$$f(x^*) = f\left(\sum_{k=1}^{n-1} \mu_k x_k\right) \leq \sum_{k=1}^{n-1} \mu_k f(x_k).$$

Определим $\alpha = \lambda_n$, тогда $1 - \alpha = \lambda^*$. В следующем преобразовании первое неравенство есть условие выпуклости функции f, а второе неравенство — индукционное предположение:

$$\begin{aligned} f\left(\sum_{k=1}^{n}\lambda_k x_k\right) &= f\left(\sum_{k=1}^{n-1}\lambda^*\mu_k x_k+\alpha x_n\right)=f((1-\alpha)x^*+\alpha x_n) \\ &\le (1-\alpha)f(x^*)+\alpha f(x_n)\le \lambda^*\sum_{k=1}^{n-1}\mu_k f(x_k)+\alpha f(x_n) \\ &= \sum_{k=1}^{n-1}\lambda_k f(x_k)+\lambda_n f(x_n)=\sum_{k=1}^{n}\lambda_k f(x_k), \end{aligned}$$

что доказывает индукционный шаг.

Если функция f строго выпуклая, то одно из двух неравенств в последнем переходе — строгое. □

Замечание *о вогнутых функциях.*

Если функция f вогнутая, то функция $(-f)$ выпуклая, и наоборот. Поэтому для вогнутых функций доказанные выше теоремы гарантируют следующие свойства:

1. $f''(x)\le 0\ \forall x\in P$, а если $f''<0$ на P, то f строго вогнута.
2. Любая хорда графика функции лежит ниже соответствующей дуги графика.
3. График лежит ниже любой своей касательной.
4. Неравенство Иенсена для вогнутых функций:

$$f\left(\sum_{k=1}^{n}\lambda_k x_k\right)\ge\sum_{k=1}^{n}\lambda_k f(x_k),\qquad 0<\lambda_k<1,\quad 1\le k\le n,\qquad \sum_{k=1}^{n}\lambda_k=1.$$

Пример 1. *Неравенство Коши.*

Рассмотрим функцию $f(x)=\log x$, заданную на интервале $(0,+\infty)$. Ее вторая производная $f''(x)=-1/x^2$ отрицательна на всей области определения. Поэтому функция f строго вогнута.

Выберем произвольные положительные числа x_1, x_2, …, x_n, и запишем неравенство Иенсена:

$$\log\left(\sum_{k=1}^{n}\lambda_k x_k\right)\ge\sum_{k=1}^{n}\lambda_k\log x_k,$$

где $0<\lambda_k<1$ при $k=1,2,\ldots,n$ и $\lambda_1+\lambda_2+\ldots+\lambda_n=1$. В неравенстве Иенсена знак $\ge$ можно заменить на $>$, если не все числа x_k совпадают.

Сумма логарифмов может быть преобразована к логарифму произведения, поэтому

$$x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n}\le\lambda_1x_1+\lambda_2x_2+\ldots+\lambda_nx_n.$$

Если выбрать $\lambda_k=\frac{1}{n}$ при $k=1,2,\ldots,n$, то получится неравенство между средним геометрическим и средним арифметическим:

$$\sqrt[n]{x_1x_2\cdots x_n}\le\frac{x_1+x_2+\cdots+x_n}{n},$$

причем неравенство строгое, если среди чисел x_k есть различные.

Пример 2. *Неравенство Минковского.* [H. Minkowski]

Рассмотрим функцию $f(x)=x^p$ при $p>1$, заданную на интервале $(0,+\infty)$. Вторая производная $f''(x)=p(p-1)x^{p-2}$ положительна на всей области определения, поэтому функция f строго выпуклая.

Выберем произвольные положительные числа x_1, x_2, …, x_n, и запишем неравенство Иенсена:

$$\left(\sum_{k=1}^{n} \lambda_k x_k\right)^p \le \sum_{k=1}^{n} \lambda_k x_k^p,$$

где $0 < \lambda_k < 1$ при $k = 1, 2, \ldots, n$ и $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$. В неравенстве Иенсена знак $\le$ можно заменить на $<$, если не все числа x_k совпадают.

Выберем произвольный набор положительных чисел $(\mu_k)_{k=1}^n$ и обозначим их сумму через $\bar{\mu}$. Тогда числа $\lambda_k = \mu_k/\bar{\mu}$, $1 \le k \le n$ обладают свойством $\lambda_1 + \cdots + \lambda_n = 1$. Поэтому

$$\frac{\left(\sum_{k=1}^n \mu_k x_k\right)^p}{\bar{\mu}^p} \le \frac{\sum_{k=1}^n \mu_k x_k^p}{\bar{\mu}},$$

что можно также преобразовать к неравенству

$$\sum_{k=1}^{n} \mu_k x_k \le \left(\sum_{k=1}^{n} \mu_k x_k^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} \mu_k\right)^{\frac{p-1}{p}} \qquad \forall \mu_k > 0, \ \ \forall x_k > 0.$$

Числа $p > 1$ и $q = \frac{p}{p-1}$ называются сопряженными, потому что $\frac{1}{p} + \frac{1}{q} = 1$.

Введем новые переменные вместо x_k и μ_k. Определим положительные числа a_k и b_k равенствами $\mu_k = b_k^q$, $x_k = a_k b_k^{1-q}$ при $1 \le k \le n$. После тождественных преобразований получится неравенство Минковского:

$$\sum_{k=1}^{n} a_k b_k \le \left(\sum_{k=1}^{n} a_k^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} b_k^q\right)^{\frac{1}{q}}, \qquad \frac{1}{p} + \frac{1}{q} = 1, \qquad a_k > 0, \ b_k > 0.$$

Пример 3. *Неравенство Коши–Буняковского.* [A.L.Caushy, 1821; интегральное обобщение доказал В.Я.Буняковский в 1859 г.; иногда это неравенство называют неравенством Шварца (H.A.Schwarz), опубликовавшего результат без ссылок на предшественников в 1884 г.]

Это частный случай неравенства Минковского. Если $p = 2$, то $q = 2$. Получается, что для любых положительных чисел a_1, a_2, …, a_n и b_1, b_2, …, b_n:

$$\sum_{k=1}^{n} a_k b_k \le \sqrt{\sum_{k=1}^{n} a_k^2} \ \sqrt{\sum_{k=1}^{n} b_k^2}.$$

Очевидно, что если среди чисел a_k или b_k есть отрицательные, то это неравенство тем более выполнено. Таким образом, оно справедливо при любых вещественных числах a_k и b_k.

Выражение в левой части неравенства интерпретируется как скалярное произведение векторов $a = (a_k)_{k=1}^n$ и $b = (b_k)_{k=1}^n$. Тогда в правой части стоит произведение длин этих векторов. Поэтому неравенство Коши–Буняковского часто формулируют в геометрических терминах: скалярное произведение векторов не превосходит произведения их длин.

Комплексные числа

Множеством *комплексных чисел* C называется множество всех упорядоченных пар вещественных чисел (a, b), для которых определены операции сложения и умножения:

$$
\begin{aligned}
(a_1, b_1) + (a_2, b_2) &= (a_1 + a_2,\ b_1 + b_2), \\
(a_1, b_1) \cdot (a_2, b_2) &= (a_1 a_2 - b_1 b_2,\ a_1 b_2 + a_2 b_1).
\end{aligned}
$$

Число a называется *вещественной частью* комплексного числа $z = (a, b)$, а число b — *мнимой частью*. Обозначение: $a = \operatorname{Re} z$, $b = \operatorname{Im} z$.

В подмножестве C, состоящем из чисел с нулевой мнимой частью, т.е. вида $z = (a, 0)$, сложение и умножение осуществляются так же, как и для вещественных чисел, являющихся их вещественными частями:

$$
\begin{aligned}
(a_1, 0) + (a_2, 0) &= (a_1 + a_2,\ 0), \\
(a_1, 0) \cdot (a_2, 0) &= (a_1 a_2,\ 0).
\end{aligned}
$$

Поэтому комплексное число $(a, 0)$ может быть отождествлено с вещественным числом a. В частности, комплексные числа (x, y) можно умножать на вещественные a, считая их комплексными числами $(a, 0)$:

$$a\,(x,\ y) = (a,\ 0)\,(x,\ y) = (ax,\ ay),$$

т.е. отдельно умножается вещественная и мнимая части.

Любое комплексное число можно представить в виде

$$(a,\ b) = (a,\ 0) + b\,(0,\ 1).$$

Комплексное число $(0,\ 1)$ называется *мнимой единицей* и обычно обозначается i (иногда j). В этих обозначениях произвольное комплексное число записывается в виде

$$(a,\ b) = a + bi.$$

Именно эта форма записи комплексных чисел является наиболее употребительной. Она называется *алгебраической записью* комплексного числа.

Мнимая единица i обладает следующим свойством: $i \cdot i = -1$, и поэтому это число называют также квадратным корнем из -1. Комплексные числа с нулевой вещественной частью, т.е. вида $z = bi$, называют *чисто мнимыми*.

Комплексное число z называется *комплексно сопряженным* по отношению к комплексному числу $c = (a,\ b)$, если $z = (a,\ -b)$. Стандартное обозначение: $z = \bar{c}$.

Очевидно, что второе комплексное сопряжение совпадает с исходным числом: $\bar{\bar{c}} = c$. Кроме того, комплексное число $c\,\bar{c} = (a^2 + b^2, 0)$ является вещественным.

Свойства комплексных чисел.

Для любых комплексных чисел z_1, z_2, z_3 выполнены следующие равенства:

1. $z_1 + z_2 = z_2 + z_1$ — коммутативность сложения;
2. $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$ — ассоциативность сложения;
3. $z_1 z_2 = z_2 z_1$ — коммутативность умножения;
4. $z_1(z_2 z_3) = (z_1 z_2) z_3$ — ассоциативность умножения;
5. $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ — дистрибутивность умножения относительно сложения.

Комплексно сопряженные числа после сложения или умножения остаются комплексно сопряженными:

6. $\bar{z}_1 + \bar{z}_2 = \overline{z_1 + z_2}$;
7. $\bar{z}_1\, \bar{z}_2 = \overline{z_1\, z_2}$.

Комплексные числа можно не только умножать, но и делить. Для того, чтобы вычислить отношение комплексных чисел $a + bi$ и $c + di \neq 0$, нужно числитель и знаменатель умножить на число, сопряженное к знаменателю:

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{(a+bi)(c-di)}{c^2+d^2}.$$

Число $c^2 + d^2$ — вещественное и положительное. Поэтому его можно обратить и вынести общим множителем:

$$\frac{a+bi}{c+di} = \frac{1}{c^2+d^2}(a+bi)(c-di).$$

Остается вычислить произведение комплексных чисел и разделить вещественную и мнимую части на $c^2 + d^2$.

