А. Г. Мордкович, А. С. Солодовников

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

y x z

А.Г. Мордкович, А.С. Солодовников

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Допущено
Министерством приборостроения,
средств автоматизации и систем управления
в качестве учебника
для учащихся техникумов,
обучающихся по специальности
«Прикладная математика»

Москва «Высшая школа» 1990

ББК 22.16
М 79
УДК 517

Рецензенты: д-р физ.-мат. наук, проф. Э. М. Карташов (Московский институт тонкой химической технологии); препод. М. М. Селиванова (Московский математический техникум)

Мордкович А. Г., Солодовников А. С.
М79 Математический анализ: Учеб. для техникумов. — М.: Высш. шк., 1990. — 416 с.: ил.
ISBN 5-06-001008-2

Учебник написан в соответствии с программой курса «Математический анализ» для техникумов по специальности «Прикладная математика». Он отличается высоким научно-методическим уровнем изложения материала. Помимо математического анализа, учебник включает такие разделы, как элементы математической логики, теории функций комплексной переменной. Изложение иллюстрируется большим количеством примеров.

М $\frac{1602070000-482}{001(01)-90}$ 95—90

ББК 22.16
517.2

ISBN 5-06-001008-2

Оглавление

Предисловие

Настоящий учебник написан в соответствии с программой курса математического анализа для учащихся техникумов по специальности «Прикладная математика». Этот курс, согласно программе, включает не только традиционные разделы, такие, как введение в анализ, дифференциальное исчисление, интегральное исчисление и т. д., но и два других раздела — элементы математической логики и элементы теории функций комплексной переменной. Включение в курс математического анализа первого из них объясняется тем, что он дает учащемуся более полное представление о логических средствах, используемых в математических рассуждениях; кроме того, математическая логика используется в теории вычислительных машин, теории автоматов, в некоторых вопросах экономики. Вместе с тем учебник содержит также ряд тем школьного курса алгебры и начал анализа (прогрессии, степенная, показательная, логарифмическая, тригонометрические функции, тригонометрические уравнения, показательные и логарифмические уравнения и неравенства, преобразования графиков элементарных функций). Это, естественно, связано с особенностями программы курса, ориентированной на лиц, имеющих неполное среднее образование.

Методика изложения материала в учебнике определяется прежде всего тем, что книга предназначена для учащихся, не имеющих пока среднего образования, а потому не готовых к полноценному усвоению математического анализа в традиционном изложении. Поэтому авторы стремились к более тщательной проработке ведущих понятий, основанной на неформальных содержательных рассуждениях, активно использовали геометрические иллюстрации (их существенно больше, чем в традиционных курсах). Кроме того, авторы не ставили своей целью доказать абсолютно все утверждения: наиболее важные, принципиальные результаты доказаны подробно и обстоятельно, но опущен ряд доказательств, не содержащих новых, по сравнению с уже использованными, идей и отличающихся от ранее проведенных доказательств лишь техническими сложностями (например, доказано, что предел произведения равен произведению пределов, но опущено аналогичное доказательство о пределе частного, или доказано, что $\lim\limits_{n \to \infty}\left(1 + \frac{1}{n}\right)^n = e$, но приведено без доказательства обобщение этого результата: $\lim\limits_{x \to \infty}\left(1 + \frac{1}{x}\right)^x = e$ и т. д.).

Отметим методические особенности книги, которые, на наш взгляд, помогут учащимся в изучении курса: сложные доказательства во многих случаях разбиты на четкие смысловые этапы; имеется ряд учебных алгоритмов, представляющих собой ориентировочную основу действий для решения того или иного круга задач; практически в каждом параграфе приводится достаточно большое число примеров с подробными решениями.

Изложенный в учебнике материал распределился между авторами следующим образом: гл. 3—7 написаны проф. А. Г. Мордковичем, а гл. 1, 2, 8—10 — проф. А. С. Солодовниковым.

Авторы выражают признательность рецензентам проф. Э. М. Карташову и препод. М. М. Селивановой, высказавшим ряд полезных замечаний, которые были учтены при подготовке книги к изданию.

Авторы

1

Элементы математической логики

§ 1. Высказывания

1. Понятие высказывания

Любая научная теория воспринимается нами как некоторая система утверждений. Истинность каждого из них, вообще говоря, нуждается в доказательстве. В отдельных случаях такое доказательство может проводиться опытным путем, но чаще всего оно достигается с помощью логических средств. Именно эти логические средства и изучает раздел математики, называемый *математической логикой*. Исходным понятием математической логики является понятие высказывания.

Определение 1. *Высказыванием* называется предложение, которое может быть либо истинным, либо ложным.

Примеры. 1. Предложение «Снег — белый» есть истинное высказывание.

2. Предложение «Волга впадает в Средиземное море» — ложное высказывание.

3. Предложение «$2 + 2 = 10$» — ложное высказывание.

Далеко не всякое предложение является высказыванием. В частности, вопросительные и восклицательные предложения не относятся к высказываниям. Например, по поводу предложения «Который час?» не имеет смысла ставить вопрос, истинно оно или ложно; то же самое относится, скажем, к предложению «Мойте руки перед едой!» Не являются высказываниями и такие предложения, которые служат определениями чего-либо, например: «Трапецией называется четырехугольник, две стороны которого параллельны».

Существуют предложения, которые безусловно являются истинными или ложными, однако в силу недостаточности наших знаний мы не можем в данный момент сказать точно, истинны они или ложны. Например, «Земля — единственная обитаемая планета во Вселенной» или «Всякое четное число есть сумма двух простых» (нерешенная до конца проблема теории чисел). Предложения такого типа мы также считаем высказываниями.

Из всех свойств высказывания нас будет в дальнейшем интересовать только одно: истинно оно или ложно. Все же прочие свойства высказывания, например особенности его грамматической формы, смысловое значение отдельных слов и всего высказывания в целом, будут оставаться как бы вне поля зрения.

В дальнейшем будем обозначать высказывания заглавными буквами латинского алфавита: P, Q, R и т. д.

2. Значение истинности высказывания

Условимся каждому истинному высказыванию сопоставлять число 1, а ложному — число 0. Иначе говоря, введем на множестве всех высказываний функцию $\alpha(P)$, которая принимает значения 1 или 0 в зависимости от того, истинно высказывание P или ложно:

$$\alpha(P) = \begin{cases} 1, & \text{если высказывание } P \text{ истинно;} \\ 0, & \text{если высказывание } P \text{ ложно.} \end{cases}$$

Определение 2. Число $\alpha(P)$ называется *значением истинности* высказывания P.

Так, для высказывания P: «В неделе 7 дней» $\alpha(P) = 1$, а для высказывания Q: «Волга впадает в Средиземное море» $\alpha(Q) = 0$.

§ 2. Операции над высказываниями

Будем считать, что имеется некоторая первоначальная совокупность высказываний, называемых *элементарными* (или *исходными*). Исходя из этих высказываний, с помощью так называемых *логических операций* строят новые (сложные) высказывания. Перейдем к точному описанию этих операций.

1. Отрицание высказывания

Определение 1. *Отрицанием* высказывания P называется новое высказывание, обозначаемое $\bar{P}$ (читается: «Не P» или «Неверно, что P»), которое считается истинным, если высказывание P ложно, и ложным, если P истинно.

Иначе говоря, значения истинности высказываний P и $\bar{P}$ связаны между собой, как указано в следующей таблице:

$\alpha(P)$	$\alpha(\bar{P})$
1	0
0	1

Эта таблица читается по строкам. Например, первая строка под горизонтальной чертой означает: если $\alpha(P) = 1$, то $\alpha(\bar{P}) = 0$. Приведенная таблица называется *таблицей истинности для отрицания.*

Заметим, что построение высказывания $\bar{P}$ с помощью приписывания предлога «не» к высказыванию P часто приводит к неблагозвучным словосочетаниям. Например, для высказывания P: «Снег — белый» построенное указанным способом высказывание $\bar{P}$ выглядит так: «Не снег — белый». В подобных случаях частицу «не» переносят на другое, более подходящее с точки зрения грамматических норм, место в предложении (обычно — перед сказуемым). Например, для указанного выше высказывания «Снег — белый» его отрицание может быть записано так: «Снег — не белый».

По сравнению с частицей «не» оборот «Неверно, что» более универсален: его можно приписать к любому высказыванию и получить осмысленный текст. Например: «Неверно, что снег — белый».

2. Конъюнкция высказываний

Определение 2. *Конъюнкцией* высказываний P и Q называется новое высказывание, обозначаемое $P \wedge Q$ (читается «P и Q»), которое считается истинным, если истинны оба высказывания P и Q, и ложным во всех остальных случаях.

Таким образом, значение истинности высказывания $P \wedge Q$ связано со значениями истинности высказываний P и Q. Эта связь выражается таблицей:

$\alpha(P)$	$\alpha(Q)$	$\alpha(P \wedge Q)$
1	1	1
1	0	0
0	1	0
0	0	0

Приведенная таблица называется *таблицей истинности для конъюнкции.*

Данное выше определение конъюнкции вполне отвечает тому смыслу, который придается в рассуждениях союзу «и». Действительно, привычная логика рассуждений требует, чтобы утверждение «P и Q» было истинно лишь в одном случае: когда истинны оба утверждения P и Q.

Примеры. 1. Высказывание «Число 2 четное и простое» является конъюнкцией высказываний: «Число 2 четное» и «Число 2 простое». Так как оба последних высказывания истинны, то истинна и их конъюнкция.

2. Высказывания «2 меньше 5» и «5 меньше 10» истинны, поэтому истинна и их конъюнкция «2 меньше 5 и 5 меньше 10». Последнее высказывание записывают обычно так: «$2<5<10$».

3. Дизъюнкция высказываний

Определение 3. *Дизъюнкцией* высказываний P и Q называется новое высказывание, обозначаемое $P \vee Q$ (читается «P или Q»), которое истинно в тех случаях, если истинно хотя бы одно из высказываний P или Q, и ложно, если ложны оба высказывания P и Q.

Значение истинности высказывания $P \vee Q$ связано со значениями истинности высказываний P и Q с помощью таблицы:

$\alpha(P)$	$\alpha(Q)$	$\alpha(P \vee Q)$
1	1	1
1	0	1
0	1	1
0	0	0

Эта таблица называется *таблицей истинности для дизъюнкции.*

Приведенное определение дизъюнкции вполне отвечает обычному употреблению союза «или». Действительно, в практике рассуждений утверждение «P или Q» считается верным в любом из случаев, когда верно P или Q; если же оба утверждения P и Q неверны, то неверно и «P или Q».

Примеры. 1. Высказывание «В неделе 10 дней или в году 12 месяцев» является дизъюнкцией двух высказываний: «В неделе 10 дней» *и* «В году 12 месяцев». Несмотря на кажущуюся странность такого высказывания, мы все же признаем его истинным, поскольку истинно одно из составляющих его высказываний («В году 12 месяцев»).

2. Высказывание «$2 \leqslant 3$» является дизъюнкцией высказываний «$2 < 3$» и «$2 = 3$», из которых первое истинно, а второе ложно. Следовательно, истинна и сама дизъюнкция.

4. Импликация высказываний

Определение 4. *Импликацией* высказываний P и Q называется высказывание, обозначаемое $P \Rightarrow Q$ (читается: «Если P, то Q», или «Из P следует Q», или «P влечет за собой Q»), которое ложно лишь в том случае, если P истинно, а Q ложно.

Таблица истинности для импликации имеет вид

$\alpha(P)$	$\alpha(Q)$	$\alpha(R \Rightarrow Q)$
1	1	1
1	0	0
0	1	1
0	0	1

Данное выше определение импликации в основном отражает тот смысл, который придается в обычных рассуждениях связке «если..., то...». Единственное возражение может вызвать, пожалуй, лишь та строка таблицы, где $\alpha(P) = 0$, $\alpha(Q) = 1$, $\alpha(P \Rightarrow Q) = 1$. Однако с таким пониманием импликации приходится все же согласиться, поскольку принцип «Из лжи следует что угодно» представляется вполне оправданным.

Заметим, что при рассмотрении импликации $P \Rightarrow Q$ высказывание P называют *посылкой* (или *условием*) импликации, а высказывание Q — ее *заключением* (или *следствием*).

Примеры. 1. Высказывание «Если Земля круглая, то $2 \times 2 = 4$» является импликацией высказываний «Земля круглая» и «$2 \times 2 = 4$». Оно истинно, так как истинны оба последних высказывания.

2. Высказывание «Если $2 \times 2 = 5$, то число 5 — простое» есть импликация высказываний «$2 \times 2 = 5$» и «5 — простое». Оно истинно, поскольку посылка «$2 \times 2 = 5$» — ложное высказывание.

5. Эквивалентность высказываний

Определение 5. *Эквивалентностью* (или *эквиваленцией*) высказываний P и Q называется новое высказывание, обозначаемое $P \Leftrightarrow Q$ (читается «P эквивалентно Q», или «P тогда и только тогда, когда Q»), которое истинно в том и только в том случае, если P и Q одновременно истинны или одновременно ложны.

Таблица истинности для эквивалентности выглядит следующим образом:

$\alpha(P)$	$\alpha(Q)$	$\alpha(P \Leftrightarrow Q)$
1	1	1
1	0	0
0	1	0
0	0	1

Примеры. 1. Высказывание «$2 \times 2 = 4$ тогда и только тогда, когда Земля — шар» представляет собой эквиваленцию двух высказываний: «$2 \times 2 = 4$» и «Земля — шар». Оно истинно, поскольку истинны оба этих высказывания.

2. Высказывание «Небо синее в том и только в том случае, когда снег черный» является эквиваленцией высказываний «Небо синее» и «Снег черный». Оно ложно, так как одно из двух последних высказываний истинно, а другое ложно.

6. Логические операции как операции на множестве $\{0, 1\}$

Рассмотрим любую из логических операций, например операцию конъюнкции $\wedge$. Поскольку число $\alpha(P \wedge Q)$ полностью определяется числами $\alpha(P)$ и $\alpha(Q)$, мы можем оперировать не с высказываниями, а с числами 0 и 1, определив конъюнкцию над ними с помощью таблицы

$$\begin{gathered} 1 \wedge 1 = 1, \\ 1 \wedge 0 = 0, \\ 0 \wedge 1 = 0, \\ 0 \wedge 0 = 1. \end{gathered}$$

Аналогичные замечания можно сделать и по отношению к остальным логическим операциям.

Например,

$$\overline{1} = 0,\ \overline{0} = 1,\ 1 \vee 0 = 1,\ 0 \Rightarrow 0 = 1$$

и т. д.

Таким образом, каждой логической операции над высказываниями соответствует некоторая функция, определенная на двухэлементном множестве $\{0,1\}$ и принимающая значения в том же множестве. Эту функцию мы будем называть тем же термином (отрицание, конъюнкция, дизъюнкция и т. д.), что и соответствующую логическую операцию.

§ 3. Формулы алгебры высказываний

1. Определение и примеры формул

С помощью логических операций, рассмотренных в предыдущем параграфе, можно, исходя из простейших высказываний, строить новые, более сложные. Например, исходя из высказываний P: «Пушкин — русский поэт», Q: «Гаусс — немецкий математик», R: «$2 \times 2 = 4$», можно построить новое высказывание: «Если Пушкин — русский поэт и Гаусс — немецкий математик, то $2 \times 2 = 4$». Это новое высказывание имеет вид

$$(P \wedge Q) \Rightarrow R. \tag{1}$$

Выражение (1), если отвлечься от конкретного смысла высказываний P, Q, R, можно рассматривать как некоторую схему, позволяющую, исходя из любых высказываний P, Q, R, строить новое высказывание. Именно такие схемы и будут нас сейчас интересовать. Они называются *формулами алгебры высказываний*. Мы, конечно, привыкли к несколько другому толкованию понятия «формула»; для нас формулы — это равенства типа $S = \pi r^2$ (формула площади круга), $a^2 + b^2 = c^2$ (формула, выражающая теорему Пифагора) и т. д. Тем не менее выражение $(P \wedge Q) \Rightarrow R$ также можно рассматривать как своего рода формулу — формулу конструирования составного высказывания из более простых.