Геометрическая интерпретация комплексных чисел.

Любая упорядоченная пара чисел однозначно определяет точку на плоскости, на которой фиксирована декартова система координат. Поэтому комплексному числу соответствует точка на плоскости. Множество комплексных чисел может быть отождествлено с плоскостью, на которой введены операции умножения и сложения точек. Поэтому множество C комплексных чисел называют также комплексной плоскостью.

На комплексной плоскости ось абсцисс содержит все вещественные числа, и поэтому называется *вещественной осью*. Ось ординат содержит все чисто мнимые числа, и поэтому называется *мнимой осью*.

На комплексной плоскости можно ввести полярную систему координат. Каждой точке $(x,\, y)$ ставится в соответствие длина радиуса–вектора $r = \sqrt{x^2+y^2}$ и полярный угол ϕ, отсчитываемый от оси абсцисс против часовой стрелки.

Таким образом, каждой комплексное число $z = x + iy$ может быть также задано числами r и ϕ. Число r называется *модулем* комплексного числа z и обозначается $|z|$. Число ϕ называется *аргументом* z.

Вещественная и мнимая части x, y, а также модуль r и аргумент ϕ комплексного числа связаны друг с другом уравнениями

$$\left\{\begin{array}{l} x = r\,\cos\phi, \\ y = r\,\sin\phi, \end{array}\right. \qquad \left\{\begin{array}{l} r = \sqrt{x^2+y^2}, \\ \sin\phi = \frac{y}{\sqrt{x^2+y^2}}, \\ \cos\phi = \frac{x}{\sqrt{x^2+y^2}}. \end{array}\right.$$

Полярный угол ϕ определяется комплексным числом неоднозначно: при добавлении $2\pi k$ точка плоскости не меняется. Множество всех чисел ϕ, удовлетворяющих последней системе уравнений, обозначается $\operatorname{Arg} z$. То значение аргумента, которое расположено в промежутке $[-\pi,\, \pi)$, называется *главным значением аргумента* z. Иногда вместо промежутка $[-\pi,\, \pi)$ для главного значения выбирают промежуток $[0, 2\pi)$. В любом случае

$$\operatorname{Arg} z = \{\arg z + 2\pi k \mid k = 0, \pm 1, \pm 2, \ldots\}.$$

Если в алгебраическую форму комплексного числа z вместо вещественной и мнимой частей x и y подставить их выражения через модуль и аргумент, то получится равенство

$$z = |z|\,(\cos\phi + i\sin\phi), \qquad (\text{где } \phi = \arg z)$$

которое называется *тригонометрической формой* комплексного числа z.

Произведение чисел, записанных в тригонометрической форме, можно преобразовать следующим образом. Пусть $|z_1| = r_1$, $\arg z_1 = \phi_1$, $|z_2| = r_2$, $\arg z_2 = \phi_2$. Тогда по правилам произведения комплексных чисел:

$$\begin{aligned} z_1\,z_2 &= r_1\,r_2\,(\cos\phi_1 + i\sin\phi_1)\,(\cos\phi_2 + i\sin\phi_2) \\ &= r_1\,r_2\,(\cos\phi_1\,\cos\phi_2 - \sin\phi_1\,\sin\phi_2) + i(\sin\phi_1\,\cos\phi_2 + \cos\phi_1\,\sin\phi_2)) \\ &= r_1\,r_2\,(\cos(\phi_1+\phi_2) + i\sin(\phi_1+\phi_2)). \end{aligned}$$

Полученное равенство может быть записано в более компактной форме:

$$|z_1\,z_2| = |z_1|\,|z_2|, \qquad \operatorname{Arg}(z_1\,z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2.$$

В произведении комплексных чисел модули перемножаются, а аргументы складываются.

Формулы Эйлера и Муавра.

Отображение от аргумента комплексного числа к самому числу обладает тем же свойством, что и показательная функция: сумма аргументов преобразуется в произведение значений. В связи с этим было введено следующее обозначение:

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

Из формул для произведения комплексных чисел в тригонометрической форме следует, что

$$e^{i\phi_1}e^{i\phi_2} = e^{i(\phi_1+\phi_2)},$$

в полном соответствии со свойствами показательной функции.

Подстановка определения $e^{i\phi}$ в тригонометрическую форму числа z дает

$$z = |z|e^{i\arg z},$$

что называется *показательной формой* комплексного числа z.

Из определения функции $e^{i\phi}$ следуют обратные формулы

$$\cos\phi = \frac{e^{i\phi} + e^{-i\phi}}{2}, \qquad \sin\phi = \frac{e^{i\phi} - e^{-i\phi}}{2i},$$

которые называются *формулами Эйлера.*

Если комплексное число возвести в натуральную степень n, то его модуль возведется в эту жу степень, а аргумент n раз сложится сам с собой. Если $z = re^{i\phi}$, то

$$z^n = r^n e^{in\phi} \qquad \text{или} \qquad z^n = r^n(\cos n\phi + i\sin n\phi).$$

Это *формулы Муавра.*

Показательная функция продолжается на все комплексные числа по формуле

$$e^{x+iy} = e^x e^{iy} = e^x(\cos y + i\sin y).$$

Эта функция сохраняет свое основное свойство

$$e^{z_1}e^{z_2} = e^{z_1+z_2} \qquad \forall z_1, z_2 \in \mathsf{C}.$$

Нетрудно видеть, что $|e^{x+iy}| = e^x$ и $\operatorname{Arg} e^{x+iy} = \{y + 2\pi k,\ k = 0, \pm 1, \pm 2, \ldots\}$.

Корни из комплексного числа.

Пусть z — комплексное число, а n — натуральное число. Комплексное число ζ называется корнем n–ой степени из z, если $\zeta^n = z$. Найдем способ вычисления корней из комплексных чисел.

Запишем числа z и ζ в показательной форме:

$$z = re^{i\phi}, \qquad \zeta = \rho e^{i\psi}.$$

Тогда условие $\zeta^n = z$ можно записать в виде

$$\rho^n = r, \qquad \{\phi + 2\pi m,\ m = 0, \pm 1, \ldots\} \supset \{n\psi + 2\pi l n,\ l = 0, \pm 1, \ldots\}.$$

Из первого равенства следует, что $\rho = \sqrt[n]{r}$, это положительное число (при $z \neq 0$). Из равенства аргументов следует, что различные (с точностью до $2\pi k$) значения величины ψ могут быть записаны в виде

$$\psi = \frac{\phi + 2\pi k}{n}, \qquad 0 \le k \le n - 1.$$

Таким образом, у любого комплексного числа $z \neq 0$ существует ровно n различных корней n–ой степени. Это

$$\zeta_k = \sqrt[n]{r} e^{\frac{\phi + 2\pi k}{n} i}, \qquad 0 \le k \le n - 1.$$

Логарифм и возведение в комплексную степень.

Комплексное число ζ называется логарифмом комплексного числа z, если $z = e^{\zeta}$. Пусть $\zeta = \xi + i\eta$ — алгебраическая форма ζ и $z = re^{i\phi}$ — показательная форма z. Здесь $r = |z|$, $\phi = \arg z$, $\xi = \operatorname{Re} \zeta$, $\eta = \operatorname{Im} \zeta$. Тогда

$$r\, e^{i\phi} = e^{\xi}\, e^{i\eta},$$

Поэтому $r = e^{\xi}$ и $\eta - \phi = 2\pi k$ при целом k.

В частности, при $k = 0$ получаем функцию

$$\ln z = \ln |z| + i \arg z,$$

которая называется *главным значением логарифма* z. Все остальные значения логарифма образуют множество

$$\operatorname{Ln} z = \{\ln z + 2\pi i k,\ k = 0, \pm 1, \pm 2, \ldots\} = \{\ln |z| + i(\arg z + 2\pi k),\ k = 0, \pm 1, \pm 2, \ldots\}.$$

Таким образом, функция "логарифм", как и функция "аргумент", имеет счетное число значений, и одно из них называется главным.

Функция $z_1^{z_2}$ для комлексных чисел z_1 и z_2 определяется при помощи замены на $e^{z_2 \operatorname{Ln} z_1}$. Получается следующее множество значений, в котором обычно не выделяют какого–либо главного значения:

$$z_1^{z_2} = \{e^{z_2(\ln |z_1| + i(\arg z_1 + 2\pi k))}, \quad k = 0, \pm 1, \pm 2, \ldots\}$$

Многочлены

Многочленом с вещественными коэффициентами (или *многочленом над полем вещественных чисел*) называется функция $P : \mathsf{R} \to \mathsf{R}$, определяемая правилом

$$P(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

где a_0, a_1, …, a_n — вещественные числа. *Степенью* многочлена $P(x)$ называется наибольшее натуральное число m, для которого $a_m \neq 0$. В частности, если $a_n = 0$, то степень P меньше n. По определению, степенью многочлена $P \equiv 0$ считается $-\infty$.

Степень обозначается $\deg P$ от английского слова degree. Многочлены иногда называют также полиномами от английского polynomial.

Числа a_0, a_1, …, a_n называются *коэффициентами многочлена*, и если $m = \deg P$, то число a_m называется *старшим коэффициентом.*

Многочлены с комплексными коэффициентами (или *над полем комплексных чисел*) отличаются от многочленов с вещественными коэффициентами только тем, что их коэффициенты могут быть комплексными (и, в частном случае, вещественными). Они могут считаться отображением $P : \mathsf{C} \to \mathsf{C}$, т.е. переменная x может принимать комплексные значения. В дальнейшем, если не оговорено особо, будем считать, что многочлены имеют комплексные коэффициенты.

Два многочлена равны тогда и только тогда, когда у них равны коэффициенты при одинаковых степенях. Многочлены можно складывать, умножать на числа и друг на друга. При этом будут получаться снова многочлены. Многочлены образуют алгебраическую структуру, называемую *кольцом.*

Рациональной функцией называется отношение двух многочленов:

$$R(x) = \frac{P(x)}{Q(x)}, \qquad x \in \mathsf{R} \backslash Q^{-1}(\{0\}),$$

где P и Q — многочлены, причем Q не есть тождественный ноль. Рациональная функция называется *правильной*, если $\deg P < \deg Q$. Если же $Q(x) = \mathrm{const} \neq 0$, то рациональная функция является также и многочленом.