Прежде чем дать общее определение формулы алгебры высказываний, условимся о нижеследующем. *Высказывательными переменными* будем называть такие переменные, которые могут принимать в качестве своих значений любые конкретные высказывания. Будем обозначать такие переменные заглавными латинскими буквами X, Y, Z, U, V, ..., или теми же буквами с индексами: X_1, X_2, ... Введем, кроме того, еще две специфические высказывательные переменные И и Л; вместо первой можно подставлять любое истинное высказывание, вместо второй — любое ложное.

Полное описание понятия формулы дают следующие соглашения:

1^0. Каждая отдельно взятая высказывательная переменная есть формула.

2^0. Если F_1 и F_2 — две формулы, то выражения $\bar{F}_1$, $\bar{F}_2$, $(F_1 \wedge F_2)$, $(F_1 \vee F_2)$, $(F_1 \Rightarrow F_2)$, $(F_1 \Leftrightarrow F_2)$ также являются формулами.

3^0. Не существует никаких других формул, кроме тех, которые получаются в результате применения конечного числа раз пп.1^0 и 2^0.

Так, например, формулами являются следующие выражения: $\bar{X}$, $(X \wedge Y)$, $(\bar{X} \vee Y)$, $((X \Rightarrow Y) \Rightarrow Z)$ и т. д.

Для большей отчетливости укажем примеры выражений, не являющихся формулами: XY, $(X \Rightarrow) \vee Y$, $X \wedge Y$.

То, что мы не признаем формулой последнее из написанных выражений, может вызвать сначала недоумение. Однако если строго следовать данному выше определению, то выражение $X \wedge Y$ — не формула; чтобы стать формулой, ему нехватает скобок, так как, согласно п. 2^0, формулой должно быть выражение $(X \wedge Y)$, а не $X \wedge Y$. Различие между выражениями $X \wedge Y$ и $(X \wedge Y)$ станет особенно существенным, если мы включим выражение $X \wedge Y$ в качестве составляющей части в более сложную формулу: сравните, например, выражение $((X \wedge Y) \Rightarrow Z)$, являющееся формулой, с выражением $X \wedge Y \Rightarrow Z$ (не формулой); неясно, как следует понимать второе выражение (какая операция следует за какой: $\wedge$ после $\Rightarrow$ или $\Rightarrow$ после $\wedge$?).

Итак, *запись внешних скобок у формулы будем считать необязательной, если только эта формула не входит составной частью в более сложную формулу.*

2. Таблицы истинности для формул

Рассмотрим какую-нибудь формулу алгебры высказываний, например $(X \wedge Y) \Rightarrow Z$. Обозначим эту формулу сокращенно $F(X, Y, Z)$. Значение истинности формулы $F(X, Y, Z)$ полностью определяется значениями истинности переменных X, Y, Z. Это обстоятельство позволяет составить таблицу, дающую значение истинности для $F(X, Y, Z)$ в зависимости от значений истинности для X, Y, Z. Такая таблица должна состоять из четырех столбцов: трех — для переменных X, Y, Z и одного — для самой формулы. Так как каждая из переменных X, Y, Z может принимать два значения (1 или 0), то для тройки X, Y, Z получается $2 \times 2 \times 2 = 8$ различных возможностей. Это означает, что таблица должна иметь 8 строк. Для заполнения последнего столбца таблицы подставляем значения X, Y и Z в формулу $F(X, Y, Z)$. Например, при $X = 1$, $Y = 1$, $Z = 1$ имеем

$$F = (1 \wedge 1) \Rightarrow 1 = 1 \Rightarrow 1 = 1,$$

при $X = 1$, $Y = 1$, $Z = 0$ находим

$$F = (1 \wedge 1) \Rightarrow 0 = 1 \Rightarrow 0 = 0$$

и т. д. В результате заполнения получаем следующую таблицу:

X	Y	Z	$(X \wedge Y) \Rightarrow Z$
1	1	1	1
1	1	0	0
1	0	1	1
1	0	0	1
0	1	1	1
0	1	0	1
0	0	1	1
0	0	0	1

Вообще, для каждой формулы $F(X_1, X_2, \ldots, X_n)$ алгебры высказываний можно составить таблицу, дающую значение истинности формулы в зависимости от значений истинности переменных $X_1, X_2, \ldots, X_n$. Такая таблица называется ***таблицей истинности для формулы*** $F(X_1, X_2, \ldots, X_n)$.

Процедуру составления таблицы истинности можно упростить, используя некоторые приемы. Проиллюстрируем это на следующем примере.

Пример. Составить таблицу истинности для формулы

$$(X \vee Y) \Leftrightarrow (\bar{X} \Rightarrow \bar{Y})$$

(в силу принятого соглашения внешние скобки у формулы опущены).

○ Первый шаг заключается в установлении последовательности операций. Для этого над знаком каждой логической операции, встречающейся в формуле, ставим номер, означающий очередность выполнения этой операции. В данном случае возможна, например, такая нумерация:

$$(X \overset{\textcircled{1}}{\vee} Y) \overset{\textcircled{5}}{\Leftrightarrow} (\overset{\textcircled{2}}{\bar{X}} \overset{\textcircled{4}}{\Rightarrow} \overset{\textcircled{3}}{\bar{Y}}).$$

В первой (заглавной) строке таблицы запишем X, Y, а также данную формулу. Под переменными X и Y выписываем всевозможные наборы их логических значений. Тогда получим таблицу

X	Y	$(X \overset{\textcircled{1}}{\vee} Y) \overset{\textcircled{5}}{\Leftrightarrow} (\overset{\textcircled{2}}{\bar{X}} \overset{\textcircled{4}}{\Rightarrow} \overset{\textcircled{3}}{\bar{Y}})$
1	1	1 1
1	0	1 0
0	1	0 1
0	0	0 0

которую продолжаем заполнять дальше. Под номером ① запишем значения, принимаемые формулой $Y \vee Y$ для соответствующих значений X и Y. Затем точно так же заполняем столбцы под номерами ②, ③, ④, ⑤. В результате получаем следующую таблицу:

X	Y	$(X \overset{\textcircled{1}}{\vee} Y) \overset{\textcircled{5}}{\Leftrightarrow} (\overset{\textcircled{2}}{\bar{X}} \overset{\textcircled{4}}{\Rightarrow} \overset{\textcircled{3}}{\bar{Y}})$
1	1	1 1 1 1 0 1 0
1	0	1 1 0 1 0 1 1
0	1	0 1 1 0 1 0 0
0	0	0 0 0 0 1 1 1

Столбец под номером ⑤ (полученный последним) дает значения истинности данной формулы. ●

3. Тавтологии

Определение 1. Формула $F(X_1, \ldots, X_n)$ алгебры высказываний называется *тождественно истинной* (или *тавтологией*), если ее значение истинности равно 1 при любых значениях истинности для $X_1, \ldots, X_n$.

Роль тавтологий прежде всего заключается в том, что они дают схемы построения истинных высказываний, независимо от содержания и истинности составляющих высказываний.

Например, тавтологией является формула

$$X \vee \bar{X} \tag{2}$$

(«X или не X»). Действительно, какое бы конкретное высказывание ни было подставлено вместо X, высказывание (2) истинно, поскольку $1 \vee 0 = 1$ и $0 \vee 1 = 1$.

Однако значение тавтологий состоит не только в том, что с их помощью строятся истинные высказывания; не меньшее значение имеет и то, что *тавтологии дают правильные способы умозаключения.* Проиллюстрируем это на примере формулы

$$((\bar{X} \Rightarrow Y) \wedge (\bar{X} \Rightarrow \bar{Y})) \Rightarrow X.$$

Эта формула является тавтологией, что легко проверяется, если составить для нее таблицу истинности (в столбце значений для всей формулы окажутся одни единицы). Схема логического умозаключения, выражаемая этой тавтологией, часто используется в математике: эта схема носит название «доказательство от противного». А именно, пусть требуется доказать некоторое утверждение X. Рассуждаем так: допустим, что X неверно (т. е. что верно $\bar{X}$). Далее с помощью некоторого рассуждения (в рамках той теории, которая изучается) доказываем, что из $\bar{X}$ следует некоторое утверждение Y, а также — что из $\bar{X}$ следует противоположное утверждение $\bar{Y}$. Так как одновременная справедливость утверждений Y и $\bar{Y}$ невозможна (тавтология (2)), то из проведенного рассуждения делаем вывод о справедливости (истинности) X.

Укажем некоторые особо важные тавтологии.

1^0. *Законы коммутативности конъюнкции и дизъюнкции:*

$$(X \wedge Y) \Leftrightarrow (Y \wedge X), \quad (X \vee Y) \Leftrightarrow (Y \vee X).$$

2^0. *Законы ассоциативности конъюнкции и дизъюнкции:*

$$((X \wedge Y) \wedge Z) \Leftrightarrow (X \wedge (Y \wedge Z)), \quad ((X \vee Y) \vee Z) \Leftrightarrow (X \vee (Y \vee Z)).$$

3^0. *Законы дистрибутивности:*

$$(X \wedge (Y \vee Z)) \Leftrightarrow ((X \wedge Y) \vee (X \wedge Z)), \quad (X \vee (Y \wedge Z)) \Leftrightarrow ((X \vee Y) \wedge (X \vee Z)).$$

4^0. *Законы де Моргана:*

$$(\overline{X \wedge Y}) \Leftrightarrow (\bar{X} \vee \bar{Y}), \quad (\overline{X \vee Y}) \Leftrightarrow (\bar{X} \wedge \bar{Y}).$$

5^0. *Закон исключенного третьего:*

$$X \vee \bar{X}.$$

6^0. *Закон контрапозиции:*

$$(X \Rightarrow Y) \Leftrightarrow (\bar{Y} \Rightarrow \bar{X}).$$

7^0. *Правило цепного заключения (закон силлогизма):*

$$((X \Rightarrow Y) \wedge (Y \Rightarrow Z)) \Rightarrow (X \Rightarrow Z).$$

8°. *Правило «модус поненс»:*

$$(X \wedge (X \wedge Y)) \Rightarrow Y.$$

9°. *Схема доказательства «от противного»:*

$$(\bar{X} \Rightarrow Y) \wedge (\bar{X} \Rightarrow \bar{Y}) \Rightarrow X.$$

Доказательство того, что каждая из указанных формул является тавтологией, рекомендуем провести самостоятельно в качестве упражнения.

4. Равносильность формул

Определение 2. Две формулы $F(X_1, ..., X_n)$ и $H(X_1, ..., X_n)$ алгебры высказываний называются *равносильными,* если при любых логических значениях переменных $X_1, ..., X_n$ логические значения высказываний F и H совпадают.

Равносильность формул F и H записывается так: $F \cong H$.

Существует тесная связь между понятием равносильности формул и понятием тавтологии. Она заключается в следующем: *формулы F и H равносильны тогда и только тогда, когда формула $F \Leftrightarrow H$ является тавтологией.*

Справедливость этого утверждения следует непосредственно из самих определений равносильности формул и тавтологии.

Примеры. 1. Доказать равносильность формул $X \Leftrightarrow Y$ и $(\bar{X} \vee Y) \wedge (\bar{Y} \vee X)$.

○ Для каждой из данных формул составим таблицу истинности:

X	Y	$X \Leftrightarrow Y$
1	1	1
1	0	0
0	1	0
0	0	1

X	Y	$(\bar{X} \vee Y) \wedge (\bar{Y} \vee X)$
1	1	1
1	0	0
0	1	0
0	0	1

Сравнивая таблицы, видим, что указанные формулы равносильны.●

2. Доказать равносильность формул $X \Rightarrow Y$ и $\bar{X} \vee Y$.

○ Разумеется, можно сравнить таблицы истинности данных формул. Однако можно рассуждать и так: формула $X \Rightarrow Y$ ложна лишь в случае $X = 1$, $Y = 0$, а формула $\bar{X} \vee Y$ — лишь в случае $\bar{X} = 0$, $Y = 0$, т. е. при $X = 1$, $Y = 0$. Таким образом, обе формулы ложны или истинны одновременно.●

Целый ряд равносильностей можно получить, исходя из приведенных в п. 3 тавтологий. Например, формулы $\overline{X \wedge Y}$ и $\bar{X} \vee \bar{Y}$ равносильны, поскольку формула $(\overline{X \wedge Y}) \Leftrightarrow (\bar{X} \vee \bar{Y})$ является тавтологией (см. тавтологии 4°).

Следует заметить, что выражение $F \cong H$ не является формулой. Оно представляет собой просто запись того факта, что между формулами F и H имеется определенного рода связь (а именно, что F равносильна H).

§ 4. Применение алгебры высказываний в математических рассуждениях

1. Понятие логического следования

Одна из главных задач науки состоит в установлении причинно-следственных связей. В решении этой задачи важнейшая роль принадлежит логике, поскольку она устанавливает правила логического следования, логического умозаключения.

Когда мы говорим, что из нескольких предложений $A_1, \ldots, A_n$ следует предложение B, то подразумеваем под этим следующее: всякий раз, когда оказываются истинными все предложения $A_1, \ldots, A_n$ (или, что то же, когда оказывается истинным предложение $A_1 \wedge A_2 \wedge \ldots \wedge A_n$), истинно и предложение B. Например: «Если я успешно сдам все экзамены за первый курс (предложение A), то буду переведен на второй курс (предложение B)», «Если воду нагреть до температуры 100° (предложение A), то она превратится в пар (предложение B)».

Установление истинности такого рода предложений не относится к компетенции логики: этот вопрос решается на основании анализа конкретного содержания составляющих высказываний. В чем же заключается роль математической логики в вопросах логического следования? Ответ на этот вопрос следующий: логика должна указать такие формы построения высказываний $A_1, \ldots, A_n, B$, которые гарантировали бы, что *высказывание B является следствием высказываний $A_1, \ldots, A_n$, независимо от конкретного содержания высказываний $A_1, \ldots, A_n$, B*. Простейший пример: из высказываний $A \Rightarrow B$, $B \Rightarrow C$ всегда следует $A \Rightarrow C$ независимо от того, какое конкретное содержание имеют высказывания A, B, C.

Определение. Формула $H(X_1, \ldots, X_n)$ называется ***логическим следствием*** формул $F_1(X_1, \ldots, X_n), \ldots, F_s(X_1, \ldots, X_n)$, если она превращается в истинное высказывание при замене $X_1, \ldots, X_n$ любыми конкретными высказываниями $A_1, \ldots, A_n$, такими, что каждое из высказываний $F_1(A_1, \ldots, A_n), \ldots, F_s(A_1, \ldots, A_n)$ истинно.

Тот факт, что формула H является логическим следствием формул $F_1, \ldots, F_s$, записывают следующим образом:

$$F_1, \ldots, F_s \models H.$$

При этом формулы $F_1, \ldots, F_s$ называют *посылками*, а формулу H — *следствием*.

Для установления логического следования одних формул из других можно использовать таблицы истинности. Предположим, например, что требуется выяснить, является ли формула $H(X, Y)$ логическим следствием формул $F_1(X, Y)$ и $F_2(X, Y)$. Для этого дополняем таблицу значений аргументов X и Y тремя столбцами: первым — для формулы F_1, вторым — для F_2, третьим — для H. В полученной таблице отмечаем те строки, в которых одновременно $F_1 = 1$ и $F_2 = 1$. Если в каждой из таких строк окажется, что и $H = 1$, то формула $H(X, Y)$ является логическим следствием формул $F_1(X, Y)$ и $F_2(X, Y)$.

Пример 1. Показать, что формула $\bar{X}$ является логическим следствием формул $F_1: X \Rightarrow Y$ и $F_2: X \Rightarrow \bar{Y}$.

○ Составим «сводную» таблицу истинности для всех данных формул:

X	Y	F_1	F_2	H
1	1	1	0	0
1	0	0	1	0
0	1	1	1	1
0	0	1	1	1

Убеждаемся, что в тех строках, где $F_1 = 1$, $F_2 = 1$, также и $H = 1$. Значит, формула H есть логическое следствие формул F_1 и F_2. ●

Заметим, что установленное в примере 1 логическое следование есть в сущности одна из форм доказательства «от противного» (если из X следуют два исключающих друг друга высказывания Y и $\bar{Y}$, то X не может быть истинным).

Если для установления логического следования мы располагаем, по существу, только одним универсальным способом (таблицами истинности), то для установления того факта, что формула H не является логическим следствием формул F_1, ..., F_s, можно использовать и другие приемы. Для этого достаточно построить хотя бы один набор значений аргументов X_1, ..., X_n (т. е. набор из n нулей и единиц), при котором $F_1 = 1$, ..., $F_s = 1$, в то время как $H = 0$. Часто такой метод оказывается более экономным, чем составление таблиц истинности.

Пример 2. Показать, что формула $H: \bar{X} \Rightarrow Y$ не является логическим следствием формул $F_1: (X \wedge \bar{Y}) \Rightarrow Z$ и $F_2: \bar{Z} \Rightarrow X$.

○ Достаточно указать хотя бы один набор значений X, Y, Z, для которого обе формулы F_1, F_2 принимают значение 1, а формула H принимает значение 0. Из вида формулы H следует, что $H = 0$ только в случае $X = 0$, $Y = 0$. Теперь остается подобрать значение Z таким образом, чтобы $F_1 = 1$, $F_2 = 1$. Из вида F_2 следует, что для этого должно быть $Z = 1$; при таком выборе Z получим, что и $F_1 = 1$. Итак, при $X = 0$, $Y = 0$, $Z = 1$ имеем $F_1 = 1$, $F_2 = 1$, $H = 0$. Это показывает, что H не является логическим следствием F_1 и F_2. ●

В следующем параграфе будет указан способ, позволяющий для данного набора формул F_1, ..., F_s получить все формулы, являющиеся логическими следствиями F_1, ..., F_s.

2. Некоторые виды математических теорем

Математические теоремы, как правило, имеют структуру, выражаемую схемой

$$X \Rightarrow Y. \tag{1}$$

При этом предложение X называется *условием* теоремы *(посылкой)*, а предложение Y — *заключением*. Например: «Если число делится на 2, то его квадрат делится на 4» и т. п.

Если верна теорема (1), то истинность высказывания Y называют *необходимым условием* для истинности высказывания X («без Y нет X»), а истинность X — *достаточным условием* для истинности Y. Не всякое необходимое условие является достаточным и не всякое достаточное условие является необходимым; другими словами, из справедливости теоремы $X \Rightarrow Y$ не всегда следует справедливость теоремы $Y \Rightarrow X$. Например, справедливо предложение «Если две прямые параллельны, то они лежат в одной плоскости», но неверно, что «Если две прямые лежат в одной плоскости, то они параллельны».

С каждой теоремой вида $X \Rightarrow Y$ можно связать еще три теоремы: обратную, противоположную и обратно-противоположную.

Противоположной к теореме $X \Rightarrow Y$ называется теорема

$$\bar{X} \Rightarrow \bar{Y}, \tag{2}$$

в которой условие X и заключение Y исходной теоремы заменяются их отрицаниями.

Обратной к теореме $X \Rightarrow Y$ называется теорема

$$Y \Rightarrow X, \tag{3}$$

в которой условие и заключение (по сравнению с исходной теоремой) меняются ролями.

Обратно-противоположной к теореме $X \Rightarrow Y$ называется теорема

$$\bar{Y} \Rightarrow \bar{X}. \tag{4}$$

Из справедливости данной теоремы всегда вытекает справедливость обратно-противоположной теоремы. Этот факт сразу следует из общего положения математической логики, согласно которому формулы $X \Rightarrow Y$ и $\bar{Y} \Rightarrow \bar{X}$ равносильны (закон контрапозиции).

Что же касается обратной теоремы (3), а также противоположной теоремы (2), то их истинность не следует автоматически из истинности исходной теоремы (1). Так, справедлива теорема «Если число оканчивается цифрой 2, то его квадрат оканчивается цифрой 4», но не верна обратная теорема «Если квадрат числа оканчивается на 4, то само число оканчивается на 2» (пример: $8^2 = 64$); неверна и противоположная теорема, которая в данном случае имеет вид «Если число не оканчивается на 2, то его квадрат не оканчивается на 4» (тот же пример: $8^2 = 64$).

Рассмотрим теперь теорему более сложной структуры, например: «В равных треугольниках против равных сторон лежат равные углы». Чтобы четко выделить условие теоремы и ее заключение, сформулируем ее в следующем виде: «Если два треугольника равны (условие X), то из равенства двух углов этих треугольников (условие Y) следует равенство противоположных сторон (заключение Z)». Теперь мы видим, что теорема имеет строение, описываемое схемой

$$X \Rightarrow (Y \Rightarrow Z).$$

Впрочем, учитывая равносильность формул $X \Rightarrow (Y \Rightarrow Z)$, $Y \Rightarrow (X \Rightarrow Z)$, $(X \wedge Y) \Rightarrow Z$, можно сформулировать эту теорему и по-другому: «Если в двух треугольниках равны два угла, то из равенства этих треугольников следует равенство сторон, противолежащих равным углам», или: «Если два треугольника равны и если в этих треугольниках равны два угла, то равны и противолежащие этим углам стороны».

Если мы попробуем сформулировать теперь теорему, обратную данной, то обнаружим, что это можно сделать не одним, а несколькими способами. Обратная теорема — это такая, в которой условие и заключение исходной теоремы поменялись местами. Следовательно, для данной теоремы возможна не одна обратная, а по крайней мере пять:

$$(Y \Rightarrow Z) \Rightarrow X,\ X \Rightarrow (Z \Rightarrow Y),\ Y \Rightarrow (Z \Rightarrow X),\ (X \Rightarrow Z) \Rightarrow Y,\ Z \Rightarrow (X \wedge Y).$$

Рекомендуем самостоятельно проверить, что лишь первая и четвертая из этих обратных теорем справедливы, а остальные неверны.

3. Принцип полной дизъюнкции

В математике часто используется способ умозаключения, логическая основа которого носит название *принципа полной дизъюнкции*. Проиллюстрируем этот принцип на примере, взятом из геометрии.

Пусть a, b, c — длины сторон треугольника и α, β, γ — величины противолежащих углов. В курсе геометрии доказываются следующие утверждения:

1. Квадрат стороны, лежащей против острого угла, меньше суммы квадратов двух других сторон, или в символической форме:

$$\alpha < \pi/2 \Rightarrow a^2 < b^2 + c^2. \tag{5}$$

2. Квадрат стороны, лежащей против прямого угла, равен сумме квадратов двух других сторон:

$$\alpha = \pi/2 \Rightarrow a^2 = b^2 + c^2. \tag{6}$$

3. Квадрат стороны, лежащей против тупого угла, больше суммы квадратов двух других сторон:

$$\alpha > \pi/2 \Rightarrow a^2 > b^2 + c^2. \tag{7}$$

Возникает вопрос, можно ли обратить каждую из импликаций (5), (6), (7). Ответ состоит в том, что такое обращение законно. Например, можно утверждать, что

$$a^2 < b^2 + c^2 \Rightarrow \alpha < \pi/2.$$

Действительно, если предположить, что для какого-то треугольника справедливо неравенство $a^2 < b^2 + c^2$, но при этом не выполняется неравенство $\alpha < \pi/2$, то придется принять, что $\alpha = \pi/2$ или $\alpha > \pi/2$. Однако в первом из этих случаев, согласно соотношению (6), должно быть $a^2 = b^2 + c^2$, что противоречит условию, во втором же случае должно быть $a^2 > b^2 + c^2$, что вследствие соотношения (7) также противоречит условию.

Логическая основа этого рассуждения сводится к следующему. Предположим, что:

1) истинно хотя бы одно из высказываний $A_1, A_2, \ldots, A_n$;
2) истинны высказывания $A_1 \Rightarrow B_1, A_2 \Rightarrow B_2, \ldots, A_n \Rightarrow B_n$;
3) среди высказываний $B_1, B_2, \ldots, B_n$ истинно не более чем одно.

Тогда истинно и каждое из высказываний

$$B_1 \Rightarrow A_1,\ B_2 \Rightarrow A_2,\ \ldots,\ B_n \Rightarrow A_n.$$

Этот принцип можно сформулировать в виде следующей теоремы.

Теорема (принцип полной дизъюнкции). *Пусть $X_1, \ldots, X_n$ — переменные высказывания. Тогда формулы*

$$X_1 \vee X_2 \vee \ldots \vee X_n, \tag{8}$$

$$X_1 \Rightarrow Y_1, \ldots, X_n \Rightarrow Y_n, \tag{9}$$

$$\overline{Y_i \wedge Y_j}\ (i, j = 1, \ldots, n;\ i \neq j) \tag{10}$$

имеют своими логическими следствиями

$$Y_1 \Rightarrow X_1, \ldots, Y_n \Rightarrow X_n. \tag{11}$$

□ Рассуждая от противного, допустим, что какая-то из формул (11), например $Y_1 \Rightarrow X_1$, не является логическим следствием формул (8), (9), (10). Это означает, что существуют такие конкретные высказывания $A_1, \ldots, A_n$, $B_1, \ldots, B_n$, для которых истинно каждое из высказываний $A_1 \vee A_2 \vee \ldots \vee A_n$, $A_1 \Rightarrow B_1, \ldots, A_n \Rightarrow B_n$, $\overline{B_i \wedge B_j}$ $(i, j = 1, \ldots, n;\ i \neq j)$, но ложно $B_1 \Rightarrow A_1$. Согласно определению импликации, ложность $B_1 \Rightarrow A_1$ означает, что B_1 — истинное высказывание, а A_1 — ложное. Тогда в силу истинности высказывания $A_1 \vee A_2 \vee \ldots \vee A_n$ истинным должно быть хотя бы одно из высказываний $A_2, \ldots, A_n$. Пусть, например, истинно A_2. Отсюда ввиду истинности высказывания $A_2 \Rightarrow B_2$ следует, что должно быть истинным и высказывание B_2. Итак, мы получили, что истинны оба высказывания B_1 и B_2, но это противоречит истинности высказывания $\overline{B_1 \wedge B_2}$. ■

§ 5. Нормальные формы для формул алгебры высказываний

В § 3 мы уже отмечали, что операции импликации и эквиваленции можно выразить через остальные логические операции — отрицание, конъюнкцию и дизъюнкцию. Напомним эти выражения:

$$(X \Rightarrow Y) \cong (\bar{X} \vee Y),\quad (X \Leftrightarrow Y) \cong (\bar{X} \vee Y) \wedge (\bar{Y} \vee X)$$

(см. примеры 1 и 2 п. 4 § 3). Отсюда вытекает, что для каждой формулы алгебры высказываний можно указать равносильную ей формулу, содержащую лишь операции отрицания, конъюнкции и дизъюнкций. Например, формула

$$X \wedge (X \Rightarrow Y) \tag{1}$$

может быть преобразована так:

$$(X \wedge (X \Rightarrow Y)) \cong (X \wedge (\bar{X} \vee Y)) \cong ((X \wedge \bar{X}) \vee (X \wedge Y)) \cong (\text{Л} \vee (X \wedge Y)) \cong$$
$$\cong (X \wedge Y).$$

Значит, формула (1) равносильна следующей формуле, не содержащей других операций, кроме отрицания, конъюнкции и дизъюнкции:

$$X \wedge Y.$$

Разумеется, для заданной формулы существует не одна; а бесчисленное множество равносильных ей формул, содержащих лишь операции $^-$, $\wedge$, $\vee$. Так, для формулы $X \wedge Y$ можно записать целый ряд равносильных ей формул, например, $(X \wedge X) \wedge Y$, $(X \wedge X \wedge X) \wedge Y$ и т. д. Однако среди всевозможных выражений данной формулы через операции $^-$, $\wedge$, $\vee$ существуют некоторые особые, играющие важную роль как в самой математической логике, так и в ее приложениях. Это так называемые совершенные нормальные формы.

1. Совершенная дизъюнктивная нормальная форма

Введем одно обозначение, которое будет полезно в дальнейшем. Положим

$$X^{\alpha} = \begin{cases} X, & \text{если } \alpha = 1; \\ \bar{X}, & \text{если } \alpha = 0. \end{cases}$$

В частности, $0^0 = 1$ и $1^1 = 1$, в то время как $0^1 = 0$ и $1^0 = 0$. Иначе говоря, $X^{\alpha} = 1$ тогда и только тогда, когда $X = \alpha$, и $X^{\alpha} = 0$ тогда и только тогда, когда $X \neq \alpha$.

Теорема 1. *Для всякой формулы $F(X_1, ..., X_n)$ алгебры высказываний справедливо следующее соотношение:*

$$F(X_1, ..., X_n) \cong \bigvee_{(\alpha_1, ..., \alpha_n)} F(\alpha_1, ..., \alpha_n) \wedge X_1^{\alpha_1} \wedge X_2^{\alpha_2} \wedge ... \wedge X_n^{\alpha_n}. \tag{2}$$

Поясним смысл выражения, стоящего в правой части (т. е. справа от знака $\cong$). Указанное выражение означает, что мы должны рассмотреть всевозможные наборы значений $\alpha_1, ..., \alpha_n$ (каждое α_i есть либо 0, либо 1), для каждого набора найти $F(\alpha_1, ..., \alpha_n)$ (значение истинности формулы F при $X_1 = \alpha_1, ..., X_n = \alpha_n$), построить конъюнктивный «одночлен» $F(\alpha_1, ..., \alpha_n) \wedge \wedge X_1^{\alpha_1} \wedge X_2^{\alpha_2} \wedge ... \wedge X_n^{\alpha_n}$ и, наконец, составить дизъюнкцию таких «одночленов» для всевозможных наборов значений $\alpha_1, ..., \alpha_n$.

Например, для формулы $F(X_1, X_2) : X_1 \Rightarrow X_2$ выражение, стоящее в правой части соотношения (2), имеет вид

$$(F(1,1) \wedge X_1 \wedge X_2) \vee (F(1,0) \wedge X_1 \wedge \bar{X}_2) \vee (F(0,1) \wedge \bar{X}_1 \wedge X_2) \vee$$
$$\vee (F(0,0) \wedge \bar{X}_1 \wedge \bar{X}_2).$$

Учитывая, что $F(1,1) = 1$, $F(1,0) = 0$, $F(0,1) = 1$, $F(0,0) = 1$, получим выражение

$$(X_1 \wedge X_2) \vee (\bar{X}_1 \wedge X_2) \vee (\bar{X}_1 \wedge \bar{X}_2).$$

□ Для упрощения записей будем считать $n = 2$, т. е. рассмотрим формулу $F(X_1, X_2)$. Чтобы установить правильность соотношения (2), нужно проверить, что при любом наборе (логических) значений $X_1=\xi_1$, $X_2=\xi_2$ левая и правая части этого соотношения принимают одинаковые значения.

Пусть $X_1=\xi_1$, $X_2=\xi_2$. Значение левой части соотношения (2) есть $F(\xi_1,\xi_2)$. Найдем значение правой части. Одночлен $F(\alpha_1,\alpha_2)\wedge X_1^{\alpha_1}\wedge X_2^{\alpha_2}$ принимает значение $F(\alpha_1,\alpha_2)\wedge\xi_1^{\alpha_1}\wedge\xi_2^{\alpha_2}$. Последнее заведомо равно 0, если $\xi_1\neq\alpha_1$ или $\xi_2\neq\alpha_2$ (тогда $\xi_1^{\alpha_1}=0$ или $\xi_2^{\alpha_2}=0$). Таким образом, из всех одночленов правой части отличным от нуля (при $X_1=\xi_1$ и $X_2=\xi_2$) может быть только тот, в котором $\alpha_1=\xi_1$, $\alpha_2=\xi_2$, т. е. $F(\xi_1,\xi_2)\wedge 1\wedge 1=F(\xi_1,\xi_2)$. ■

Определение 1. *Совершенным конъюнктивным одночленом* от переменных $X_1, X_2, \ldots, X_n$ называется выражение вида

$$X_1^{\alpha_1}\wedge X_2^{\alpha_2}\wedge\ldots\wedge X_n^{\alpha_n},$$

где каждое α_i ($i=1,\ldots,n$) есть либо 1, либо 0. *Совершенной дизъюнктивной нормальной формой* (сокращенно СДНФ) называется формула вида

$$K_1\vee K_2\vee\ldots\vee K_s,$$

где каждая из формул $K_1,\ldots,K_s$ представляет собой совершенный конъюнктивный одночлен.