Теорема *о делении многочленов.*

Для любых многочленов P и Q существуют и единственны многочлены S и r, такие, что

$$P(x) = S(x)Q(x) + r(x) \quad \text{ и } \quad \deg r < \deg Q.$$

Доказательство. Пусть $\deg P = n$. Докажем существование S и r индукцией по n. При $n = 0$ утверждение очевидно: если $\deg Q > 0$, то $S = 0$, $r = P$; если $\deg Q = 0$, то $r = 0$, $S = Q/P$.

Докажем индукционный переход. Пусть утверждение доказано для многочленов P степени n и пусть

$$P(x) = p_0 + p_1 x + \ldots + p_{n+1} x^{n+1}, \qquad Q(x) = p_0 + p_1 x + \ldots + p_m x^m,$$

причем $p_{n+1} \neq 0$ и $q_m \neq 0$. Если $n + 1 < m$, то можно выбрать $S = 0$ и $r = P$.

Пусть $n + 1 \geq m$. Тогда

$$P(x) = \frac{p_{n+1}}{q_m} x^{n+1-m} Q(x) + P_1(x), \qquad P_1(x) = P(x) - \frac{p_{n+1}}{q_m} x^{n+1-m} Q(x).$$

Многочлен $P_1(x)$ является разностью двух многочленов степени $n+1$, причем старшие коэффициенты этих многочленов равны и поэтому сокращаются. Следовательно, $\deg P_1 \le n$ и по индукционному предположению существуют такие многочлены S_1 и r, что $P_1 = S_1 Q + r$. Подставим это выражение в предыдущую формулу:

$$P(x) = \left(\frac{p_{n+1}}{q_m} x^{n+1-m} + S_1(x)\right) Q(x) + r(x) = S(x)Q(x) + r(x), \qquad \deg r < \deg Q,$$

что доказывает индукционный переход.

Докажем единственность многочленов S и r. Пусть таких пар две:

$$P = S_1 Q + r_1 = S_2 Q + r_2, \qquad \deg r_1 < \deg Q, \quad \deg r_2 < \deg Q.$$

Тогда

$$(S_1 - S_2)Q = r_2 - r_1.$$

Если $S_1 \neq S_2$, то степень многочлена в левой части этого равенства не меньше $\deg Q$, а в правой части — строго меньше, чем $\deg Q$. Противоречие доказывает единственность многочленов S и r. □

Утверждение теоремы о делении многочленов часто записывают в виде равенства рациональных функций:

$$\frac{P(x)}{Q(x)} = S(x) + \frac{r(x)}{Q(x)}, \qquad \deg r < \deg Q.$$

Согласно теореме, любая рациональная функция может быть представлена в виде суммы многочлена и правильной рациональной функции, причем такое представление единственно.

Из доказательства теоремы следует, что если коэффициенты многочленов P и Q вещественные, то коэффициенты частного S и остатка r тоже вещественные.

Теорема Безу. [E.Bézout, 1779]

Остаток от деления многочлена $P(x)$ на $x - a$ равен значению $P(a)$.

Доказательство. Пусть a — произвольное комплексное число. По теореме о делении многочленов существуют такие многочлены S и r, что

$$P(x) = (x-a)S(x) + r(x), \qquad \deg r < 1.$$

Но многочлен r нулевой степени есть константа: $r(x) = r$. Последнее уравнение переходит в числовое тождество при подстановке вместо x любого числа, комплексного или вещественного. Подставим $x = a$. Тогда получится $P(a) = r$. □

Следствие *о линейном множителе.*

Если комплексное число a есть корень многочлена P, то этот многочлен разлагается на множители, один из которых есть $x - a$. □

Теорема. *Основная теорема высшей алгебры.*

Любой многочлен над полем комплексных чисел положительной степени имеет хотя бы один комплексный корень. □

Без доказательства.

Теорема *о разложении многочлена на множители над полем комплексных чисел.*

Пусть P — многочлен степени n со старшим коэффициентом p_n. Тогда существуют такие различные комплексные числа c_1, c_2, …, c_m и соответствующие им натуральные числа k_1, k_2, …, k_m, что

$$P(x) = p_n(x - c_1)^{k_1}(x - c_2)^{k_2} \cdots (x - c_m)^{k_m},$$

причем $k_1 + k_2 + \ldots + k_m = n$.

Доказательство. По основной теореме высшей алгебры многочлен P имеет корень. Обозначим его a_1. По следствию из теоремы Безу многочлен P можно разложить на множители: $P(x) = (x - a_1)P_1(x)$. При этом степень P_1 будет равна $n - 1$, а старший коэффициент останется равным p_n.

Далее, многочлен P_1 снова разлагаем на множители в соответствии с основной теоремой высшей алгебры и следствием из теоремы Безу: $P_1(x) = (x - a_2)P_2(x)$. Так же поступаем и с многочленом $P_2(x)$ и т. д. Поскольку после каждой итерации степень многочлена понижается на 1, то через n итераций она станет равной нулю. Это значит, что многочлен $P_n(x)$ равен своему старшемуц коэффициенту, а это p_n. Следовательно,

$$P(x) = p_n(x - a_1)(x - a_2) \cdots (x - a_n).$$

Если среди чисел a_1, a_2, …, a_n встречаются равные, то соответствующие сомножители сгруппируем вместо. Получится формула из утверждения теоремы. $\square$

Теорема *о разложении многочлена на множители с вещественными коэффициентами.*

Пусть P — многочлен с вещественными коэффициентами. Тогда существуют такие наборы вещественных чисел $(\alpha_i)_{i=1}^m$, $(p_j, q_j)_{j=1}^n$ и наборы натуральных чисел $(k_i)_{i=1}^m$, $(l_j)_{j=1}^n$, что

$$P(x) = d\prod_{i=1}^{m}(x - \alpha_i)^{k_i} \prod_{j=1}^{n}(x^2 + p_j x + q_j)^{l_j},$$

где d — старший коэффициент P и все дискриминанты квадратных трехчленов отрицательны: $D_j = p_j^2 - 4q_j < 0$, $1 \le j \le n$. Такое разложение на множители единственно. $\square$

Доказательство. По теореме о разложении многочлена на множители над полем комплексных чисел существуют такие различные комплексные числа c_1, c_2, …, c_N и соответствующие им натуральные числа k_1, k_2, …, k_N, что

$$P(x) = d(x - c_1)^{k_1}(x - c_2)^{k_2} \cdots (x - c_m)^{k_N},$$

причем $k_1 + k_2 + \ldots + k_m = n$. Среди чисел c_1, …, c_N, могут встретиться вещественные. Тогда эти числа обозначим α_1, …, α_m. Множители $x - c_j$, у которых мнимая часть c_j не равна нулю, преобразуются далее так, чтобы получились квадратичные сомножители с вещественными коэффициентами.

Докажем сначала, что если комплексное число c является корнем многочлена $P(x) = a_0 + a_1x + \ldots + a_nx^n$ с вещественными коэффициентами, то и комплексно–сопряженное число $\bar{c}$ также является корнем:

$$\begin{aligned} P(\bar{c}) &= a_0 + a_1\bar{c} + a_2\bar{c}^2 + \ldots + a_n\bar{c}^n = a_0 + a_1\bar{c} + a_2\overline{c^2} + \ldots + a_n\overline{c^n} \\ &= \overline{a_0} + \overline{a_1 c} + \overline{a_2 c^2} + \ldots + \overline{a_n c^n} = \overline{a_0 + a_1 c + a_1 c^2 + \ldots + a_n c^n} \\ &= \overline{P(c)} = \bar{0} = 0. \end{aligned}$$

Таким образом, если $c_k = a_k + ib_k$ с $b_k \neq 0$, то среди чисел $c_1, \ldots, c_N$, есть сопряженное число $\overline{c_k} = a_k - ib_k$. Преобразуем произведение соответствующих множителей:

$$(x - c_k)(x - \overline{c_k}) = (x - a_k - ib_k)(x - a_k + ib_k) = (x - a_k)^2 - (ib_k)^2 = x^2 - 2a_k x + a_k^2 + b_k^2.$$

Получился квадратный трехчлен, коэффициенты которого вещественные. Кроме того, его дискриминант отрицательный.

Перемножим пары скобок с сопряженными числами c_k столько раз, сколько это возможно сделать. Тогда многочлен $P(x)$ окажется разложенным на множители, часть из которых имеет вещественные коэффициенты, а часть — комплексные с ненулевой мнимой частью. Произведение первой группы множителей обозначим $Q(x)$, а второй — $S(x)$.

Тогда $P(x) = Q(x)S(x)$. Многочлены $P(x)$ и $Q(x)$ имеют вещественные коэффициенты. По теореме о делении многочленов многочлен $S(x)$ является частным от деления P на Q. Следовательно, его коэффициенты также вещественные.

Однако S может иметь только комплексные корни с ненулевой мнимой частью, а сопряженных пар комплексных корней нет, так как все множителями с сопряженными корнями были перемножены и вошли в многочлен Q. Следовательно, многочлен Q вообще не имеет корней и равен константе. Это значит, что многочлен P уже разложен на множители указанного вида с вещественными коэффициентами. □

Теорема *о разложении правильной рациональной функции на простейшие.*
Пусть $R(x) = P(x)/Q(x)$ — правильная рациональная функция и

$$Q(x) = \prod_{i=1}^{m} (x - \alpha_i)^{k_i} \prod_{j=1}^{n} (x^2 + p_j x + q_j)^{l_j},$$

причем $p_j^2 < 4q_j$ при $1 \leq j \leq n$.

Тогда существуют и единственны наборы вещественных чисел $(A_{i,s})$, $1 \leq s \leq k_i$, $1 \leq i \leq m$ и $(B_{i,s}, C_{i,s})$, $1 \leq s \leq l_j$, $1 \leq j \leq n$, такие, что

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{m} \sum_{s=1}^{k_i} \frac{A_{i,s}}{(x - \alpha_i)^s} + \sum_{j=1}^{n} \sum_{s=1}^{l_j} \frac{B_{j,s} x + C_{j,s}}{(x^2 + p_j x + q_j)^s}.$$

Каждая из дробей в правой части разложения называется *простейшей рациональной функцией.*

Доказательство. Выберем произвольный вещественный корень α_i многочлена Q и обозначим через $Q_1(x)$ многочлен, получаемый перемножением всех множителей в разложении Q, кроме $(x - \alpha_i)^{k_i}$, т.е. $Q(x) = (x - \alpha_i)^{k_i} Q_1(x)$. Докажем, что существует и единственны вещественное число A и многочлен P_1, такие, что

$$\frac{P(x)}{Q(x)} = \frac{A}{(x - \alpha_i)^{k_i}} + \frac{P_1(x)}{(x - \alpha_i)^{k_i - 1} Q_1(x)}.$$

Другими словами, в знаменателе степень множителя $x - \alpha_i$ можно понизить, выделив одну из простейших дробей, стоящую в утверждении теоремы.