Примером совершенной дизъюнктивной нормальной формы от двух переменных X_1, X_2 может служить выражение

$$(X_1\wedge X_2)\vee(X_1\wedge\bar{X}_2)\vee(\bar{X}_1\wedge X_2).$$

Условимся называть формулу $F(X_1,\ldots,X_n)$ *тождественно ложной*, если при любых значениях для $X_1,\ldots,X_n$ высказывание $F(X_1,\ldots,X_n)$ является ложным.

Теорема 2. *Всякая не тождественно ложная формула F алгебры высказываний равносильна некоторой совершенной дизъюнктивной нормальной форме* (или, как мы будем дальше говорить, имеет некоторую совершенную дизъюнктивную нормальную форму).

□ Доказательство непосредственно следует из теоремы 1. Действительно, если формула F не тождественно ложна, то среди ее значений $F(\alpha_1,\ldots,\alpha_n)$ некоторые обязательно равны 1; для таких значений одночлен $F(\alpha_1,\ldots,\alpha_n)\wedge\wedge X_1^{\alpha_1}\wedge\ldots\wedge X_n^{\alpha_n}$ есть $1\wedge X_1^{\alpha_1}\wedge\ldots\wedge X_n^{\alpha_n}$, что равносильно конъюнктивному одночлену $X_1^{\alpha_1}\wedge\ldots\wedge X_n^{\alpha_n}$. Если же какое-то значение $F(\alpha_1,\ldots,\alpha_n)$ равно нулю, то соответствующий одночлен есть $0\wedge X_1^{\alpha_1}\wedge\ldots\wedge X_n^{\alpha_n}$, что тождественно равно нулю. ■

Приведенное доказательство дает одновременно и способ нахождения совершенной дизъюнктивной нормальной формы для данной формулы: *нужно выделить все наборы значений $\alpha_1,\ldots,\alpha_n$, для которых $F(\alpha_1,\ldots,\alpha_n)=1$, затем для каждого такого набора построить конъюнктивный одночлен $X_1^{\alpha_1}\wedge\ldots\wedge X_n^{\alpha_n}$ и, наконец, составить дизъюнкцию всех таких одночленов.*

Пример. Для формулы $\bar{X}\Rightarrow\bar{Y}$ составить совершенную дизъюнктивную нормальную форму.

○ Таблица истинности для данной формулы имеет вид

X	Y	$\bar{X}\Rightarrow\bar{Y}$
1	1	1
1	0	1
0	1	0
0	0	1

Как видно из этой таблицы, формула принимает значение 1 для следующих наборов значений переменных: $X=1$, $Y=1$; $X=1$, $Y=0$; $X=0$, $Y=0$. Отсюда получаем следующую дизъюнктивную нормальную форму:

$$(X \wedge Y) \vee (X \wedge \bar{Y}) \vee (\bar{X} \wedge \bar{Y}).$$

Итак,

$$(\bar{X} \to \bar{Y}) \cong (X \wedge Y) \vee (X \wedge \bar{Y}) \vee (\bar{X} \wedge \bar{Y}). \quad \bullet$$

2. Совершенная конъюнктивная нормальная форма

Понятия и теоремы этого пункта носят, как говорят, *двойственный характер* по отношению к соответствующим понятиям и теоремам из предыдущего пункта.

Теорема 1′. *Для всякой формулы $F(X_1, \ldots, X_n)$ алгебры высказываний справедливо следующее соотношение:*

$$F(X_1, \ldots, X_n) \cong \bigwedge_{(\alpha_1, \ldots, \alpha_n)} F(\alpha_1, \ldots, \alpha_n) \vee X_1^{\alpha_1} \vee X_2^{\alpha_2} \vee \ldots \vee X_n^{\alpha_n}. \quad (2)$$

Определение 1′. *Совершенным дизъюнктивным одночленом* от переменных $X_1, X_2, \ldots, X_n$ называется выражение вида

$$X_1^{\alpha_1} \vee X_2^{\alpha_2} \vee \ldots \vee X_n^{\alpha_n},$$

где каждое α_i ($i = 1, \ldots, n$) есть либо 1, либо 0. *Совершенной конъюнктивной нормальной формой* (сокращенно СКНФ) называется формула вида

$$D_1 \wedge D_2 \wedge \ldots \wedge D_t,$$

где каждая из формул $D_1, \ldots, D_t$ представляет собой совершенный дизъюнктивный одночлен.

Теорема 2′. *Всякая не тождественно истинная формула алгебры высказываний равносильна некоторой совершенной конъюнктивной нормальной форме.*

Мы ограничимся лишь формулировкой теорем 1′ и 2′, поскольку их доказательства аналогичны доказательствам теорем 1 и 2.

3. Нахождение всех следствий из данных посылок

Сведения, которыми мы теперь располагаем, позволяют решить следующий важный вопрос. Пусть даны несколько формул $F_1, \ldots, F_s$. Как найти все без исключения формулы, являющиеся логическими следствиями посылок $F_1, \ldots, F_s$?

Прежде всего заметим, что можно ограничиться случаем одной посылки ($s = 1$). Действительно, формула H является следствием формул $F_1, \ldots, F_s$ тогда и только тогда, когда H есть следствие (одной-единственной) формулы $F_1 \wedge F_2 \wedge \ldots \wedge F_s$.

Ответ на поставленный вопрос дает следующее утверждение.

Теорема 3. *Пусть каждая из формул $H(X_1, \ldots, X_n)$ и $F(X_1, \ldots, X_n)$ не является тождественно ложной. Тогда H есть следствие F тогда и только тогда, когда все конъюнктивные одночлены, входящие в совершенную дизъюнктивную нормальную форму для F, входят и в совершенную дизъюнктивную форму для H* (или, говоря проще, когда совершенная дизъюнктивная нормальная форма для F является «частью» аналогичной формы для H).

□ Пусть H — следствие F. Тогда наборы значений $X_1 = \alpha_1, ..., X_n = \alpha_n$, для которых $F = 1$, содержатся среди наборов, для которых $H = 1$. Если принять во внимание способ построения СДНФ, то это и означает, что СДНФ для F является «частью» СДНФ для H.

Обратно, пусть СДНФ для F является «частью» СДНФ для H. Согласно способу построения СДНФ, это означает, что при любом наборе значений $X_1 = \alpha_1, ..., X_n = \alpha_n$, для которого $F = 1$, имеет место и $H = 1$, т. е. H является следствием F. ■

Пример. Найти все следствия из посылок X и $X \Rightarrow Y$.

○ Поставленная задача равнозначна следующей: найти все следствия из формулы F: $X \wedge (X \Rightarrow Y)$.

Составим для F таблицу истинности:

X	Y	$X \wedge (X \Rightarrow Y)$
1	1	1
1	0	0
0	1	0
0	0	0

Как показывает эта таблица, СДНФ для F имеет вид $X \wedge Y$. Поэтому следствиями формулы F являются всевозможные формулы, для которых СДНФ содержит одночлен $X \wedge Y$.

Всего имеется четыре совершенных конъюнктивных одночлена от X и Y:

$$K_1 : X \wedge Y,\ K_2 : \bar{X} \wedge Y,\ K_3 : X \wedge \bar{Y},\ K_4 : \bar{X} \wedge \bar{Y}.$$

Поэтому любое следствие формулы F есть либо K_1, либо формула вида $K_1 \vee ...$, где многоточие обозначает любую дизъюнктивную сумму, которую можно составить из K_2, K_3, K_4. Таким образом, получаем 8 различных следствий из F:

$$K_1,\ K_1 \vee K_2,\ K_1 \vee K_3,\ K_1 \vee K_4,\ K_1 \vee K_2 \vee K_3,\ K_1 \vee K_2 \vee K_4,\ K_1 \vee K_3 \vee K_4,$$
$$K_1 \vee K_2 \vee K_3 \vee K_4.$$

Последняя из этих формул — тождественно истинная. ●

§ 6. Предикаты

1. Понятие предиката. Область определения и область истинности предиката

Рассмотрим предложение «$x^2 = 9$». Ясно, что это предложение не является высказыванием: о его истинности или ложности невозможно судить до тех пор, пока не будет указано, какое число подразумевается под x. При одних значениях x (например, при $x = 3$) оно превращается в истинное высказывание, при других — в ложное.

Аналогично обстоит дело, скажем, с предложением «Город x находится в СССР». Оно превращается в истинное высказывание, если вместо x подставить, например, названия городов «Москва» или «Владивосток», но оно будет ложным высказыванием, если вместо x подставить «Токио» или «Париж».

Определение 1. Предложение, в которое входят переменные и которое при замене переменных возможными для них значениями становится высказыванием, называется *высказывательной формой* (или *предикатом*).

При задании предиката должно указываться множество X тех значений, которые могут принимать переменные; оно называется *областью определения предиката.*

Если предикат содержит лишь одну переменную, то он называется *одноместным*; при наличии n переменных предикат называется *n-местным.* Например, предложение «Река x впадает в Каспийское море» — одноместный предикат, а предложение «Река x впадает в море y» — двуместный. Предложение «$\sin^2 x + \cos^2 x = 1$» есть одноместный предикат, а «$x^2 + y^2 = 1$» — двуместный.

Пусть X — область определения предиката.

Определение 2. Подмножество множества X, состоящее из тех значений переменных, при которых данный предикат превращается в истинное высказывание, называется *областью истинности предиката.*

В дальнейшем одноместные предикаты с переменной x будем обозначать через $P(x)$, $Q(x)$, ...; с двумя переменными x, y — через $P(x, y)$, $Q(x, y)$, ... Впрочем, для сокращения записей вместо $P(x)$, $P(x, y)$, ... часто будем писать просто P. Множество истинности предиката P будем обозначать той же буквой, что и сам предикат, но с верхним индексом $^+$, т. е. P^+.

Примеры. 1. Уравнение «$x^2 + 5x + 6 = 0$», где x является действительным числом, — одноместный предикат, определенный на множестве $\mathbb{R}$ действительных чисел. Его множество истинности состоит из двух чисел -2 и -3 (корней уравнения); следовательно, обозначив данный предикат через $P(x)$, можем записать $P^+ = \{-2, -3\}$.

2. Уравнение «$x^2 + y^2 = z^2$», где x, y, z являются целыми числами, представляет собой трехместный предикат на множестве Z всех целых чисел. Множество истинности этого предиката состоит из всех троек x, y, z целых чисел, для которых выполняется равенство $x^2 + y^2 = z^2$. Например, в это множество входят тройки (3, 4, 5), (5, 12, 13) и т. д.

3. Неравенство «$|x| > 3$» есть одноместный предикат на множестве $\mathbb{R}$ действительных чисел. Его множество истинности — объединение двух лучей $(3, +\infty)$ и $(-\infty, -3)$.

2. Тождественно истинные и тождественно ложные предикаты. Равносильность и следование предикатов

Определение 3. Предикат с областью определения X называется: *тождественно истинным*, если при любых значениях переменных из X он превращается в истинное высказывание; *тождественно ложным*, если при любых значениях переменных из X он превращается в ложное высказывание.

Примером тождественно истинного предиката может служить предикат «$\sin^2 x + \cos^2 x = 1$», определенный на множестве $\mathbb{R}$ действительных чисел, или предикат «$(x+y)^2 = x^2 + 2xy + y^2$», определенный для действительных значений x и y. Напротив, предикат «$x^2 + y^2 < 0$», где x и y — любые действительные числа, тождественно ложен.

Определение 4. Два предиката с одной и той же областью определения X называются *равносильными*, если они имеют одинаковые множества истинности. Равносильность предикатов P и Q обозначается так: $P \cong Q$.

Переход от некоторого предиката P к равносильному ему предикату Q называется *равносильным преобразованием* предиката P. Понятие равносильности играет важную роль при рассмотрении уравнений, поскольку каждое уравнение представляет собой некоторый предикат, а задача решения уравнения — задачу нахождения множества истинности соответствующего предиката. В процессе решения мы производим над уравнением различные преобразования; при этом важно, чтобы получаемый предикат был равносилен исходному, т. е. чтобы область истинности не изменялась. Аналогичное соображение относится и к решению неравенств, содержащих неизвестные.

Приведем пример цепочки равносильных преобразований в случае уравнения:

$$x^2=9 \cong x^2-9=0 \cong (x-3)(x+3)=0 \cong (x=3) \vee (x=-3).$$

Следовательно, множество истинности предиката $x^2=9$ есть $\{3, -3\}$.

Определение 5. Пусть P и Q — два предиката с общей областью определения X. Говорят, что Q есть *следствие* P, если область истинности предиката P является частью области истинности предиката Q (или совпадает с ней), т. е. если $P^+ \subset Q^+$.

Примеры. 1. Предикат «$x^2=25$» с областью определения $\mathbf{R}$ есть следствие предиката «$x=5$» (с той же областью определения). Эти два предиката не равносильны: множество истинности первого из них имеет вид $\{-5, 5\}$, а второго — $\{5\}$.

2. Предикат Q: «Углы треугольника x равны в некотором порядке углам треугольника y» есть следствие предиката P: «Стороны треугольника x параллельны в некотором порядке сторонам треугольника y». В то же время эти предикаты не равносильны, поскольку существуют треугольники с соответственно равными углами, но расположенные так, что стороны одного не параллельны ни в каком порядке сторонам другого.

Непосредственно из определений равносильности и следования вытекает следующее предложение: предикат P равносилен предикату Q тогда и только тогда, когда Q есть следствие P, а P есть следствие Q.

3. Логические операции над предикатами

Над предикатами можно производить те же логические операции, что и над высказываниями: отрицание, конъюнкцию, дизъюнкцию, импликацию, эквиваленцию.

Для определенности будем рассматривать предикаты с одной переменной x.

Определение 6. *Отрицанием* предиката $P(x)$ с областью определения X называется предикат с той же областью определения, обозначаемый $\bar{P}(x)$ (читается: «Неверно, что $P(x)$»), который превращается в истинное высказывание для тех и только тех значений $x \in X$, при которых $P(x)$ есть ложное высказывание.

Очевидно, множество истинности предиката $\bar{P}(x)$ представляет собой дополнение множества истинности предиката $P(x)$ (имеется в виду, конечно, дополнение до X).

Примеры. 1. Отрицанием предиката «$x \geqslant 0$» с областью определения $\mathbf{R}$ служит предикат «$x<0$» с той же областью определения. Последний предикат можно записать и по-другому: например, «$x^3<0$» или «$e^x<1$» (равносильные записи).

2. Отрицанием предиката «$|x|>3$» с областью определения $\mathbf{R}$ является предикат «$|x| \leqslant 3$» с той же областью определения. Равносильная запись последнего предиката: «$-3 \leqslant x \leqslant 3$».

Определение 7. *Конъюнкцией* двух предикатов $P(x)$ и $Q(x)$, определенных соответственно на множествах X_1 и X_2, называется третий предикат, обозначаемый $P(x) \wedge Q(x)$ (читается: «$P(x)$ и $Q(x)$») с областью определения $X=X_1 \cap X_2$, который превращается в истинное высказывание для тех и только тех значений $x \in X$, при которых оба исходных предиката являются истинными высказываниями.

Множество истинности предиката $P(x) \wedge Q(x)$ представляет собой пересечение множеств истинности для предикатов $P(x)$ и $Q(x)$ (т. е. $P^+ \cap Q^+$).

Пример 3. Конъюнкцией предикатов «$x<2$» и «$x>-2$» (с областью определения $\mathbf{R}$) является предикат «$-2<x<2$» или, в эквивалентной записи, «$|x|<2$».

Понятие конъюнкции предикатов тесно связано с понятием системы уравнений. В самом деле, рассмотрим, например, систему двух уравнений с двумя неизвестными:

$$\begin{cases} F(x, y) = 0, \\ \Phi(x, y) = 0. \end{cases} \qquad (1)$$

Задача решения такой системы есть, по существу, задача нахождения множества истинности для предиката, являющегося конъюнкцией двух предикатов: «$F(x, y) = 0$» и «$\Phi(x, y) = 0$». В этом смысле часто говорят, что система (1) есть *конъюнкция уравнений* $F(x, y) = 0$ и $\Phi(x, y) = 0$.

Определение 8. *Дизъюнкцией* двух предикатов $P(x)$ и $Q(x)$, имеющих соответственно области определения X_1 и X_2, называется третий предикат, обозначаемый $P(x) \vee Q(x)$ (читается: «$P(x)$ или $Q(x)$»), с областью определения $X = X_1 \cap X_2$, который превращается в истинное высказывание для тех и только тех значений $x \in X$, при которых хотя бы одно из высказываний $P(x)$ и $Q(x)$ истинно.

Множество истинности предиката $P(x) \vee Q(x)$ есть объединение множеств истинности для предикатов $P(x)$ и $Q(x)$ (т. е. $P^+ \cup Q^+$).

Пример 4. Дизъюнкцией предикатов «$x > 2$» и «$x < -2$» (с областью определения $\mathbb{R}$) является предикат «$(x > 2) \vee (x < -2)$» или, что равносильно, «$|x| > 2$».

Понятие дизъюнкции предикатов тесно связано с уравнениями. Рассмотрим, например, уравнение вида

$$F(x)\,\Phi(x) = 0, \qquad (2)$$

левая часть которого есть произведение двух функций $F(x)$ и $\Phi(x)$. Так как произведение двух чисел равно нулю тогда и только тогда, когда равен нулю один из сомножителей, то множество корней уравнения (2) есть объединение множеств корней уравнений $F(x) = 0$ и $\Phi(x) = 0$. Другими словами, множество корней уравнения (2) совпадает с множеством истинности предиката $(F(x) = 0) \vee (\Phi(x) = 0)$. В этом смысле говорят, что уравнение (2) есть *дизъюнкция уравнений* $F(x) = 0$ и $\Phi(x) = 0$.

Имеется также связь между дизъюнкцией предикатов и неравенствами, содержащими неизвестные. Например, чтобы решить неравенство вида

$$F(x)\,\Phi(x) > 0, \qquad (3)$$

нужно объединить множества решений двух систем

$$\begin{cases} F(x) > 0, \\ \Phi(x) > 0, \end{cases} \qquad \begin{cases} F(x) < 0, \\ \Phi(x) < 0. \end{cases}$$

Это означает, что множество всех решений неравенства (3) есть множество истинности предиката, являющегося дизъюнкцией двух предикатов

$$(F(x) > 0) \wedge (\Phi(x) > 0), \quad (F(x) < 0) \wedge (\Phi(x) < 0).$$

Так, неравенство $(x-2)(x+1) > 0$ представляет собой дизъюнкцию двух систем

$$\begin{cases} x-2 > 0, \\ x+1 > 0, \end{cases} \qquad \begin{cases} x-2 < 0, \\ x+1 < 0, \end{cases}$$

первая из которых равносильна неравенству $x > 2$, а вторая — неравенству $x < -1$. Множество решений исходного неравенства есть объединение лучей $(-\infty, -1)$ и $(2, +\infty)$.

Определение 9. Пусть $P(x)$ и $Q(x)$ — два предиката, определенные соответственно на множествах X_1 и X_2. *Импликацией* $P(x)$ и $Q(x)$ называется третий предикат, обозначаемый $P(x) \Rightarrow Q(x)$ (читается: «Если $P(x)$, то $Q(x)$»), с областью определения $X = X_1 \cap X_2$, который превращается в ложное высказывание для тех и только тех значений $x \in X$, при которых первый предикат является истинным высказыванием, а второй — ложным.

Как следует из данного определения, предикат $P(x) \Rightarrow Q(x)$ превращается в ложное высказывание для тех и только тех значений $x \in X$, при которых высказывание $P(x)$ — истинное, а $Q(x)$ — ложное. Иначе говоря, множество истинности предиката $P(x) \Rightarrow Q(x)$ есть дополнение (до X) множества $P^+ \cap Q^-$ (где Q^- — дополнение к Q^+), т. е. множество $P^- \cup Q^+$.

Предоставляем самостоятельно убедиться в справедливости следующего предложения: если предикат $Q(x)$ есть следствие предиката $P(x)$, то предикат $P(x) \Rightarrow Q(x)$ тождественно истинен; обратно, если предикат $P(x) \Rightarrow Q(x)$ тождественно истинен, то $Q(x)$ есть следствие $P(x)$.

Определение 10. *Эквиваленцией* предикатов $P(x)$ и $Q(x)$, имеющих соответственно области определения X_1 и X_2, называется предикат, обозначаемый $P(x) \Leftrightarrow Q(x)$ (читается: «$P(x)$ тогда и только тогда, когда $Q(x)$»), который превращается в истинное высказывание для тех и только тех значений $x \in X_1 \cap X_2$, при которых оба данных предиката превращаются одновременно либо в истинные, либо в ложные высказывания.

Множество истинности предиката $P(x) \Leftrightarrow Q(x)$ является объединением двух множеств: $P^+ \cap Q^+$ и $P^- \cap Q^-$.

§ 7. Кванторные операции над предикатами

Рассмотренные в предыдущем параграфе логические операции над предикатами в определенном смысле аналогичны соответствующим операциям над высказываниями. Однако поскольку понятие предиката существенно богаче, чем понятие высказывания (предикат превращается в высказывание лишь при фиксированных значениях входящих в него переменных), для предикатов можно ввести операции, не имеющие аналогов среди операций над высказываниями. Это — так называемые *кванторные операции*. Каждая из них применяется к одноместному предикату и превращает его в высказывание.

Определение 1. Операцией *квантор общности* называется правило, которое каждому одноместному предикату $P(x)$, определенному на множестве X, сопоставляет высказывание, обозначаемое $(\forall x)(P(x))$ (читается: «Для всякого x справедливо $P(x)$»), которое истинно тогда и только тогда, когда предикат $P(x)$ тождественно истинен*.

Сам символ $\forall$ также называют квантором общности, а присоединение его к предикату $P(x)$ часто называют «навешиванием» квантора общности на предикат $P(x)$.

Пример. Пусть $P(x)$ есть предикат «$x^2 > 0$», определенный на множестве $\mathbb{R}$ действительных чисел. Тогда $(\forall x)(P(x))$ является ложным высказыванием, поскольку неравенство $x^2 > 0$ справедливо не для всех $x \in \mathbb{R}$ (оно неверно при $x = 0$).

В случае, когда множество X, на котором определен предикат $P(x)$, — конечное и состоит из элементов $a_1, a_2, \ldots, a_n$, высказывание $(\forall x)(P(x))$ рав-

* Символ $\forall$ происходит от первой буквы английского слова All («все»).

носильно конъюнкции всех высказываний $P(a_1), \ldots, P(a_n)$; иначе говоря, высказывание $(\forall x)(P(x))$ имеет то же значение истинности, что и высказывание $P(a_1) \wedge P(a_2) \wedge \ldots \wedge P(a_n)$.

Замечание. В высказывании $(\forall x)(P(x))$ переменная x уже перестает быть переменной в обычном смысле этого слова: вместо нее уже нельзя подставлять конкретные элементы из множества X. Как говорят в подобных случаях, переменная x — *связанная* (кажущаяся, немая). Вместо x можно употребить любую другую букву, например y; другими словами, высказывания $(\forall x)(P(x))$ и $(\forall y)(P(y))$ одинаковы.

Определение 2. Операцией *квантор существования* называется правило, которое каждому одноместному предикату $P(x)$, определенному на множестве X, сопоставляет высказывание, обозначаемое $(\exists x)(P(x))$ (читается: «Существует значение $x \in X$, такое, что верно $P(x)$»), которое ложно тогда и только тогда, когда предикат $P(x)$ тождественно ложен*.

Пример. Пусть $P(x)$ есть предикат «x — простое число», определенный на множестве целых чисел. Тогда $(\exists x)(P(x))$ является истинным высказыванием.

Из определений кванторов общности и существования вытекают следующие две равносильности:

$$\overline{(\forall x)(P(x))} \cong (\exists x)\overline{(P(x))}, \quad \overline{(\exists x)(P(x))} \cong (\forall x)\overline{(P(x))}.$$

Поясним первое из этих утверждений (второе поясняется аналогично). Высказывание, стоящее слева от знака $\cong$, истинно тогда и только тогда, когда $P(x)$ верно не для всех $x \in X$; высказывание же, стоящее справа от знака $\cong$, истинно тогда и только тогда, когда существует значение $x \in X$, при котором $P(x)$ неверно. Отсюда вытекает, что левая и правая части истинны или ложны одновременно.

«Навешивание» кванторов возможно и в случае многоместных предикатов. При этом каждая переменная должна быть связана только одним квантором. Например, в случае двуместного предиката $P(x, y)$ можно рассмотреть следующие высказывания: $(\exists x)(\exists y)(P(x, y))$, $(\exists x)(\forall y)(P(x, y))$, $(\forall x)(\exists y)(P(x, y))$, $(\forall x)(\forall y)(P(x, y))$ (Первое из этих высказываний читается так: «Существуют такое x и такое y, при которых справедливо $P(x, y)$»; аналогично читаются остальные три высказывания.)

Нетрудно видеть, что *одноименные кванторы можно менять местами:* это означает, например, что высказывания $(\forall x)(\forall y)(P(x, y))$ и $(\forall y)(\forall x)(P(x, y))$ равносильны, т. е. истинны или ложны одновременно. Что же касается *разноименных кванторов* ($\forall x$ и $\exists y$ или $\exists x$ и $\forall y$), *то при их перестановке может получиться высказывание, не равносильное исходному.* Например, в случае предиката $P(x, y)$: «Река x впадает в море y» высказывание $(\forall x)(\exists y)(P(x, y))$ истинно, так как для всякой реки существует море, в которое она впадает; в то же время высказывание $(\exists y)(\forall x)(P(x, y))$ ложно, поскольку оно означает, что существует море, в которое впадают все реки.

Для более твердого усвоения кванторных операций рекомендуем самостоятельно проверить следующие равносильности:

$$(\forall x)(P(x)) \cong \overline{(\exists x)(\overline{P(x)})}; \; (\exists x)(P(x)) \cong \overline{(\forall x)(\overline{P(x)})};$$

$$(\forall x)(P(x)) \wedge Q(x) \cong (\forall x)(P(x)) \wedge (\forall y)(Q(y));$$

$$(\exists x)(P(x)) \vee Q(x)) \cong (\exists x)(P(x)) \vee (\exists y)(P(y)).$$

* Символ $\exists$ происходит от первой буквы английского слова Existence («существование»).

2

Действительные числа

§ 8. Положительные действительные числа

1. Измерение отрезков. Определение положительного действительного числа

Для чего нужны числа? В самой общей форме ответ на этот вопрос может быть сформулирован так: чтобы фиксировать результаты измерений (количества предметов, размеры фигур, массы тел и т. д.).

Вспомним, в частности, как осуществляется измерение отрезков.

Пусть требуется измерить некоторый отрезок AB с помощью отрезка OE, длина которого принимается равной 1.

Для этого на отрезке AB, начиная от точки A, откладываем отрезок OE столько раз, чтобы полученный остаток A_1B был меньше 1 (рис. 1). Пусть α_0 — число таких откладываний. Тогда целое число α_0 выражает результат измерения с точностью до 1.

Чтобы измерить отрезок AB с большей точностью, берем отрезок, равный $1/10$ единицы измерения, и откладываем его на отрезке A_1B (от точки A_1) до тех пор, пока не получим остаток A_2B, меньший $1/10$. Пусть α_1 — число таких откладываний; очевидно, $0 \leqslant \alpha_1 \leqslant 9$. Тогда десятичная дробь α_0, α_1 выражает результат измерения отрезка AB с точностью до $1/10$.

Этот процесс можно продолжить. В результате будем получать числа

$$\alpha_0;\ \alpha_0, \alpha_1;\ \ \alpha_0, \alpha_1\alpha_2;\ \ldots,$$

характеризующие результат измерения с точностью до 1, $1/10$, $1/10^2$ и т. д.

Имеются две возможности:

1. При откладывании на отрезке A_nB очередной доли единичного отрезка, равной $1/10^n$, впервые не получится никакого остатка. Тогда десятичная дробь $\alpha_0, \alpha_1\alpha_2 \ldots \alpha_n$ уже не приближенно, а точно выражает длину отрезка AB.

2. Ни один из остатков A_1B, A_2B, A_3B, ... не равен нулю. Тогда естественно считать, что длина отрезка AB выражается бесконечной дробью $\alpha_0, \alpha_1\alpha_2 \ldots \alpha_n \ldots$.

Например, если указанным способом измерить отрезок, равный $1/3$, то получим бесконечную десятичную дробь 0,333 ... (содержащую 3 в периоде).

Нетрудно установить, что при описанном способе измерения отрезков не может появиться бесконечная десятичная дробь с «хвостом» из девяток, скажем, дробь 0,2999... . Действительно, появление такой дроби означало бы, что в первом из остатков A_1B отрезок $1/10$ укладывается на самом деле не 2,

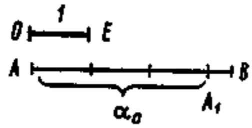


Рис. 1

а ровно 3 раза, вследствие чего результат измерения должен выражаться числом 0,3.

Определение 1. Любая бесконечная десятичная дробь $\alpha_0, \alpha_1\alpha_2 \dots \alpha_n \dots$, где $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n, \dots$ — неотрицательные целые числа, не равные одновременно нулю, причем $\alpha_i \leqslant 9$ для всех $i = 1, 2, \dots$, называется *положительным действительным числом.*

Что касается конечных десятичных дробей, то условимся каждую такую дробь $\alpha_0, \alpha_1\alpha_2 \dots \alpha_n$, где $\alpha_n \neq 0$, отождествлять с любой из двух бесконечных десятичных дробей:

$$\alpha_0, \alpha_1\alpha_2 \dots \alpha_n 00 \dots \text{ и } \alpha_0, \alpha_1\alpha_2 \dots (\alpha_n - 1)99 \dots .$$

Например, $2{,}5 = 2{,}500 \dots = 2{,}499 \dots$.

Определение 2. Два положительных действительных числа $a = \alpha_0, \alpha_1\alpha_2 \dots$ и $b = \beta_0, \beta_1\beta_2 \dots$ будем считать *равными*, если имеет место один из двух случаев:

1) $\alpha_0 = \beta_0,\ \alpha_1 = \beta_1,\ \alpha_2 = \beta_2, \dots$;

2) для некоторого номера n справедливы равенства $\alpha_0 = \beta_0,\ \alpha_1 = \beta_1, \dots, \alpha_n = \beta_n,\ \alpha_{n+1} = \beta_{n+1} + 1$, причем «хвост» числа a после цифры α_{n+1} состоит из нулей, а «хвост» числа b — из девяток.

Например, $12{,}3700 \dots = 12{,}3699 \dots$.

Очевидно, такое определение равенства полностью согласуется с принятым выше соглашением о конечных десятичных дробях.

2. Сравнение положительных действительных чисел

Пусть снова $a = \alpha_0, \alpha_1\alpha_2 \dots$ и $b = \beta_0, \beta_1\beta_2 \dots$ — два положительных действительных числа. Предположим дополнительно, что каждая из написанных бесконечных десятичных дробей не имеет «хвоста» из девяток.

Определение 3. Условимся считать, что $a > b$, если имеет место один из случаев:

1) $\alpha_0 > \beta_0$;

2) $\alpha_0 = \beta_0$, но $\alpha_1 > \beta_1$;

3) $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$, но $\alpha_2 > \beta_2$;

. . . .

Например: $12, \dots > 7, \dots$ (многоточия обозначают любые наборы цифры): $6{,}245 \dots > 6{,}241 \dots$.