Определим $A = P(\alpha_i)/Q_1(\alpha_i)$. Тогда у многочлена $P(x) - AQ_1(x)$ число α_i является корнем. По следствию из теоремы Безу этот многочлен делится на $(x - \alpha_i)$, т.е. существует такой многочлен $P_1(x)$ с вещественными коэффициентами, что

$$P(x) - AQ_1(x) = (x - \alpha_i)P_1(x).$$

Разделим это равенство на $Q(x) = (x-\alpha_i)^{k_i} Q_1(x)$ и перенесем простейшую дробь в правую часть:

$$\frac{P(x)}{Q(x)} = \frac{A}{(x-\alpha_i)^{k_i}} + \frac{P_1(x)}{(x-\alpha_i)^{k_i-1}Q_1(x)},$$

что и требовалось доказать. Обозначим $A_{i,k_i} = A$.

Рациональную функцию $\frac{P_1(x)}{(x-\alpha_i)^{k_i-1}Q_1(x)}$ можно преобразовать по той же схеме, если $k_i - 1 > 0$:

$$\frac{P_1(x)}{(x-\alpha_i)^{k_i-1}Q_1(x)} = \frac{A_{i,k_i-2}}{(x-\alpha_i)^{k_i-1}} + \frac{P_2(x)}{(x-\alpha_i)^{k_i-2}Q_1(x)},$$

и т. д., пока множитель $(x-\alpha_i)$ вообще не исчезнет в знаменателе.

Поступая таким образом, разложим рациональную функцию $P(x)/Q(x)$ в сумму простейших дробей и правильной рациональной функции $\widetilde{P}(x)/\widetilde{Q}(x)$, у которой знаменатель содержит лишь произведения множителей $(x^2+p_jx+q_j)^{l_j}$.

Пусть $\widetilde{Q}(x) = (x^2+p_jx+q_j)^{l_j}\widetilde{Q}_1(x)$. Докажем, что существуют такие вещественные числа B, C и многочлен $\widetilde{P}_1$, что

$$\frac{\widetilde{P}(x)}{\widetilde{Q}(x)} = \frac{Bx+C}{(x^2+p_jx+q_j)^{l_j}} + \frac{\widetilde{P}_1(x)}{(x^2+p_jx+q_j)^{l_j-1}\widetilde{Q}_1(x)}.$$

Умножим это равенство на знаменатель последней дроби:

$$\frac{\widetilde{P}(x)}{x^2+p_jx+q_j} = (Bx+C)\frac{\widetilde{Q}_1(x)}{x^2+p_jx+q_j} + \widetilde{P}_1(x).$$

По теореме о делении многочленов существуют такие вещественные числа r_0, r_1, ρ_0, ρ_1 и многочлены с вещественными коэффициентами $S_0(x)$ и $S_1(x)$, что

$$\frac{\widetilde{P}(x)}{x^2+p_jx+q_j} = S_0(x) + \frac{r_0+r_1x}{x^2+p_jx+q_j}, \qquad \frac{\widetilde{Q}_1(x)}{x^2+p_jx+q_j} = S_1(x) + \frac{\rho_0+\rho_1x}{x^2+p_jx+q_j}.$$

При этом числа ρ_0 и ρ_1 не могут быть равны нулю одновременно, так как многочлен $\widetilde{Q}_1(x)$ не содержит множителем $x^2+p_jx+q_j$.

Тогда уравнение относительно многочленов записывается в виде

$$\frac{r_0+r_1x}{x^2+p_jx+q_j} = (Bx+C)\frac{\rho_0+\rho_1x}{x^2+p_jx+q_j} + \widetilde{P}_1(x) - S_0(x) + (Bx+C)S_1(x).$$

Это уравнение легко решается, если заметить, что

$$(Bx+C)\frac{\rho_0+\rho_1x}{x^2+p_jx+q_j} = B\rho_1 + \frac{\rho_0C - q_j\rho_1B + (\rho_1C+\rho_0B-p_j\rho_1B)}{x^2+p_jx+q_j}.$$

Приравняем коэффициенты числителей в двух оставшихся правильных дробях:

$$\begin{cases} \rho_0C - q_j\rho_1B & = & r_0, \\ \rho_1C + (\rho_0 - p_j\rho_1)B & = & r_1. \end{cases}$$

Это система уравнений относительно чисел B и C. Определитель матрицы системы равен $\Delta = \rho_0^2 - p_j\rho_0\rho_1 + q_j\rho_1^2$. Если $\rho_1 = 0$, то $\Delta = \rho_0^2 > 0$. Если $\rho_1 \neq 0$, то $\Delta = (\rho^2+p_j\rho+q_j)\rho_1^2$, где $\rho = -\rho_0/\rho_1$. Снова получаем, что $\Delta > 0$, так как дискриминант многочлена $x^2+p_jx+q_j$ отрицательный и, следовательно, $\rho^2+p_j\rho+q_j > 0$.

Определитель Δ системы линейных уравнений не равен нулю, следовательно, система имеет единственное решение (B, C). Остается определить

$$\widetilde{P}_1(x) = S_0(x) - (Bx + C)S_1(x) - B\rho_1.$$

Таким образом, в правильной рациональной функции $\widetilde{P}(x)/\widetilde{Q}(x)$ удается выделить простейшую дробь так, чтобы у оставшейся рациональной функции в знаменателе множитель $x^2 + p_j x + q_j$ имел меньшую кратность. Как и в случае множителей вида $(x-\alpha_i)^{k_i}$, можно последовательно выделять простейшие дроби, понижая степень сомножителя $x^2 + p_j x + q_j$ в знаменателе, до тех пор, пока он совсем не исчезнет. Так можно поступить с каждым множителем $x^2 + p_j x + q_j$ при $j = 1, 2, \dots, n$. Оставшаяся правильная рациональная функция не будет иметь корней знаменателя, и поэтому равна нулю. $\square$

Числа $A_{i,s}$, $B_{j,s}$ и $C_{j,s}$ называются *неопределенными коэффициентами*, поскольку сначала они записываются в виде неизвестных констант, которые затем находятся из системы уравнений.

Практическое вычисление неопределенных коэффициентов состоит в следующем. Сумма дробей в правой части приводится к общему знаменателю. При этом в знаменателе получится в точности многочлен $Q(x)$. В числителе получится многочлен степени не выше $\deg Q - 1$, коэффициенты которого будут линейными комбинациями неопределенных коэффициентов. Поскольку числитель должен быть равен $P(x)$ при всех x, то можно приравнять коэффициенты при одинаковых степенях x.

Метод доказательства теоремы о разложении рациональной функции на простейшие содержит еще один алгоритм расчета неопределенных коэффициентов. Этот способ обычно наиболее эффективен, если все корни знаменателя рациональной функции — простые.

Интеграл

Определение *первообразной функции на промежутке.*

Пусть функции f и F заданы на промежутке P, а F дифференцируема на P. Функция F называется *первообразной функцией* функции f, если $F' = f$. □

Операция вычисления первообразной — обратная к операции дифференцирования. Поэтому таблицы производных, переписанные справа налево, являются таблицами первообразных.

Замечание *о линейности первообразных.*

Пусть функции f и g заданы на промежутке P, а F и G — соотвествующие первообразные. Пусть $\alpha, \beta \in \mathsf{R}$. Тогда функция $\alpha F + \beta G$ является первообразной функцией функции $\alpha f + \beta g$.

Доказательство. По свойствам линейности производной непосредственно проверяется, что

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g.$$

Теорема *о множестве первообразных.*

Пусть F — первообразная функции f на промежутке P. Тогда для любой постоянной функции $C(x) = \mathrm{const}$ функция $F + C$ также является первообразной функции f. И наоборот, для любой первообразной F_1 функции f существует постоянная функция $C_1(x) = \mathrm{const}$, такая, что $F_1 = F + C_1$.

(Первообразная функция единственна с точностью до постоянного слагаемого.)

Доказательство. Первое утверждение очевидно: $(F + C)' = F' + C' = f$.

Докажем второе утверждение. Пусть $g(x) = F(x) - F_1(x)$, $x \in P$. Тогда $g'(x) = F'(x) - F_1'(x) = 0$ при всех $x \in P$. Докажем, что $g(x) = \mathrm{const}$. Выберем произвольное число $a \in P$. Тогда по теореме Лагранжа о конечных приращениях для любого $x \in P$ существует такое число $\xi \in P$, что $g(x) - g(a) = g'(\xi)(x - a)$. Но $g'(\xi) = 0$, поэтому $g(x) = g(a) = \mathrm{const}$. □

Определение *неопределенного интеграла.*

Множество всех первообразных функций функции f называется *неопределенным интегралом* функции f.

В соответствии с предыдущей теоремой это множество может быть записано как $\{F + C, C \in \mathsf{R}\}$, где F — некоторая первообразная функции f. □

В отличие от производных для вычисления первообразных нет правила преобразования первообразной сложной функции к первообразным составляющих ее функций. Более того, первообразная от функции, выраженной через элементарные функции, может вообще не разлагаться в композицию элементарных функций.

Под вычислением неопределенного интеграла обычно понимают задачу нахождения первообразной, которая записывается через элементарные функции. Многочисленные способы решения этой задачи называют обычно техникой интегрирования.

Интегральные суммы

Пусть на замкнутом отрезке $[a, b]$ выбран набор из $n+1$ точки $\tau = (x_k)_{k=0}^{n}$, причем $x_0 = a$, $x_n = b$ и $x_0 < x_1 < \ldots < x_n$. Такой набор τ называется *дроблением* отрезка $[a, b]$.

Поскольку точки дробления x_k упорядочены, то $\sum_{k=0}^{n-1}(x_{k+1}-x_k) = b-a$. Наибольшее из расстояний между соседними точками дробления

$$\lambda(\tau) = \max_{0\le k\le n-1}(x_{k+1}-x_k)$$

называется *мелкостью* дробления.

Пусть на каждом из промежутков между точками дробления выбрано число $\xi_k \in [x_k, x_{k+1}]$, $0 \le k \le n-1$. Тогда набор этих чисел $\xi = (\xi_k)_{k=0}^{n-1}$ называется *оснащением* дробления τ.