Замечание. Условие относительно «хвоста» из девяток важно: оно гарантирует, что если $a > b$, то числа a и b не могут быть равными. Снятие этого условия сделало бы возможным одновременное выполнение неравенства $a > b$ и равенства $a = b$. Например, так получилось бы в случае чисел $a = 0{,}3000 \dots$, $b = 0{,}2999 \dots$ (здесь $\alpha_0 = \beta_0$, но $\alpha_1 > \beta_1$).

Можно показать (убедитесь в этом самостоятельно), что введенное выше отношение «$>$» обладает следующими свойствами:

1^0. *Для любых двух положительных действительных чисел a и b справедливо одно и только одно из соотношений $a = b$; $a > b$, $b > a$.*

2^0. *Если $a > b$ и $b > c$, то $a > c$.*

3^0. *Для любых положительных действительных чисел a и d справедливо неравенство $a + d > a$.*

4^0. *Если $a > b$, то найдется такое положительное число t, что $a > b + t$* (в качестве t может быть взято число $1/10^n$ при достаточно большом n).

Наряду с отношением «$>$» мы будем использовать и отношение «$<$». По определению, считаем $a < b$, если $b > a$.

3. Точная верхняя граница множества

Пусть M — некоторое множество положительных действительных чисел. Будем говорить, что множество M *ограничено сверху*, если существует такое действительное число t, что $a \leqslant t$ для любого $a \in M$. В этом случае число t называется *верхней границей множества M* (или верхней границей для множества M).

Разумеется, если число t является верхней границей для M, то любое число t', большее t, также является верхней границей для M.

Определение 4. Пусть множество M ограничено сверху. Тогда наименьшая из всех его верхних границ называется *точной верхней границей* (или *верхней гранью*) для множества M.

Теорема. *Всякое ограниченное сверху множество положительных действительных чисел имеет точную верхнюю границу.*

□ Среди верхних границ множества M существуют, очевидно, и целые числа. Пусть β_0 — наименьшее среди таких чисел (очевидно, $\beta_0 > 0$). Таким образом, целое число β_0 является верхней границей для M, а целое число $\beta_0 - 1$ не является.

Положим $\alpha_0 = \beta_0 - 1$ и разделим отрезок $[\alpha_0, \beta_0]$ (длины 1) на 10 равных частей, т. е. рассмотрим числа

$$\alpha_0, \alpha_0 + \frac{1}{10}, \ \alpha_0 + \frac{2}{10}, \ldots, \alpha_0 + \frac{9}{10}, \ \beta_0.$$

Первое из них, как мы уже отмечали, не является верхней границей для M, а последнее — является. Перебирая эти числа слева направо, найдем среди них такие два соседних числа α_0, α_1 и $\alpha_0, \alpha_1 + \frac{1}{10}$, что меньшее из них не является верхней границей для M, а большее — является.

Снова разделим отрезок между числами α_0, α_1 и $\alpha_0, \alpha_1 + \frac{1}{10}$ на 10 равных частей. Повторяя прежнее рассуждение, найдем два числа $\alpha_0, \alpha_1\alpha_2$ и $\alpha_0, \alpha_1\alpha_2 + \frac{1}{10^2}$, из которых меньшее не является верхней границей для M, а большее — является.

Продолжая этот процесс, в результате получим бесконечную десятичную дробь $\alpha_0, \alpha_1\alpha_2 \ldots$, иначе говоря, некоторое действительное число a. Покажем, что это число и является точной верхней границей для M.

Прежде всего отметим следующее. Пусть $a_n = \alpha_0, \alpha_1\alpha_2 \ldots \alpha_n$. По построению число $a_n + \frac{1}{10^n}$ при любом n является верхней границей для M. Может ли оказаться, что при некотором n это число принадлежит M? Разумеется, так может быть; но в этом случае число $a_n + \frac{1}{10^n}$ будет наибольшим числом в M, а значит, $a = a_n + \frac{1}{10^n}$ (все цифры $\alpha_{n+1}, \alpha_{n+2}, \ldots$ будут в этом случае равны 9). Итак, если для некоторого n имеем $a_n+\frac{1}{10^n}\in M$, то $a=a_n+\frac{1}{10^n}$.

Теперь докажем, что a является верхней границей для M. Пусть $b=\beta_0, \beta_1\beta_2 \dots$ — любое число из M. Если при некотором n имеем $b=a_n+\frac{1}{10^n}$, то, как уже отмечалось, $b=a$. Если же для всех n имеем $b<a_n+\frac{1}{10^n}$, то, очевидно, $\beta_0\leqslant \alpha_0$, $\beta_1\leqslant\alpha_1, \dots, \beta_{n-1}\leqslant\alpha_{n-1}$; поскольку такие неравенства имеют место при любом n, то $b\leqslant a$. Итак, в обоих случаях имеем $b\leqslant a$. Это доказывает, что a есть верхняя граница для M.

Пусть теперь $c=\gamma_0, \gamma_1\gamma_2 \dots$ — любая верхняя граница для M. Если при некотором n имеем $c\geqslant a_n+\frac{1}{10^n}$, то $c\geqslant a$. Если же для всех n выполняется неравенство $c<a_n+\frac{1}{10^n}$, то ввиду $a_n<c<a_n+\frac{1}{10^n}$ будет $\gamma_0=\alpha_0$, $\gamma_1=\alpha_1$, $\dots, \gamma_n=\alpha_n$, а значит (в силу произвольности n), $c=a$. В обоих случаях имеем $c\geqslant a$. Это доказывает, что число a является не просто верхней границей для M, но и точной верхней границей. ■

4. n-е приближение к положительному действительному числу

Пусть $a=\alpha_0, \alpha_1\alpha_2 \dots$ — положительное действительное число, записанное без «хвоста» из девяток. Конечную десятичную дробь $\alpha_0, \alpha_1\alpha_2 \dots \alpha_n$ условимся называть n-м *приближением к числу* a. Здесь n может принимать значения 0, 1, 2, ... В частности, нулевое приближение к a есть целое число α_0; оно называется целой *частью числа* a.

Будем обозначать n-е приближение к a через a_n. Очевидно, что всегда

$$a_n\leqslant a\leqslant a_n+\frac{1}{10^n}.$$

Отметим следующий факт, непосредственно вытекающий из определения знака «$>$» для действительных чисел: *если $a>b$, то $a_n\geqslant b_n$ для любого номера n.*

5. Действия над положительными действительными числами

Будем считать известными действия над конечными десятичными дробями. С их помощью определяются действия над положительными действительными числами.

Пусть a и b — два положительных действительных числа. Рассмотрим всевозможные конечные десятичные дроби a' и b' такие, что $a'<a$, $b'<b$. Множество M, состоящее из всевозможных чисел вида $a'+b'$, ограничено сверху. В самом деле, если a'' и b'' — какие-либо конечные десятичные дроби, для которых $a<a''$ и $b<b''$, то $a'+b'<a''+b''$, откуда видно, что число $a''+b''$ является верхней границей для множества M.

Определение 5. *Суммой* $a+b$ **называется точная верхняя граница множества чисел вида $a'+b'$, где a' и b' — любые две конечные десятичные дроби, такие, что $a'<a$, $b'<b$.**

Исходя из данного определения суммы, легко доказать, что если $a\leqslant c$ и $b\leqslant d$, то $a+b\leqslant c+d$, т. е. *неравенства одинакового смысла можно складывать почленно.*

В частности, складывая при любом n неравенства

$$a_n\leqslant a\leqslant a_n+\frac{1}{10^n}, \quad b_n\leqslant b\leqslant b_n+\frac{1}{10^n},$$

получим

$$a_n+b_n \leqslant a+b \leqslant a_n+b_n+\frac{2}{10^n}.$$

С ростом n эти неравенства определяют все более тесно сближающиеся границы, между которыми заключено число $a+b$.

Операция умножения определяется аналогично операции сложения.

Определение 6. *Произведением* ab называется точная верхняя граница множества чисел вида $a'b'$, где a' и b' — те же, что и в предыдущем определении суммы.

Для введенных таким образом операций сложения и умножения справедливы основные законы арифметики:

$a+b=b+a$	*(коммутативность сложения);*
$ab=ba$	*(коммутативность умножения);*
$(a+b)+c=a+(b+c)$	*(ассоциативность сложения);*
$(ab)c=a(bc)$	*(ассоциативность умножения);*
$a(b+c)=ab+ac$	*(дистрибутивность умножения относительно сложения).*

Чтобы доказать любое из этих равенств, следует принять во внимание, что оно справедливо для конечных десятичных дробей, а затем учесть данные выше определения.

Рассмотрим теперь операцию деления. Сначала определим число b^{-1}, где b — положительное действительное число. По определению, b^{-1} есть точная верхняя граница чисел вида $1/b''$, где b'' — любая конечная десятичная дробь, большая чем b. Можно показать, что

$$b \cdot b^{-1}=1,$$

т. е. число b^{-1} является обратным к b в обычном смысле. Далее определим *частное* a/b двух положительных действительных чисел с помощью формулы

$$\frac{a}{b}=ab^{-1}.$$

Учитывая равенство $b \cdot b^{-1}=1$, имеем

$$\frac{a}{b}b=1.$$

Это показывает, что введенная операция деления обратна операции умножения.

§ 9. Действительные числа любого знака

1. Отрицательные действительные числа

Как известно, потребности математики и ее приложений не обеспечиваются одними лишь положительными числами: необходимы также отрицательные числа и нуль.

Определение 1. *Отрицательным действительным числом* будем называть символ $-a$, где a — положительное действительное число.

Разумеется, такое определение еще недостаточно: необходимо распространить арифметические операции (прежде всего, сложение и умножение) на отрицательные действительные числа. Это будет сделано ниже (см. п. 3).

Положительные и отрицательные действительные числа, а также число 0 образуют *множество всех действительных чисел.*

2. Лемма о разности положительных чисел

Лемма. *Если a и b — два положительных действительных числа, причем $a<b$, то существует единственное положительное действительное число x такое, что $a+x=b$.*

□ Рассмотрим положительные действительные числа t такие, что $a+t<b$. Обозначим множество всех таких чисел через M. Оно непусто (см. свойство 4° в п. 2 § 8) и ограничено сверху (например, числом b); следовательно, оно имеет точную верхнюю границу. Обозначим ее через c и покажем, что $a+c=$ $=b$.

Для любого t из M справедливо неравенство $a+t<b$. Это означает, что число b является верхней границей множества чисел $a+t$, где $t\in M$. Так как верхняя граница всегда больше или равна точной верхней границе, то $b\geqslant a+$ $+c$ или, что то же, $a+c\leqslant b$. Остается проверить невозможность неравенства $a+c<b$.

Рассуждая от противного, допустим, что $a+c<b$. Согласно свойству отношения «$<$» (см. § 8, п. 2) найдется такое положительное действительное число d, что $(a+c)+d<b$, или, что то же самое, $a+(c+d)<b$. Это означает, что число $c+d$ принадлежит множеству M. Но $c+d>c$. Мы пришли к противоречию, так как одно из чисел множества M (число $c+d$) оказалось больше, чем верхняя граница M. Значит, $a+c=d$.

Единственность решения уравнения $a+x=b$ устанавливается совсем просто. Предположим, что d — еще одно решение, т. е. $a+c=b$ и $a+d=b$. Если числа c и d различны, то одно из них больше другого. Пусть, например, $c>d$. Тогда $a+c>a+d$, что невозможно, так как обе суммы равны b. ■

Определение 2. Число x, являющееся решением уравнения $a+x=b$, где a и b — два положительных действительных числа таких, что $b>$ $>a$, называется *разностью* чисел b и a и обозначается $b-a$.

В следующем пункте понятие разности будет распространено на общий случай (когда a и b — любые действительные числа).

3. Распространение арифметических операций и неравенств на все действительные числа

Сначала рассмотрим операцию сложения. Так как сумма двух положительных действительных чисел уже определена, то остается определить сумму для остальных случаев, а именно:

$$t+0,\ 0+t,\ (-a)+b,\ a+(-b),\ (-a)+(-b),$$

где t — любое действительное число, а a и b — положительные действительные числа.

Прежде всего положим

$$t+0=t,\ 0+t=0.$$

Далее, положим

$$(-a)+b=\begin{cases} b-a, & \text{если } b>a;\\ -(a-b), & \text{если } a>b;\\ 0, & \text{если } a=b.\end{cases}$$

Аналогично определяется $a+(-b)$. Наконец, по определению полагаем

$$(-a)+(-b)=-(a+b).$$

Для введенной таким образом операции сложения действительных чисел остаются справедливыми коммутативный и ассоциативный законы сложения.

Перейдем теперь к операции вычитания. По определению, полагаем

$$t-s=t+(-s),$$

где t и s — любые действительные числа. При этом если s — отрицательное число, т. е. число вида $-a$, где a положительно, то под $-s$ понимается число a.

Легко проверить выполнение равенства

$$s+(t-s)=t,$$

означающего, что вычитание есть действие, обратное сложению.

Далее, рассмотрим операцию умножения. По определению,

$$t\cdot 0=0,\ 0\cdot t=0,$$

где t — любое действительное число. Далее, если a и b — два положительных действительных числа, то полагаем

$$(-a)b=-(ab),\ a(-b)=-(ab),\ (-a)(-b)=ab.$$

Таким образом, определяя операцию умножения, мы руководствуемся обычными «правилами знаков»:

$$(-)(+)=-,\ (+)(-)=-,\ (-)(-)=+.$$

Законы коммутативности и ассоциативности умножения, а также закон дистрибутивности умножения относительно сложения остаются верными. Кроме того, справедливо равенство

$$t(s-u)=ts-tu$$

(дистрибутивность умножения относительно вычитания).

Рассмотрим, наконец, операцию деления. Частное t/s в случае, когда $t>0$, $s>0$, определено ранее. В остальных случаях определим частное с помощью «правил знаков»:

$$\frac{(+)}{(-)}=-,\ \frac{(-)}{(+)}=-,\ \frac{(-)}{(-)}=+.$$

Например, третье из этих правил означает, что

$$\frac{-a}{-b}=\frac{a}{b},$$

где a и b — любые положительные действительные числа.

Как и в случае положительных чисел, имеем

$$\frac{t}{s}s=t,$$

что характеризует деление как действие, обратное сложению.

Частное $t/0$ не определено (деление на нуль невозможно).

Определение 3. Пусть a и b — два действительных числа. Условимся считать, что $a>b$, если разность $a-b$ есть положительное действительное число. Вместо $a>b$ будем также писать $b<a$.

Основные свойства отношения «$>$»:

1°. *Если $a>b$ и $b>c$, то $a>c$.*

2°. *Если $a>b$, то $a+c>b+c$, где c — любое действительное число.*

4. Модуль (абсолютная величина) действительного числа

Модуль действительного числа t обозначается $|t|$. Согласно определению,

$$|a| = a \text{ и } |-a| = a, \tag{1}$$

где a — любое положительное действительное число или нуль. Например, $|2{,}375| = 2{,}375$, $|-6{,}1| = 6{,}1$.

Перечислим ряд свойств модуля.

1^0. *Модуль произведения равен произведению модулей:*

$$|ts| = |t| \cdot |s| \tag{2}$$

2^0. *Модуль частного равен отношению модулей делимого и делителя:*

$$\left|\frac{t}{s}\right| = \frac{|t|}{|s|}. \tag{3}$$

3^0. *Модуль суммы не превосходит суммы модулей:*

$$|t+s| \leqslant |t| + |s|. \tag{4}$$

4^0. *Модуль суммы не меньше, чем разность модулей слагаемых:*

$$|t+s| \geqslant |t| - |s|. \tag{5}$$

Приведем доказательство свойства 3^0 (доказательства остальных свойств несложны).

□ Возможны четыре случая: 1) $a \geqslant 0$, $b \geqslant 0$; 2) $a \geqslant 0$, $b < 0$; 3) $a < 0$, $b \geqslant 0$; 4) $a < 0$, $b < 0$. Рассмотрим каждый из них в отдельности.

I случай: $a \geqslant 0$, $b \geqslant 0$. Тогда $a + b \geqslant 0$; значит, $|a+b| = a+b$, $|a| = a$, $|b| = b$, и неравенство (4) принимает вид $a+b \leqslant a+b$, т. е. получаем верное утверждение.