Определение *интегральной суммы Римана.* [B.Riemann, 1853]

Пусть функция f определена на отрезке $[a,b]$. Пусть $\tau = (x_k)_{k=0}^{n}$ — некоторое дробление $[a,b]$ и $\xi = (\xi_k)_{k=0}^{n-1}$ — некоторое оснащение этого дробления. Тогда число

$$\sigma(f,\tau,\xi) = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1}-x_k)$$

называется *суммой Римана* (или *интегральной суммой*), соответствующей функции f, дроблению τ и оснащению ξ. □

Дробление вместе с оснащением порождают кусочно–постоянную функцию, равную $f(\xi_k)$ на промежутке $[x_k, x_{k+1})$. Интегральная сумма равна площади ступенчатой области, расположенной между графиком этой функции и осью абсцисс, причем площадь берется со знаком минус, если график расположен ниже оси абсцисс.

Теорема *о пределе интегральных сумм Римана.*

Пусть функция f определена и непрерывна на отрезке $[a,b]$. Тогда существует предел интегральных сумм Римана при стремлении к нулю мелкости дробления промежутка $[a,b]$. Этот предел не зависит от выбора дробления и от выбора оснащения.

Другими словами, существует такое число $I \in \mathsf{R}$, что

$$\begin{aligned}\forall\varepsilon>0\quad \exists\delta>0\ :\ &\forall \text{ дробления } \tau \text{ отрезка } [a,b] \text{ и } \forall \text{ оснащения } \xi \text{ дробления } \tau\\ &\text{если мелкость } \lambda(\tau)<\delta, \quad \text{то } |\sigma(f,\tau,\xi)-I|<\varepsilon. \square\end{aligned}$$

Определение *интеграла функции по замкнутому промежутку.*

Пусть функция f непрерывна на отрезке $[a,b]$. Тогда предел интегральных сумм I при стремлении к нулю мелкости дробления промежутка $[a,b]$ называется *интегралом* (или *определенным интегралом*) функции f по промежутку $[a,b]$ (или от a до b). Обозначение:

$$\int_a^b f \qquad \text{или} \qquad \int_{[a,b]} f \qquad \text{или} \qquad \int_a^b f(x)\,dx.$$

Функция f называется подынтегральной функцией, а промежуток $[a,b]$ — множеством интегрирования. Вместо отрезка $[a,b]$ пишут также открытый интервал (a,b) или полуоткрытый промежуток. Число a называется нижним пределом интегрирования, а число b — верхним пределом интегрирования. □

Доказательство теоремы о пределе интегральных сумм Римана.

Сначала докажем сходимость в себе интегральных сумм:

$$\begin{aligned}\forall\varepsilon>0\qquad \exists\delta>0\ :\ &\forall \text{ дроблений } \tau_1,\tau_2 \text{ если мелкости } \lambda(\tau_1)<\delta,\ \ \lambda(\tau_2)<\delta,\\ &\text{то } |\sigma(f,\tau_1,\xi_1)-\sigma(f,\tau_2,\xi_2)|<\varepsilon \quad \forall \text{ оснащений } \xi_1,\xi_2 \text{ дроблений } \tau_1,\tau_2.\end{aligned}$$

Сходящиеся в себе последовательности имеют предел по признаку Коши. Это же верно и для сходящихся в себе обобщенных последовательностей, какими являются интегральные суммы при стремлении к нулю мелкости дробления.

Фиксируем число $\varepsilon > 0$ и найдем $\delta > 0$, для которого выполнено указанное выше условие.

Непрерывная на отрезке $[a, b]$ функция f по теореме Кантора равномерно непрерывна. По определению равномерной непрерывности для положительного числа $\varepsilon_1 = \varepsilon/(2(b-a))$ существует такое число $\delta > 0$, что $|f(x_1) - f(x_2)| < \varepsilon_1$ для любой пары точек $x_1, x_2 \in [a, b]$, отстоящих друг от друга менее, чем на δ, т.е. $|x_1 - x_2| < \delta$. Фиксируем это число δ.

Рассмотрим два дробления: $\tau_1 = (x_i)_{i=0}^m$ и $\tau_2 = (y_k)_{k=0}^n$ отрезка $[a, b]$, у которых мелкости меньше, чем δ, т.е. $|\lambda(\tau_1)| < \delta$ и $|\lambda(\tau_2)| < \delta$. Предположим сначала, что дробление τ_1 является частью дробления τ_2, т.е. $x_i = y_{k_i}$, $0 \le i \le m$ при некоторой последовательности номеров $(k_i)_{i=0}^m$.

Выберем два произвольных оснащения $\xi = (\xi_i)_{i=0}^m$ дробления τ_1 и $\eta = (\eta_k)_{k=0}^n$ дробления τ_2. Выберем номер i между 0 и $m-1$ и рассмотрим промежуток $[x_i, x_{i+1}] = [y_{k_i}, y_{k_{i+1}}]$.

Поскольку $|x_{i+1} - x_i| \le \lambda(\tau_1) < \delta$, то на этом промежутке модуль разности значений функции f не может быть больше, чем ε_1. В частности,

$$|f(\xi_i) - f(\eta_k)| < \epsilon_1, \qquad k = k_i, k_{i+1}, \ldots, k_{i+1} - 1,$$

так как числа ξ_i и η_k при $k_i \le k < k_{i+1}$ лежат на промежутке $[x_i, x_{i+1}]$. Отметим также, что

$$x_{i+1} - x_i = y_{k_{i+1}} - y_{k_i} = \sum_{k=k_i}^{k_{i+1}-1} (y_{k+1} - y_k).$$

Полученные неравенства верны при всех $i = 0, 1, \ldots, m-1$, что позволяет непосредственно преобразовать разность интегральных сумм:

$$\begin{aligned}
|\sigma(f, \tau_1, \xi) - \sigma(f, \tau_2, \eta)| &= \left| \sum_{i=0}^{m-1} f(\xi_i)(x_{i+1} - x_i) - \sum_{k=0}^{n-1} f(\eta_k)(y_{k+1} - y_k) \right| \\
&= \left| \sum_{i=0}^{m-1} \left(f(\xi_i)(x_{i+1} - x_i) - \sum_{k=k_i}^{k_{i+1}-1} f(\eta_k)(y_{k+1} - y_k) \right) \right| \\
&= \left| \sum_{i=0}^{m-1} \sum_{k=k_i}^{k_{i+1}-1} (f(\xi_i) - f(\eta_k))(y_{k+1} - y_k) \right| \\
&< \varepsilon_1 \sum_{i=0}^{m-1} \sum_{k=k_i}^{k_{i+1}-1} (y_{k+1} - y_k) = \varepsilon_1(b-a) = \frac{\varepsilon}{2}.
\end{aligned}$$

Таким образом, требуемое неравенство $|\sigma(f, \tau_1, \xi) - \sigma(f, \tau_2, \eta)| < \varepsilon$ в этом случае доказано.

Рассмотрим теперь общий случай: дробление τ_1 не является частью дробления τ_2. Тогда мы можем сформировать дробление τ_3, которое содержит все точки обоих дроблений. Выберем произвольные оснащения ξ_1, ξ_2, ξ_3 соответствующих дроблений.

По построению дробление τ_1 является частью дробления τ_3. В соответствии с доказанным,

$$|\sigma(f, \tau_1, \xi_1) - \sigma(f, \tau_3, \xi_3)| < \frac{\varepsilon}{2}.$$

Аналогично, дробление τ_2 является частью дробления τ_3 и

$$|\sigma(f,\tau_2,\xi_2)-\sigma(f,\tau_3,\xi_3)|<\frac{\varepsilon}{2}.$$

Следовательно,

$$|\sigma(f,\tau_1,\xi_1)-\sigma(f,\tau_2,\xi_2)|\le|\sigma(f,\tau_1,\xi_1)-\sigma(f,\tau_3,\xi_3)|+|\sigma(f,\tau_3,\xi_3)-\sigma(f,\tau_2,\xi_2)|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

и сходимость в себе интегральных сумм доказана. Существование предела следует из критерия Коши. □

Замечание *о направлении интегрирования.*

Если $a>b$, то интеграл функции f от a до b определяется равенством

$$\int_a^b f=-\int_b^a f.$$

Если $a=b$, то интеграл функции f от a до b по определению равен нулю. □

Замечание *об интегрируемых функциях.*

Пределы интегральных сумм Римана при стремлении к нулю мелкости дроблений промежутка интегрирования существуют не только для непрерывных функций f. Кусочно–непрерывные и многие другие функции обладают таким же свойством, и для них существует интеграл.

Функции на промежутке P, для которых интегральные суммы Римана имеют предел, называются интегрируемыми по Риману. Несколько более широкий класс функций, интегрируемых по Лебегу [H.Lebesgue, 1902], имеет стандартное обозначение $L(P)$.

Свойства интеграла

1. Аддитивность интеграла.

Пусть функция f непрерывна на отрезке $[a,b]$ и $a<c<b$. Докажем, что тогда

$$\int_a^b f=\int_a^c f+\int_c^b f.$$

К каждому разбиению τ отрезка $[a,b]$ добавим точку c, если ее нет среди точек разбиения τ. Новое разбиение можно разделить на две части: одна составляет разбиение отрезка $[a,c]$, а вторая — отрезка $[c,b]$. Предел суммы равен сумме пределов при $\lambda(\tau)\to 0$.

Заметим, что доказанное равенство с учетом определения интеграла с обратным направлением интегрирования может быть записано в формах

$$\int_a^c f=\int_a^b f+\int_b^c f,\qquad \int_c^b f=\int_c^a f+\int_a^b f.$$

Таким образом, исходное условие $a<c<b$ несущественно, и равенство верно при любом взаимном расположении чисел a, b и c.

Замечание *об определении интеграла по множеству, составленному из промежутков.*

При фиксированной функции f интеграл можно рассматривать как отображение, переводящее множество интегрирования в значение интеграла. Это отображение задано на множестве промежутков вещественной оси и принимает значение в R.

Произвольное отображение Φ, переводящее подмножества вещественной оси в R называется *аддитивным*, если $\Phi(A \cup B) = \Phi(A) + \Phi(B)$ при условии, что A и B не пересекаются (т.е. $A \cap B = \emptyset$), и отображение Φ определено на A, B и $A \cup B$.

Доказанная формула

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f.$$

означает, что интеграл является аддитивной функцией от промежутка интегрирования. На основании этого свойства легко определить интеграл на объединении произвольных непересекающихся промежутков.

Если множество A есть объединение произвольных непересекающихся промежутков P_1, P_2, ..., P_n, то по определению

$$\int_A f = \int_{P_1} f + \int_{P_2} f + \ldots + \int_{P_n} f.$$

Замечание *об определении интеграла от кусочно–непрерывной функции.*

Функция f называется кусочно–непрерывной на промежутке $[a, b]$, если лишь в конечном множестве точек на $[a, b]$ функция f имеет разрывы, и это разрывы первого рода. Обозначим эти точки $c_1 < c_2 < \cdots < c_n$. Добавим сюда точки $c_0 = a$ и $c_{n+1} = b$.

Сужение функции f на каждый из промежутков (c_k, c_{k+1}) является непрерывной функцией. Поэтому с учетом свойства аддитивности можно определить интеграл от кусочно–непрерывной функции f по промежутку $[a, b]$ равенством

$$\int_a^b f = \sum_{k=0}^{n} \int_{c_k}^{c_{k+1}} f.$$

2. Линейность интеграла.

Пусть функции f и g интегрируемы по Риману на множестве $A \subset \mathsf{R}$ и α, β — вещественные числа. Тогда

$$\int_A (\alpha f + \beta g) = \alpha \int_A f + \beta \int_A g.$$

Для доказательства этого равенства рассмотрим интегральную сумму, соответствующую функции $\alpha f + \beta g$, разбиению $\tau = (x_k)_{k=0}^{n}$ и оснащению $\xi = (\xi_k)_{k=0}^{n-1}$:

$$\begin{aligned}
\sigma(\alpha f + \beta g, \tau, \xi) &= \sum_{k=0}^{n-1} (\alpha f(\xi_k) + \beta g(\xi_k))(x_{k+1} - x_k) \\
&= \alpha \sum_{k=0}^{n-1} f(\xi_k) + \beta \sum_{k=0}^{n-1} g(\xi_k) = \alpha\sigma(f, \tau, \xi) + \beta\sigma(g, \tau, \xi).
\end{aligned}$$

Предел интегральных сумм $\sigma(f, \tau, \xi)$ и $\sigma(g, \tau, \xi)$ существуют и равны интегралам функций f и g. Остается применить теорему о пределе суммы и произведения.

3. Монотонность интеграла.

Пусть функции f и g интегрируемые по Риману на множестве $A \subset \mathsf{R}$ и $f(x) \leq g(x)$ при всех $x \in A$. Тогда

$$\int_A f \leq \int_A g.$$

Это свойство называют также теоремой об интегрировании неравенств.

В частности, интеграл от функции тождественно равной нулю, равен нулю. Следовательно,

$$\text{если } \forall x \in A \;\; f(x) \geq 0, \qquad \text{то} \quad \int_A f \geq 0,$$

т.е. интеграл от неотрицательной функции неотрицателен.

Последнее свойство очевидно следует из неотрицательности любой интегральной суммы, в которой складываются неотрицательные слагаемые.

Докажем, что неравенства вида $f \leq g$ можно интегрировать. Если $f(x) \leq g(x)$ при $x \in A$, то функция $h(x) = g(x) - f(x)$ неотрицательна и интегрируема по Риману на множестве A. Поэтому интеграл от h неотрицателен:

$$\int_A h = \int_A (g - f) = \int_A g - \int_A f \geq 0,$$

что равносильно неравенству $\int_A g \geq \int_A f$.

Теорема о среднем

Интеграл от постоянной функции $f(x) = c = \mathrm{const}$ по промежутку $[a, b]$ легко вычисляется. Для любого дробления $\tau = (x_k)_{k=0}^{n}$ и оснащения $\xi = (\xi_k)_{k=0}^{n-1}$ интегральная сумма

$$\sigma(f, \tau, \xi) = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k) = c \sum_{k=0}^{n-1} (x_{k+1} - x_k) = c(b - a)$$

есть величина постоянная. Естественно, она имеет предел:

$$\int_a^b f = c(b - a).$$

Таким образом, интеграл от постоянной функции f по промежутку $[a, b]$ равен произведению значения c этой функции на длину $b - a$ промежутка.

Если функция f не является постоянной, но непрерывна на $[a, b]$, то последнее равенство все же имеет место, но в нем число c есть значение функции f в некоторой "средней"точке промежутка $[a, b]$.

Теорема *о среднем.*

Пусть функция f определена и непрерывна на промежутке $[a, b]$. Тогда существует такое число $\xi \in [a, b]$, что

$$\int_a^b f = f(\xi)(b - a).$$

□

Геометрическая интерпретация теоремы о среднем состоит в следующем: площадь под графиком функции f равна площади прямоугольника с тем же основанием, у которого горизонтальная сторона пересекает график функции f в некоторой точке $\xi \in [a, b]$.

Теорема о среднем является частным случаем следующего утверждения.

Теорема *о среднем, обобщенная.*

Пусть функции f и ϕ определены на промежутке $[a, b]$. Пусть функция f непрерывна на $[a, b]$, а функция ϕ интегрируема и принимает значения только одного знака на $[a, b]$.

Тогда существует такое число $\xi \in [a, b]$, что

$$\int_a^b (f\phi) = f(\xi) \int_a^b \phi.$$

Доказательство. Непрерывная функция f на отрезке $[a, b]$ принимает свои наибольшее и наименьшее значения: $m \le f(x) \le M$ $\forall x \in [a, b]$. Функция ϕ либо неотрицательная, либо неположительная на $[a, b]$ по условию. Для определенности, предположим, что $\phi \ge 0$ на $[a, b]$. Тогда

$$m\phi(x) \le f(x)\phi(x) \le M\phi(x) \qquad \forall x \in [a, b].$$

Неравенства можно интегрировать, а постоянный множитель выносить из–под знака интеграла. Следовательно,

$$m \int_a^b \phi \le \int_a^b (f\phi) \le M \int_a^b \phi.$$

Если $\int_a^b \phi = 0$, то в утверждении теоремы оба интеграла равны нулю и теорема доказана. Поэтому предположим, что интеграл ненулевой, и в силу свойства монотонности интеграла $\int_a^b \phi > 0$.

Разделим последние неравенства на положительное число:

$$m \le \frac{\int_a^b (f\phi)}{\int_a^b \phi} \le M.$$

По теореме Больцано–Коши об образе отрезка при непрерывном отображении все точки промежутка $[m, M]$ являются образами некоторых точек промежутка $[a, b]$ при отображении f. Здесь используется предположение о непрерывности функции f, без которого теорема неверна.

Поскольку число

$$\frac{\int_a^b (f\phi)}{\int_a^b \phi}$$

принадлежит промежутку $[m, M]$, то оно является образом некоторой точки $\xi \in [a, b]$, т.е. оно равно $f(\xi)$. После умножения на знаменатель получается формула из утверждения теоремы. □

Формула Ньютона–Лейбница

Пусть $a \in \mathsf{R}$ и b — либо вещественное число, большее a, либо $+\infty$. Если функция f определена и непрерывна на промежутке $[a, b)$, то для любого $x \in [a, b)$ сужение функции f на отрезок $[a, x]$ непрерывно. Поэтому можно определить отображение

$$F(x) = \int_a^x f, \qquad x \in [a, b).$$

Оно называется интегралом с переменным верхним пределом.

Теорема *об интеграле с переменным верхним пределом.*

Пусть функция f непрерывна на промежутке $[a,b)$ и F — интеграл функции f с переменным верхним пределом на $[a,b)$. Тогда F есть первообразная функция функции f.

Доказательство. Выберем произвольное число $x \in [a,b)$ и докажем, что функция F дифференцируема в точке x и $F'(x) = f(x)$.

Пусть $h > 0$ и $x + h < b$. Применим теорему о среднем к приращению функции F в точке x:

$$F(x+h) - F(x) = \int_a^{x+h} f - \int_a^x f = \int_x^{x+h} f = f(\xi)h,$$

где $\xi = \xi(h) \in (x, x+h)$ — средняя точка. Если $h \to 0$, то $\xi(h) \to x$, так как $x < \xi(h) < x + h$. По условию, функция f непрерывна в точке x, и по теореме о непрерывности сложной функции: $f(\xi(h)) \to f(x)$ при $h \to 0$ справа.

Если $h < 0$ и $a < x + h$, то тоже по теореме о среднем $F(x+h) - F(x) = hf(\xi(h))$, причем $x + h < \xi(h) < x$ и $f(\xi(h)) \to f(x)$ при $h \to 0$ слева.

Таким образом,

$$\frac{F(x+h) - F(x)}{h} = f(\xi(h)) \to f(x) \qquad \text{при } \ h \to 0,$$

что и означает дифференцируемость функции f в точке x и $F'(x) = f(x)$. $\square$

В соответствии с теоремой об интеграле функции f с переменным верхним пределом этот интеграл есть первообразная функции f. При этом нижний предел интегрирования может быть любым, лишь бы функция f оставалась непрерывной на рассматриваемом интервале.

По аналогии с интегралом с переменным верхним пределом можно определить интеграл с переменным нижним пределом. Для него

$$\left(\int_x^b f \right)' = -f(x).$$

Теорема. *Формула Ньютона–Лейбница.* [I.Newton, G.Leibniz, сер. XVII в.]

Пусть функция f непрерывна на отрезке $[a,b]$ и F — некоторая первообразная функции f. Тогда

$$\int_a^b f = F(b) - F(a).$$

Доказательство. Функция

$$F_1(x) = \int_a^x f$$

является первообразной функции f на промежутке $[a,b]$ по теореме об интеграле с переменным верхним пределом. В точке a верхний и нижний пределы интегрирования совпадают. Поэтому $F_1(a) = 0$ и

$$\int_a^b f = F_1(b) - F_1(a).$$

Пусть F — произвольная первообразная функции f. По теореме о множестве первообразных функция $F - F_1 = C$ есть константа. Следовательно, $F(b) - F(a) = F_1(b) - F_1(a)$ и равно интегралу функции f по $[a, b)$. □

Разность значений первообразной функции $F(x)$ в точке b и в точке a обычно обозначают символом *подстановки от a до b*:

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

Несобственные интегралы

Определение интеграла как предел интегральных сумм Римана было введено только для непрерывных на отрезке функций. Эти функции ограничены на отрезке. Если же функция, заданная и непрерывная на интервале (a, b) неограничена на нем, то она не может быть продолжена в граничных точках с сохранением непрерывности. Однако, интеграл от такой неограниченной функции иногда можно определить, и он в этом случае называется *несобственным.*

Интеграл по бесконечному промежутку можно определить как предел интегралов по замкнутым ограниченным промежуткам. Такой интеграл также называется *несобственным.*

Интеграл от неограниченной функции по ограниченному промежутку. Пусть функция f определена и непрерывна на полуоткрытом промежутке $[a, b)$ и неограничена вблизи точки b. Тогда *несобственным интегралом* по промежутку $[a, b)$ называется предел слева

$$\int_a^b f = \lim_{b_1 \to b-0} \int_a^{b_1} f.$$

Если предел существует, то говорят, что несобственный интеграл *сходится*, а если предела не существует или он бесконечный — то несобственный интеграл *расходится.* Говорят, что несобственный интеграл имеет *особенность* в точке b, где вычисляется предел.