II случай: $a \geqslant 0$, $b < 0$. Тогда $|a| = a$, $|b| = -b$, $|a+b| = |a| - |b|$ или $|b| - |a|$ (в зависимости от того, что больше, $|a|$ или $|b|$). Так как $|a| - |b| < |a| + |b|$, $|b| - |a| \leqslant |a| + |b|$, то $|a+b| < |a| + |b|$; значит, неравенство (4) выполняется и в этом случае.

III случай: $a < 0$, $b \geqslant 0$ аналогичен предыдущему.

IV случай: $a < 0$, $b < 0$. Имеем $|a+b| = |-a-b| = |(-a)+(-b)|$. Для положительных чисел $(-a)$ и $(-b)$ неравенство (4) справедливо (см. случай I); значит, $|(-a)+(-b)| \leqslant |-a| + |-b|$. Но $|-a| = |a|$, $|-b| = |b|$. В результате получаем неравенство (4). ■

5. Обзор свойств действительных чисел

Для математического анализа важно не столько то, что представляют собой действительные числа, сколько то, какими свойствами они обладают. Подводя итог предшествующему изложению, выделим свойства действительных чисел, которые в последующих главах будут широко использоваться:

$$a+b = b+a;\ a+(b+c) = (a+b)+c;\ a+0 = a;\ a+(-a) = 0;\ ab = ba;$$

$$a(bc) = (ab)c;\ a \cdot 1 = a;\ a\frac{1}{a} = 1,\ a \neq 0;\ a(b+c) = ab+ac;$$

$$a > b,\ b > c \Rightarrow a > c;\ a > b \Rightarrow a+c > b+c;\ |ab| = |a| \cdot |b|;\ \left|\frac{a}{b}\right| = \frac{|a|}{|b|},\ b \neq 0;$$

$$|a+b| \leqslant |a| + |b|.$$

§ 10. Числовая прямая. Границы числовых множеств. Разделяющие числа

1. Числовая прямая

Напомним известный из школьного курса математики способ изображения чисел с помощью точек на прямой.

Пусть l — прямая, на которой выделена точка O, называемая *началом*. Один из двух лучей, выходящих из O, назовем *положительным*, другой — *отрицательным*. Обычно прямая l изображается горизонтальной, причем в качестве положительного луча принимается луч, идущий из O вправо.

Каждому положительному числу x можно сопоставить точку X, лежащую на положительном луче и такую, что $OX = x$ (рис. 2); условимся говорить, что эта точка изображает число x. Аналогично, каждому отрицательному числу x' сопоставим точку X', лежащую на отрицательном луче и такую, что $OX' = |x'|$; точка X' изображает число x'.

Так как всякий отрезок имеет (при выбранной единице измерения) однозначно определенную длину, то каждая точка прямой изображает некоторое

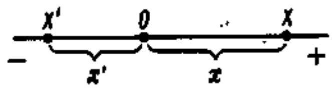


Рис. 2

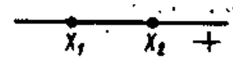


Рис. 3

однозначно определенное действительное число. Обратно, каждое действительное число изображается единственной точкой прямой. Иначе говоря, описанный способ изображения чисел точками прямой устанавливает *взаимно однозначное соответствие между множеством всех действительных чисел и множеством всех точек прямой*. Отметим, что если $x_2 > x_1$, то точка X_2 лежит с положительной стороны от точки X_1 (рис. 3).

Прямая, точки которой сопоставлены указанным выше образом с действительными числами, называется *числовой прямой*. Число, изображаемое точкой, называют *координатой* этой точки (на числовой прямой).

Способ изображения чисел точками на числовой прямой столь привычен для математики, что, как правило, не различают число x и точку X, его изображающую; обычно говорят «точка x».

Нетрудно видеть, что расстояние между точками X_1 и X_2 на числовой прямой равно модулю разности чисел x_1 и x_2, т. е. $X_1X_2 = |x_1 - x_2|$.

2. Ограниченные числовые множества и их границы

Пусть M — некоторое множество действительных чисел (множество точек числовой прямой). Говорят, что множество M *ограничено сверху*, если существует такое число t, что $a \leqslant t$ для всякого $a \in M$. Геометрически это означает, что множество M расположено целиком слева от точки t (при этом не исключается, что сама точка t входит в M). Число t называется в этом случае *верхней границей* для M. Заметим, что для множеств положительных действительных чисел понятие ограниченности сверху было определено в п. 3 § 8.

Аналогично определяют ограниченность снизу. А именно, множество M называется *ограниченным снизу*, если существует такое число s, что $s \leqslant a$ для всякого $a \in M$. Геометрически это означает, что множество M расположено правее точки s. В этом случае число s называется *нижней границей* для M.

Если множество M ограничено и сверху, и снизу, то оно называется просто *ограниченным*. Геометрический смысл ограниченности заключается в следующем: множество M целиком помещается в некотором отрезке $[s, t]$ числовой прямой.

Определение 1. Наименьшая из всех верхних границ множества M называется *точной верхней границей* (или *верхней гранью*) для M. Наибольшая из всех нижних границ для M называется *точной нижней границей* (или *нижней гранью*) для M.

Точная верхняя граница множества M обозначается $\sup M$ (от лат. supremum — «наивысшее»), а точная нижняя граница — $\inf M$ (от infinum — «наинизшее»).

Заметим, что определение точных границ, данное в § 8, относилось лишь к множествам, состоящим из положительных действительных чисел, теперь же M есть любое множество действительных чисел.

Теорема 1 (о точных границах). *Если множество M ограничено сверху, то для него существует точная верхняя граница. Если M ограничено снизу, то для M существует точная нижняя граница.*

□ Ограничимся доказательством первого утверждения теоремы (второе доказательство аналогично).

Итак, пусть M ограничено сверху. Возможны два случая.

1) В M содержится хотя бы одно положительное число. Обозначим через M' множество всех положительных чисел из M. Будучи ограниченным сверху множеством положительных действительных чисел, M' имеет по теореме из п. 3 § 8 точную верхнюю границу c. Ясно, что число c является точной верхней границей и для всего множества M.

2) В M не содержится положительных чисел. Возьмем какое-либо число a из M ($a \leqslant 0$) и «сдвинем» все числа из M на $-a+1$ (т. е. прибавим к каждому числу из M число $-a+1$); получим новое множество M^*. Если мы докажем, что множество M^* имеет точную верхнюю границу c^*, то и множество M будет иметь точную верхнюю границу (равную $c^*-(-a+1)$). Но M^* содержит положительное число 1 (результат сдвига числа a на $-a+1$), и значит, по доказанному, M^* должно иметь точную верхнюю границу. ■

Теорема о точных границах играет фундаментальную роль в математическом анализе.

3. Принцип разделяющего числа

Не меньшее значение, чем теорема о точных границах, имеет и другое предложение, которое обычно называют принципом разделяющего числа.

Определение 2. Пусть A и B — два множества действительных чисел и c — некоторое число. Говорят, что число *c разделяет* множества A и B, если для любых чисел a и b таких, что $a \in A$, $b \in B$, справедливы неравенства $a \leqslant c \leqslant b$.

Очевидно, что не для всяких двух множеств A и B существует разделяющее число. Для существования такого числа необходимо, чтобы каждое число из множества A не превосходило любое число из B. Условимся называть такие множества *разъединенными*, причем множество A будем называть *левым*, а B — *правым*.

Оказывается, условие разъединенности не только необходимо, но и достаточно для существования разделяющего числа. Справедлива следующая теорема.

Теорема 2 (принцип разделяющего числа). *Если A и B — два разъединенных множества действительных чисел, то существует число, разделяющее эти множества.*

Например, если A есть множество всех действительных чисел, меньших 1, а B — множество всех действительных чисел, бо́льших 1, то разделяющим числом является 1. В этом случае разделяющее число не принадлежит ни одному из множеств A и B. Но так будет не всегда: скажем, если A есть множество всех чисел, меньших или равных 1, а B — множество всех чисел, больших или равных 1, то разделяющее число 1 принадлежит сразу обоим множествам A и B.

□ Нетрудно видеть, что левое множество A ограничено сверху; его верхней границей служит любое число из B. Следовательно, существует точная верхняя граница для A. Обозначим эту границу через c и покажем, что c является искомым разделяющим числом.

Пусть b — любое число из множества B. Мы должны показать, что $c \leqslant b$. Допустим противное, т. е. что $c > b$. Между числами b и c обязательно найдется число из множества A: если бы это было не так, т. е. если бы все числа из A лежали слева от b, то получилось бы, что одна из верхних границ для A (а именно, число b) меньше, чем точная верхняя граница c, а это невозможно. Итак, найдется число $a \in A$, которое больше b, что противоречит условию теоремы. Следовательно, точка c разделяет множества A и B. ■

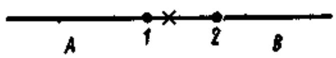


Рис. 4

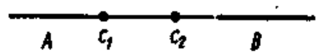


Рис. 5

Следствие. *Пусть левое множество A и правое множество B разъединены. Тогда одним из чисел, разделяющих эти множества, является точная верхняя граница множества A.*

Аналогично можно показать, что разделяющим числом служит и точная нижняя граница множества B.

Возникает вопрос: является ли построенная в доказательстве теоремы точка c **единственной** разделяющей точкой? Примеры показывают, что этого может не быть: так, если множество A состоит из всех чисел, меньших 1, а B — из чисел, бо́льших 2, то любое число из отрезка $[1, 2]$ является разделяющим (рис. 4). Интуитивно ясно, что для единственности разделяющей точки нужно, чтобы множества A и B сближались как угодно тесно. Сформулируем эту мысль более точно.

Теорема 3 (критерий единственности разделяющего числа). *Число c, разделяющее два разъединенных множества A и B (A — левое, B — правое), единственно тогда и только тогда, когда выполнено следующее условие: каково бы ни было положительное число ε, существуют два числа a и b, $a \in A$, $b \in B$, для которых* $b - a < \varepsilon$ (т. е. расстояние между a и b меньше ε).

□ Пусть выполнено условие теоремы; покажем, что разделяющее число единственно. Предположим, что это не так, т. е. что существуют два различных разделяющих числа c_1 и c_2; мы можем считать $c_1 < c_2$ (рис. 5). Так как любое число a из A расположено слева от c_1 (или совпадает с c_1), а любое число b из B расположено справа от c_2 (или совпадает с c_2), то расстояние между a и b не меньше, чем $c_2 - c_1$. Это противоречит заданному условию. Следовательно, разделяющее число единственно.

Обратно, пусть разделяющее число единственно; докажем, что выполняется условие теоремы. Пусть ε — любое положительное число. В силу предположения о единственности разделяющего числа точная верхняя граница множества A совпадает с точной нижней границей множества B. Обозначим эту общую границу через c и рассмотрим точки $c-\frac{\varepsilon}{2}$ и $c+\frac{\varepsilon}{2}$. Поскольку $c-\frac{\varepsilon}{2}<c$ и c — точная верхняя граница для A, найдется точка $a \in A$ такая, что $c-\frac{\varepsilon}{2}<a$. Аналогично найдется точка $b \in B$ такая, что $c+\frac{\varepsilon}{2}>b$ (рис. 6). Из написанных неравенств следует, что расстояние между точками a и b меньше расстояния между точками $c-\frac{\varepsilon}{2}$ и $c+\frac{\varepsilon}{2}$, т. е. меньше ε, а это и нужно было показать. ■

Принцип разделяющего числа делает наглядным одно из самых важных свойств множества действительных чисел (или, что эквивалентно, точек числовой прямой). Это свойство называется «непрерывностью» и понимается как отсутствие «пустот», «пробелов». Если представить, что для каких-то двух разъединенных множеств A и B не существовало бы разделяющего числа, то это означало бы наличие определенного «пробела» между множествами A и B на числовой прямой. Отсутствие таких пробелов и характеризуется словом «непрерывность».

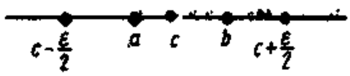


Рис. 6

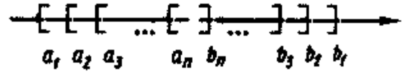


Рис. 7

Разумеется, если из множества всех действительных чисел изъять какое-либо подмножество (т. е. создать «пробелы»), то для оставшегося множества принцип разделяющего числа выполняться не будет. Например, если изъять одно-единственное число 1, то два множества $A=(-\infty, 1)$, $B=(1, \infty)$, принадлежащие оставшейся части, не будут иметь разделяющего числа.

4. Теорема о стягивающейся последовательности отрезков

Не меньшее значение, чем теорема о точных границах, имеет и другое предложение, которое обычно называют теоремой о стягивающейся последовательности отрезков.

Определение 3. Последовательность отрезков $[a_1, b_1]$, $[a_2, b_2]$, ... на числовой прямой называется *стягивающейся* (рис. 7), если:

1) каждый следующий отрезок содержится в предыдущем, т. е. $a_{n+1} \geqslant a_n$, $b_{n+1} \leqslant b_n$ $(n=1, 2, \ldots)$;
2) длины отрезков стремятся к нулю.

Второе условие означает, что каково бы ни было положительное число ε, найдется такой номер n, что длина отрезка $[a_n, b_n]$ (а значит, и длины всех последующих отрезков) будет меньше ε. Иначе говоря, разность b_n-a_n с ростом n становится как угодно малым числом.

Теорема 4 (о стягивающейся последовательности отрезков). *Какова бы ни была стягивающаяся последовательность отрезков, существует единственная точка, принадлежащая всем этим отрезкам.*

Иными словами, существует, и притом единственное, число c такое, что $a_n \leqslant c \leqslant b_n$ для всех $n=1, 2, \ldots$.

□ Множества $A=\{a_1, a_2, \ldots\}$ и $B=\{b_1, b_2, \ldots\}$ являются разъединенными, т. е. $a_n \leqslant b_m$ для любых натуральных n и m. В самом деле, если $n=m$, то $a_n<b_n$ по условию; если $n>m$, то $a_n<b_n \leqslant b_m$; если $n<m$, то $a_n \leqslant a_m<b_m$. Согласно теореме 2 существует число, разделяющее A и B; в силу теоремы 3 такое число единственно. ■

3

Числовые последовательности и их пределы

§ 11. Метод математической индукции

1. Полная и неполная индукция

В основе всякого математического рассуждения лежат дедуктивный или индуктивный методы. *Дедукция* — это умозаключение от общего к частному, например, такое: известно, что если сумма цифр натурального числа делится на 3, то и само число делится на 3; так, поскольку у числа 678 сумма цифр, равная 21, делится на 3, то и само число 678 делится на 3. *Индукция* — это умозаключение от частного к общему, когда общие выводы получают, опираясь на ряд частных утверждений.

Различают полную и неполную индукцию. *Полная индукция* состоит в том, что общее утверждение доказывается по отдельности в каждом из конечного числа возможных случаев.

Неполная индукция заключается в том, что общий вывод делается на основе не всех, а достаточно большого числа частных случаев. Результат, полученный неполной индукцией, остается, однако, лишь гипотезой, пока он не доказан строгим математическим рассуждением, охватывающим все частные случаи.

Полная индукция имеет в математике лишь ограниченное применение. Многие интересные математические утверждения охватывают бесконечное число частных случаев, а провести проверку для бесконечного числа частных случаев человек не может (примером такого утверждения может служить любое утверждение, относящееся ко всем натуральным числам). Неполная индукция часто приводит к ошибочным результатам.

Во многих случаях выход из такого рода затруднений заключается в обращении к особому методу рассуждений, называемому методом математической индукции.

2. Метод математической индукции

Пусть нужно доказать справедливость некоторого утверждения для любого натурального числа n (например, нужно доказать, что сумма первых n нечетных чисел равна n^2). Непосредственная проверка этого утверждения для каждого значения n невозможна, поскольку множество натуральных чисел бесконечно. Чтобы доказать это утверждение, сначала проверяют его справедливость при $n = 1$. Затем доказывают, что для любого натурального значения k

из справедливости рассматриваемого утверждения при $n=k$ вытекает его справедливость и при $n=k+1$. Тогда утверждение считается доказанным при всех n. В самом деле, утверждение справедливо при $n=1$. Но тогда оно справедливо и для следующего числа $n=1+1=2$. Из справедливости утверждения для $n=2$ вытекает его справедливость для $n=2+1=3$. Отсюда, в свою очередь, следует справедливость утверждения для $n=4$ и т. д. Ясно, что в конце концов мы дойдем до любого натурального числа n. Значит, утверждение верно для любого n.