Пусть F — первообразная функция функции f на промежутке $[a, b)$. Тогда по формуле Ньютона–Лейбница

$$\int_a^b f = \lim_{b_1 \to b-0} (F(b_1) - F(a)).$$

Очевидно, что добавка $F(a)$ не влияет на существование и конечность предела при $b_1 \to b$ слева. Поэтому сходимость интеграла не зависит от точки a. Она определяется свойствами функции f (и ее первообразной F) вблизи точки b. Вместо сходимости несобственного интеграла по промежутку $[a, b)$ иногда говорят о сходимости этого интеграла в точке b, подчеркивая независимость этого свойства от точки a.

Если функция f задана на промежутке $[a, b)$ и ограничена вблизи точки b, то предел F в точке b всегда существует, и такой интеграл можно определить как предел сумм Римана. Поэтому он не называется несобственным.

Аналогично определяется несобственный интеграл с особенностью на левом конце промежутка. Пусть функция f определена и непрерывна на промежутке $(a, b]$

и неограничена вблизи точки a. Тогда *несобственным интегралом* по промежутку $(a, b]$ называется предел справа

$$\int_a^b f = \lim_{a_1 \to a+0} \int_{a_1}^b f.$$

Несобственный интеграл *сходится*, если существует конечный предел, и *расходится*, если предел не существует или бесконечен.

Несобственный интеграл можно определить и по открытому интервалу (a, b), если функция f неограничена вблизи точек a и b. Тогда это двойной предел отдельно в точке a справа и в точке b слева.

Пример *несобственного интеграла от неограниченной степенной функции.*

Пусть $f(x) = \frac{1}{x^\alpha}$, $\alpha > 0$. Функция неограничена на промежутке $(0, a]$ для любого $a > 0$. Вычислим несобственный интеграл по промежутку $(0, a]$. Если $\alpha \neq 1$, то

$$\int_0^a \frac{1}{x^\alpha}\,dx = \lim_{\varepsilon \to 0+0} \int_\varepsilon^a \frac{1}{x^\alpha}\,dx = \lim_{\varepsilon \to 0+0} \frac{1}{1-\alpha} x^{1-\alpha}\Big|_\varepsilon^a = \lim_{\varepsilon \to 0+0} \frac{1}{1-\alpha}\left(a^{1-\alpha} - \varepsilon^{1-\alpha}\right).$$

Очевидно, что конечный предел существует, если $\alpha < 1$ и предел бесконечный, если $\alpha > 1$. При $\alpha = 1$ первообразной является функция $\log x$, и ее предел в нуле бесконечен.

Отсюда следует вывод: *интеграл от функции $x^{-\alpha}$ по промежутку $(0, a]$ сходится, если $\alpha < 1$ и расходится, если $\alpha \geq 1$.*

Теорема *сравнения о сходимости несобственных интегралов.*

Пусть функции f и g определены и непрерывны на промежутке $[a, b)$ и $0 \leq f(x) \leq g(x)$ при всех $x \in [a, b)$. Тогда если несобственный интеграл $\int_a^b g(x)\,dx$ сходится, то и несобственный интеграл $\int_a^b f(x)\,dx$ также сходится.

Доказательство. Для любого числа $x \in (a, b)$ интегралы

$$F(x) = \int_a^x f, \qquad G(x) = \int_a^x g$$

определены, так как функции f и g непрерывны на $[a, x]$. Эти интегралы являются первообразными функций f и g, соответственно. Первые производные $F'(x) = f(x)$ и $G'(x) = g(x)$ неотрицательны по условию. По теореме о критерии монотонности отсюда следует, что функции F и G возрастают на $[a, b)$. По теореме о пределе монотонной функции эти функции всегда имеют пределы при $x \to b$ справа, которые могут быть конечными или бесконечными. Однако, по условию, несобственный интеграл функции g сходится, т.е. существует конечный предел

$$\lim_{x \to b-0} G(x) = \lim_{x \to b-0} \int_a^x g = \int_a^b g.$$

По свойству монотонности интеграла: $0 \leq F(x) \leq G(x)$ при всех $x \in [a, b)$. Следовательно, функция F ограничена сверху числом $\int_a^b g$ и поэтому имеет конечный предел:

$$\exists \lim_{x \to b-0} F(x) = \lim_{x \to b-0} \int_a^x f = \int_a^b f \in \mathsf{R},$$

т.е. несобственный интеграл функции f по $[a, b)$ сходится. □

Следствие *о главной части несобственного интеграла.*

Если функции f и h непрерывны и неотрицательны на промежутке $[a, b)$, функция f неограничена вблизи точки b, а функция h ограничена и отделена от нуля вблизи точки b, то несобственные интегралы

$$\int_a^b f(x)\,dx \qquad \text{и} \qquad \int_a^b f(x)h(x)\,dx$$

сходятся или расходятся одновременно.

Доказательство. По условию, существуют такие числа $\delta > 0$ и $M > m > 0$, что $m \le h(x) \le M$ при всех $x \in [b-\delta, b)$. Определим $g_1(x) = mf(x)$ и $g_2(x) = Mf(x)$. Тогда $g_1(x) \le f(x)h(x) \le g_2(x)$, и для анализа сходимости несобственных интегралов этих функций по промежутку $[b-\delta, b)$ можно применить теорему сравнения. Остается вынести постоянные множители m и M из интегралов от g_1 и g_2 и учесть независимость свойства сходимости от нижнего предела интегрирования. □

В обозначениях следствия при анализе сходимости несобственного интеграла $\int_a^b f(x)h(x)\,dx$ достаточно ограничиться изучением несобственного интеграла от функции f. Поэтому в произведении $f(x)h(x)$ главной частью является функция $f(x)$, от которой зависит сходимость или расходимость интеграла.

Например, если положительная функция $h(x)$ непрерывна в окрестности точки $x = 0$, то несобственный интеграл от функции $h(x)/x^\alpha$ сходится в точке 0, если $\alpha < -1$ и расходится, если $\alpha \ge 1$.

Интеграл по бесконечному промежутку.

Пусть $a \in \mathsf{R}$, функция f задана на луче $[a, +\infty)$ и интегрируема на любом конечном промежутке $[a, b]$ при $b > a$. Тогда *несобственным интегралом по промежутку* $[a, +\infty)$ называется предел

$$\int_a^{+\infty} f = \lim_{b\to+\infty} \int_a^b f.$$

Говорят, что интеграл *сходится*, если этот предел существует и конечен, и *расходится* в противном случае.

Аналогично, интеграл по промежутку $(-\infty, b]$ от функции, интегрируемой на каждом конечном промежутке $[a, b]$ при $a < b$, определяется как предел

$$\int_{-\infty}^b f = \lim_{a\to-\infty} \int_a^b f.$$

Интеграл по промежутку $(-\infty, +\infty)$, т.е. по всей числовой оси R, определяется как двойной предел интегралов по $[a, b]$ при $a \to -\infty$ и при $b \to +\infty$.

Для определения сходимости или расходимости несобственного интеграла по бесконечному промежутку интегрирования можно применить теорему сравнения, аналогичную той, которая сформулирована для интеграла от неограниченных функций.

Нетрудно доказать, что несобственный интеграл $\int_a^{+\infty} x^{-\alpha}\,dx$ сходится, если $\alpha > 1$ и расходится, если $\alpha \le 1$.

Техника интегрирования

Первообразная функция существует у любой непрерывной на промежутке функции. В частности, любая функция, составленная из элементарных функций (являющаяся их композицией), имеет первообразную функцию на любом промежутке, где она определена. Однако эта первообразная функция совсем не обязательно может быть представлена в виде композиции элементарных функций. Например, у функций $f(x) = \frac{\sin x}{x}$ или $f(x) = e^{x^2}$ первообразная не может быть выражена через элементарные функции.

Представление первообразной функции в виде композиции элементарных функций весьма желательно, так как они легко вычисляются и их свойства хорошо известны. Поэтому нахождение представления первообразной через элементарные функции называется вычислением интеграла. "Взять интеграл"означает представить его через элементарные функции. Методы преобразования неопределенного интеграла, ведущие к его представлению через элементарные функции, называются техникой интегрирования.

Техника интегрирования основана на двух основных приемах: применение таблицы интегралов и замена переменной (или подстановка), а также на вспомогательном приеме, называемом интегрированием по частям.

Операция интегрирования является обратной к операции дифференцирования. Поэтому таблица производных, переписанная справа налево, окажется таблицей неопределенных интегралов.

Таблица интегралов

1. $\displaystyle\int e^x\,dx = e^x + C, \qquad \int a^x\,dx = \frac{1}{\ln a}a^x + C, \quad (a > 0, a \neq 1).$
2. $\displaystyle\int x^\alpha\,dx = \frac{1}{\alpha+1}x^{\alpha+1} + C, \quad (\alpha \neq -1), \qquad \int \frac{dx}{x} = \ln|x| + C.$
3. $\displaystyle\int \sin x\,dx = -\cos x + C.$
4. $\displaystyle\int \cos x\,dx = \sin x + C.$
5. $\displaystyle\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C.$
6. $\displaystyle\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C.$
7. $\displaystyle\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C.$
8. $\displaystyle\int \frac{dx}{1+x^2} = \operatorname{arctg} x + C.$

Замена переменной.

Пусть P — промежуток, функция g дифференцируема на P, функция f непрерывна на $g(P)$ и F — первообразная функции f. Тогда по правилам дифференцирования сложной функции существует производная функции $(F \circ g)$ и

$$[F(g(x))]' = f(g(x))\,g'(x) \qquad \forall x \in P.$$

Последнее равенство можно записать и в виде интегралов:

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

На этой формуле основан прием интегрирования, называемый заменой переменной в неопределенном интеграле.

Выполним вспомогательное преобразование. Дифференциал функции g определяется как функция $dg = g'(x)\,dx$. Применим эту формулу, сделав формальную замену в подынтегральном выражении:

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

В правой части переменная x встречается не иначе, как под знаком $g(x)$. Поэтому можно выполнить формальную подстановку: вместо $g(x)$ подставить новую переменную: $y = g(x)$. Получится

$$\int f(g(x))g'(x)\,dx = \int f(y)\,dy.$$

Предположим, что первообразную функцию функции f удалось вычислить, т.е. найдено выражение функции F через элементарные функции. Тогда

$$\int f(g(x))g'(x)\,dx = \int f(y)\,dy = F(y) + C = F(g(x)) + C,$$

т.е. получается формула интегрирования производной сложной функции.

Пример. Замена переменной $y = \sin x$ позволяет свести к табличному следующий интеграл:

$$\int \cos^2 x\, \sin x\,dx = \int \cos^2 x\,d(\cos x) = \int y^2\,dy = \frac{1}{3}y^3 + C = \cos^3 x + C.$$

Другой способ замены переменной связан с преобразованием интеграла $\int f(y)\,dy$ к виду $\int f(g(x))g'(x)\,dx$, если последний интеграл по каким–либо причинам вычислять проще.

Пример. Для вычисления следующего интеграла применяется замена переменной $x = \sqrt{y+1}$. При этом $y = x^2 - 1$ и $dy = 2x\,dx$:

$$\begin{aligned} \int y\sqrt{y+1}\,dy &= \int (x^2-1)x\,2x\,dx = 2\int (x^4 - x^2)\,dx = \frac{2}{5}x^5 - \frac{2}{3}x^3 + C \\ &= \frac{2}{5}\sqrt{(y+1)^5} - \frac{2}{3}\sqrt{(y+1)^3} + C. \end{aligned}$$

Интегрирование по частям.

Пусть функции f и g непрерывны на промежутке P, функции F и G — их первообразные. Тогда по правилам дифференцирования произведения

$$(F(x)G(x))' = f(x)G(x) + F(x)g(x), \qquad x \in P.$$

После тождественного преобразования:

$$F(x)g(x) = (F(x)G(x))' - f(x)G(x), \qquad x \in P.$$

Запишем это равенство в терминах неопределенных интегралов:

$$\int F(x)g(x)\,dx = F(x)G(x) - \int f(x)G(x)\,dx.$$

Способ преобразования интеграла $\int Fg$, стоящего в левой части данного равенства, называется *интегрированием по частям.*

Формулу интегрирования по частям записывают и в других формах, например,

$$\int F(x)\,dG(x) = F(x)G(x) - \int G(x)\,dF(x).$$

Такую формулу проще запомнить: внеинтегральный член есть произведение дифференцируемых функций F и G, а под интегралами стоят эти же функции, причем в одном случае под дифференциалом оказывается F, а в другом — G.

Однако, наибольшую значимость при решении задач имеет формула

$$\int F(x)G'(x)\,dx = F(x)G(x) - \int F'(x)G(x)\,dx.$$

В исходном интеграле, стоящем в левой части равенства, выделяются два сомножителя: $F(x)$ и $G'(x)$. В результате преобразования, называемого интегрированием по частям, получается интеграл, в котором первый сомножитель — $F(x)$ дифференцируется, а второй — $G(x)$ — интегрируется.

Поскольку интегрируется только часть подынтегральной функции, то прием получил название интегрирования по частям. Чаще всего это прием бывает полезным там, где производная $F'(x)$ оказывается проще, чем исходная функция $F(x)$.

Пример. В таблице интегралов отсутствуют правила интегрирования функций $\log x$ и обратных тригонометрических функций. Однако, производные этих функций уже выражаются через функции, присутствующие в таблице интегралов. В интеграле

$$\int x^2\,\ln x\,dx$$

следует выбрать $F(x) = \ln x$ с тем, чтобы эту функцию продифференцировать. Тогда оставшаяся часть $G'(x) = x^2$ должна быть проинтегрирована, что не вызывает труда: $G(x) = x^3/3$. По правилу интегрирования по частям:

$$\int x^2\,\ln x\,dx = \frac{1}{3}x^3\ln x - \frac{1}{3}\int x^3\frac{1}{x}\,dx = \frac{1}{3}x^3\ln x - \frac{1}{9}x^3 + C.$$

Интегралы, у которых в подынтегральном выражении присутствуют множителями функции $\operatorname{arctg} x$, $\arcsin x$, $\arccos x$, не устраняемые непосредственной заменой переменной, обычно вычисляются при помощи интегрирования по частям.

Интегрирование рациональных функций

Неопределенный интеграл от любой рациональной функции вычисляется, т.е. может быть представлен через элементарные функции. Пусть P и Q — многочлены, и требуется вычислить интеграл $\int P(x)/Q(x)\,dx$. Решение разбивается на несколько этапов.

1. Деление многочленов с остатком. Если степень P меньше, чем степень Q, то этот этап пропускается. Если же рациональная функция $P(x)/Q(x)$ не является правильной, то можно применить теорему о делении многочленов:

$$\int \frac{P(x)}{Q(x)}\,dx = \int S(x)\,dx + \int \frac{r(x)}{Q(x)}\,dx, \qquad \deg r < \deg Q.$$

Интеграл от многочлена вычисляется непосредственно, и остается вычислить интеграл от правильной рациональной функции.

2. Разложение знаменателя на множители. По теореме о разложении многочлена на множители с вещественными коэффициентами знаменатель $Q(x)$ можно представить в виде произведения:

$$Q(x) = c\prod_{i=1}^{m}(x-\alpha_i)^{k_i}\ \prod_{j=1}^{n}(x^2+p_jx+q_j)^{l_j},$$

где c — старший коэффициент Q и все дискриминанты квадратных трехчленов отрицательны: $D_j = p_j^2 - 4q_j < 0$, $1 \le j \le n$.

3. Разложение на простейшие. В соответствии с теоремой о разложении правильной рациональной функции на простейшие интеграл можно разбить в сумму интегралов от простейших рациональных функций:

$$\int \frac{r(x)}{Q(x)}\,dx = \sum_{i=1}^{m}\sum_{s=1}^{k_i}\int \frac{A_{i,s}}{(x-\alpha_i)^s}\,dx + \sum_{j=1}^{n}\sum_{s=1}^{l_j}\int \frac{B_{j,s}x + C_{j,s}}{(x^2+p_jx+q_j)^s}\,dx.$$

Коэффициенты разложения $A_{i,s}$, $B_{j,s}$ и $C_{j,s}$ можно определить путем приведения суммы дробей в правой части к общему знаменателю и затем приравнивания числителей левой и правой частей.

4. Вычисление интегралов от простейших дробей. Каждая из простейших дробей интегрируется индивидуально. Рассмотрим сначала дроби с вещественными корнями. Если степень знаменателя равна 1, то

$$\int \frac{A}{x-\alpha}\,dx = A\ln|x-\alpha| + C.$$

Если степень знаменателя больше 1, то

$$\int \frac{A}{(x-\alpha)^k}\,dx = \frac{A}{(1-k)(x-\alpha)^{k-1}} + C.$$

Для вычисления интеграла от простейшей рациональной функции со знаменателем $(x^2+px+q)^k$ сначала нужно избавиться от линейного члена в знаменателе. Для этого выполняется тождественное преобразование

$$x^2+px+q = \left(x+\frac{p}{2}\right)^2 + q - \frac{p^2}{5} = y^2 + |D|,$$

где $D = p^2/4 - q < 0$ — дискриминант знаменателя. Сделаем замену переменной $y = x + p/2$. Тогда

$$\int \frac{Bx+C}{(x^2+px+q)^k}\,dx = \int \frac{Fy+G}{y^2+|D|}\,dy = F\int \frac{y\,dy}{(y^2+|D|)^k} + G\int \frac{dy}{(y^2+|D|)^k},$$

где $F = B$, $G = C - Bp/2$. Далее, первый интеграл вычисляется заменой переменной $z = y^2 + |D|$. Если $k = 1$, то

$$\int \frac{y\,dy}{y^2+|D|} = \frac{1}{2}\int\frac{dz}{z} = \frac{1}{2}\ln|z| + C = \frac{1}{2}\ln(x^2+px+q) + C,$$

где C — произвольная константа интегрирования. Если $k > 1$, то

$$\int \frac{y\,dy}{(y^2+|D|)^k} = \frac{1}{2}\int \frac{dz}{z^k} = \frac{1}{2(1-k)z^{k-1}} + C = \frac{1}{2(1-k)(x^2+px+q)^{k-1}} + C.$$

Во втором инетграле сделаем замену переменной $y = \sqrt{|D|}z$:

$$\int \frac{dy}{(y^2+|D|)^k} = \int \frac{\sqrt{|D|}\,dz}{|D|^k(z^2+1)^k} = \frac{1}{|D|^{k-\frac{1}{2}}}\int \frac{dz}{(z^2+1)^k}.$$

При $k = 1$ интеграл табличный и равен $\operatorname{arctg} z + C$, поэтому

$$\int \frac{dy}{y^2+|D|} = \frac{1}{\sqrt{|D|}}\operatorname{arctg}\frac{x+\frac{p}{2}}{\sqrt{|D|}} + C.$$

Если $k > 1$, то интеграл вычисляется рекуррентно, последовательным понижением степени k, пока она не дойдет до 1. Введем обозначение:

$$\int \frac{dz}{(z^2+1)^n} = I_n, \qquad i = 1, 2, \ldots$$

Проведем один раз интегрирование по частям:

$$\begin{aligned} I_n &= \int \frac{dz}{(z^2+1)^n} = \frac{z}{(z^2+1)^n} + 2n\int \frac{z^2\,dz}{(z^2+1)^{n+1}} \\ &= \frac{z}{(z^2+1)^n} + 2n\left(\int \frac{dz}{(z^2+1)^n} - \int \frac{dz}{(z^2+1)^{n+1}}\right) = \frac{z}{(z^2+1)^n} + 2nI_n - 2nI_{n+1}. \end{aligned}$$

Отсюда получаем рекуррентную формулу расчета интегралов I_n:

$$I_{n+1} = \frac{z}{2n(z^2+1)^n} + \frac{2n-1}{2n}I_n, \qquad n = 1, 2, \ldots$$

При начальном данном $I_1 = \operatorname{arctg} z + C$ данная формула позволяет вычислить интегралы I_k при любом натуральном показателе k.

Таким образом, вычисляется интеграл от любой простейшей рациональной функции, а следовательно, и от произвольной рациональной функции.