Обобщая сказанное, сформулируем следующий общий принцип.

Принцип математической индукции. *Если предложение $A(n)$, зависящее от натурального числа n, истинно для $n=1$ и из того, что оно истинно для $n=k$ (где k — любое натуральное число), следует, что оно истинно и для следующего числа $n=k+1$* [кратко это записывают так: $A(k)\Rightarrow A(k+1)$], *то предложение $A(n)$ истинно для любого натурального числа n.*

На этом принципе основан метод математической индукции. Доказательство с помощью этого метода проводится следующим образом.

1°. Доказываемое утверждение проверяют для $n=1$. Эта часть доказательства называется *базисом индукции.*

2°. Доказывают справедливость утверждения для $n=k+1$ в предположении справедливости утверждения для $n=k$, т. е. доказывают, что $A(k)\Rightarrow A(k+1)$. Эта часть доказательства называется *индукционным шагом.*

Примеры. 1. Доказать, что

$$1+3+5+\ldots+(2n-1)=n^2, \tag{1}$$

или короче,

$$S_n=n^2, \text{ где } S_n=1+3+5+\ldots+(2n-1).$$

○ 1°. Имеем $S_1=1=1^2$. Следовательно, утверждение верно при $n=1$, т. е. A (1) истинно.

2°. Докажем, что $A(k)\Rightarrow A(k+1)$. Пусть k — любое натуральное число и пусть формула (1) верна при $n=k$, т. е. $S_k=k^2$.

Докажем, что тогда утверждение справедливо и для следующего натурального числа $n=k+1$, т. е. что $S_{k+1}=(k+1)^2$. Действительно,

$$S_{k+1}=(1+3+5+\ldots+(2k-1))+(2k+1)=S_k+(2k+1)=k^2+(2k+1)=(k+1)^2.$$

Итак, $A(k)\Rightarrow A(k+1)$. На основании принципа математической индукции заключаем, что предложение $A(n)$ истинно для любого $n\in\mathbf{N}$. ●

2. Доказать, что

$$1^2+2^2+3^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6}. \tag{2}$$

○ 1°. При $n=1$ формула верна: $1^2=\dfrac{1(1+1)(2+1)}{6}$.

2°. Предположим, что $1^2+2^2+\ldots+k^2=\dfrac{k(k+1)(2k+1)}{6}$. Докажем, что формула (2) верна и при $n=k+1$, т. е. что

$$1^2+2^2+3^2+\ldots+k^2+(k+1)^2=\frac{(k+1)(k+2)(2k+3)}{6}.$$

В самом деле, имеем

$$(1^2+2^2+\ldots+k^2)+(k+1)^2=\frac{k(k+1)(2k+1)}{6}+(k+1)^2=$$

$$=\frac{(k+1)(k(2k+1)+6(k+1))}{6}=\frac{(k+1)(2k^2+7k+6)}{6}=\frac{(k+1)(k+2)(2k+3)}{6}.$$

Значит, формула (2) верна для любого n. ●

В ряде случаев бывает нужно доказать справедливость некоторого утверждения не для всех натуральных чисел, а лишь для $n \geqslant p$, где p — фиксированное натуральное число. Тогда принцип математической индукции формулируется следующим образом: *если предложение $A(n)$ истинно при $n=p$ и если $A(k) \Rightarrow A(k+1)$ для любого $k \geqslant p$, то предложение $A(n)$ истинно для любого $n \geqslant p$.*

Иногда используется еще одна форма принципа математической индукции: *если предложение $A(n)$ истинно для $n=1$ и из того, что оно истинно для всех $n<k$, следует, что оно истинно и для $n=k$, то предложение $A(n)$ истинно для любого n.*

Пример 3. Доказать, что если $n \geqslant 2$ и $x>0$, то справедливо неравенство

$$(1+x)^n > 1+nx \tag{3}$$

(неравенство Бернулли).*

○ 1^0. При $n=2$ неравенство справедливо, так как $(1+x)^2=1+2x+x^2>1+2x$. Значит, $A(2)$ истинно.

2^0. Докажем, что $A(k) \Rightarrow A(k+1)$, если $k \geqslant 2$. Предположим, что $A(k)$ истинно, т. е. что справедливо неравенство

$$(1+x)^k > 1+kx. \tag{4}$$

Докажем, что тогда и $A(k+1)$ истинно, т. е. что справедливо неравенство

$$(1+x)^{k+1} > 1+(k+1)x.$$

В самом деле, умножив обе части неравенства (4) на положительное число $1+x$, получим

$$(1+x)^{k+1} > (1+kx)(1+x).$$

Рассмотрим правую часть последнего неравенства; имеем

$$(1+kx)(1+x)=1+(k+1)x+kx^2>1+(k+1)x.$$

В итоге получаем, что $(1+x)^{k+1}>1+(k+1)x$. Итак, $A(k) \Rightarrow A(k+1)$.

На основании принципа математической индукции можно утверждать, что неравенство Бернулли справедливо для любого $n \geqslant 2$. ●

§ 12. Основные понятия, связанные с последовательностями. Прогрессии

1. Способы задания последовательности

Ранее мы уже использовали термин «последовательность», не дав ему, однако, строгого определения. Мы говорили, например, о последовательности десятичных приближений для данного действительного числа. Так, для числа $\sqrt{2}=$ $=1,4142$ ... можно построить последовательность десятичных приближений по недостатку: $a_1=1$; $a_2=1,4$; $a_3=1,41$; $a_4=1,414$; $a_5=1,4142$ .... Каждому натуральному числу, выражающему номер шага в построении последовательности, соответствует вполне определенное число (член последовательности, десятичное приближение), полученное на рассматриваемом шаге: $1 \to 1$; $2 \to 1,4$; $3 \to 1,41$; $5 \to 1,4142$, ...

Определение 1. Если каждому натуральному числу n поставлено в соответствие некоторое действительное число, то говорят, что задана *числовая последовательность*; это число называют *n-м членом* последовательности и обозначают $f(n)$.

---

* Я. Бернулли (1654—1705) — швейцарский математик.

Часто вместо $f(n)$ пишут y_n, а последовательность обозначают $y_1, y_2, \ldots, y_n, \ldots$ или (y_n) (вместо y может быть использована любая другая буква латинского алфавита).

Для последовательностей особенно важны два способа задания.

1. *Аналитический*, т. е. с помощью формулы n-го члена: $a_n = f(n)$. Пусть, например, $a_n = 1/n$, $b_n = n^2$. Придавая аргументу n значения 1, 2, 3, ..., будем получать соответствующие значения членов последовательностей. Так, последовательность (a_n) имеет вид $1, 1/2, 1/3, 1/4, \ldots, 1/n, \ldots$, а последовательность (b_n) — вид $1, 4, 9, 16, 25, \ldots, n^2, \ldots$.

Иногда n-й член последовательности задается несколькими выражениями, например

$$y_n = \begin{cases} n, & \text{если } n \text{ — простое число;} \\ n^2, & \text{если } n \text{ — составное число.} \end{cases}$$

Тем самым задана последовательность 1, 2, 3, 16, 5, 36, 7, 64, 81, 100, 11, ...

2. Любой член последовательности, начиная с некоторого, часто выражают через предшествующие (один или несколько). Например, последовательность 1, 2, 3, 5, 8, 13, ... может быть задана следующим образом:

$$a_1 = 1,\ a_2 = 2,\ a_{n+2} = a_n + a_{n+1}.$$

Действительно, $a_3 = a_1 + a_2 = 1 + 2 = 3$, $a_4 = a_2 + a_3 = 2 + 3 = 5$, ...

Такой способ задания последовательности называется *рекуррентным**. При рекуррентном способе задания последовательности обычно указывают: а) первый член последовательности или несколько первых членов; б) формулу, позволяющую определить любой член последовательности по известным предшествующим членам.

Примеры. 1. Выписать несколько первых членов последовательности, если известно, что $a_1 = 1$, $a_n = a_{n-1} + 2$ при $n > 1$.

○ Имеем $a_1 = 1$, $a_2 = a_1 + 2 = 1 + 2 = 3$, $a_3 = a_2 + 2 = 3 + 2 = 5, \ldots$. Итак, получаем последовательность 1, 3, 5, 7, 9, ... ●

2. Выписать несколько членов последовательности, если известно, что $a_1 = 1$, $a_n = a_{n-1}/2$ при $n > 1$.

○ Имеем $a_1 = 1$, $a_2 = a_1/2 = 1/2$, $a_3 = a_2/2 = 1/4$, $a_4 = a_3/2 = 1/8 \ldots$. Итак, получаем последовательность 1, 1/2, 1/4, 1/8, ... ●

Определение 2. Числовая последовательность, каждый член которой начиная со второго равен предыдущему члену, сложенному с одним и тем же числом d, называется *арифметической прогрессией*. Число d называется *разностью* арифметической прогрессии.

Таким образом, арифметическая прогрессия определяется условиями: 1) $a_1 = a$; 2) $a_{n+1} = a_n + d$ для любого $n \geqslant 1$.

Например, если $a_1 = 2$, $a_{n+1} = a_n + 3$, то $a_2 = 2 + 3 = 5$, $a_3 = 5 + 3 = 8, \ldots$, т. е. получаем последовательность 2, 5, 8, 11, 14, 17 ... .

Приведем примеры арифметических прогрессий:

2, 5, 8, 11, 14, 17, ... ; здесь $a_1 = 2$, $a_{n+1} = a_n + 3$, $d = 3$;

17, 14, 11, 8, 5, 2, −1, −4, ... ; здесь $a_1 = 17$, $a_{n+1} = a_n - 3$, $d = -3$;

8, 8, 8, 8, 8, ... ; здесь $a_1 = 8$, $a_{n+1} = a_n + 0$, $d = 0$.

Определение 3. Числовая последовательность, первый член которой отличен от нуля, а каждый член начиная со второго равен предшествующему члену, умноженному на одно и то же не равное нулю число q, называется *геометрической прогрессией*. Число q называется *знаменателем* геометрической прогрессии.

* От лат. recurrere («возвращаться»).

Таким образом, геометрическая прогрессия определяется условиями: 1) $b_1 = b$ $(b \neq 0)$; 2) $b_{n+1} = b_n q (q \neq 0)$.

Например, если $b_1 = 1$, $b_{n+1} = b_n \cdot 2$, то $b_2 = 1 \cdot 2 = 2$, $b_3 = b_2 \cdot 2 = 4$, ..., т. е. получаем последовательность 1, 2, 4, 8, 16, ...

В определении полагают $b_1 \neq 0$, $q \neq 0$, чтобы избежать тривиального случая последовательности, состоящей из одних нулей.

Приведем примеры геометрических прогрессий:

2, 8, 32, 128, 512, ... ; здесь $b_1 = 2$, $b_{n+1} = b_n \cdot 4$, $q = 4$;
2, —8, 32, —128, 512, ... ; здесь $b_1 = 2$, $b_{n+1} = b_n \cdot (-4)$, $q = -4$;
27, 9, 3, 1, 1/3, 1/9, ... ; здесь $b_1 = 27$, $b_{n+1} = b_n \cdot (1/3)$, $q = 1/3$;
8, 8, 8, 8, 8, ... ; здесь $b_1 = 8$, $b_{n+1} = b_n \cdot 1$, $q = 1$.

Последовательности задаются не только аналитически или рекуррентно, но и «словесно». Например, для последовательности простых чисел 2, 3, 5, 7, 11, 13, ... или для последовательности 1; 1,4; 1,41; 1,414; 1,4142; ... десятичных приближений по недостатку для $\sqrt{2}$ нет ни аналитической формулы, ни рекуррентного соотношения.

2. Свойства последовательностей

Определение 4. Последовательность (y_n) называется *ограниченной сверху*, если существует число M такое, что $y_n \leqslant M$ для любого $n \in \mathbb{N}$; число M называют *верхней границей* последовательности. Последовательность (y_n) называется *ограниченной снизу*, если существует число m такое, что $y_n \geqslant m$ для любого $n \in \mathbb{N}$; число m называют *нижней границей* последовательности. Наконец, последовательность (y_n) называется *ограниченной*, если она ограничена и снизу и сверху.

Если последовательность ограничена сверху, то она имеет бесконечное множество верхних границ. В самом деле, если M — верхняя граница последовательности (y_n), т. е. $y_n \leqslant M$ для всех n, то любое число $P > M$ также является верхней границей для (y_n). Аналогично, ограниченная снизу последовательность имеет бесконечное множество нижних границ.

Примеры. 1. Последовательность 3, 2, 1, 0, —1, —2, —3, ... ограничена сверху (например, числом 3).

2. Последовательность 1, 2, 3, 4, 5, ..., n, ... ограничена снизу (например, числом 1).

3. Последовательность 1, 1/2, 1/3, 1/4, ..., $1/n$, ... ограничена и сверху (например, числом 1), и снизу (например, числом 0), т. е. является ограниченной.

4. Последовательность c, c, c, ..., c, ... (стационарная последовательность) является ограниченной.

Члены последовательности (y_n) можно изобразить точками на координатной прямой. Если последовательность (y_n) ограничена, причем m — нижняя, а M — верхняя граница, то каждый член последовательности (y_n) принадлежит отрезку $[m, M]$; в этом состоит геометрический смысл ограниченности последовательности.

Теорема 1. *Последовательность (y_n) является ограниченной тогда и только тогда, когда существует число $r > 0$ такое, что $|y_n| < r$ для всех n.*

□ Так как (y_n) — ограниченная последовательность, то все ее члены принадлежат некоторому отрезку $[m, M]$, который всегда можно поместить

Рис. 8

(рис. 8) внутри некоторого интервала $(-r, r)$. Тогда любой член последовательности (y_n) удовлетворяет неравенству $-r < y_n < r$ или, что то же самое, $|y_n| < r$.

Обратно, пусть последовательность (y_n) такова, что для некоторого $r > 0$ и для всех n выполнено неравенство $|y_n| < r$ или, что то же самое, $-r < y_n < < r$. Значит, эта последовательность (y_n) является ограниченной. ■

Теорема 2. *Свойство ограниченности последовательности (сверху, снизу, с двух сторон) не нарушится, если отбросить конечное число членов последовательности или, напротив, к данной последовательности добавить некоторое конечное число членов.*

□ Если последовательность ограничена, т. е. целиком лежит внутри некоторого интервала $(-r, r)$, то последовательность, полученная из данной отбрасыванием конечного числа членов, лежит внутри того же интервала, т. е. также ограничена.

Если к ограниченной последовательности, лежащей внутри интервала $(-r, r)$, добавить конечное число членов, то при необходимости всегда можно увеличить заданный интервал $(-r, r)$ до интервала $(-R, R)$ так, чтобы все приписанные члены оказались внутри большего интервала. Тогда вся новая последовательность будет лежать внутри интервала $(-R, R)$, т. е. она также ограничена. ■

Над последовательностями можно осуществлять арифметические операции. Например, суммой последовательностей (x_n) и (y_n) называется последовательность (z_n) такая, что $z_n = x_n + y_n$. Аналогично определяют разность, произведение и частное последовательностей.

Теорема 3. *Если (x_n) и (y_n) — ограниченные последовательности, то их сумма также является ограниченной последовательностью.*

□ Так как последовательность (x_n) ограничена, то по теореме 1 существует число $r_1 > 0$ такое, что $|x_n| < r_1$ при всех n. Аналогично, в силу ограниченности последовательности (y_n) существует число $r_2 > 0$ такое, что $|y_n| < r_2$ при всех n.

Пусть $z_n = x_n + y_n$. Тогда $|z_n| = |x_n + y_n| \leqslant |x_n| + |y_n| < r_1 + r_2$ (мы воспользовались неравенством (3) из § 9).

Полагая $r = r_1 + r_2$, для последовательности (z_n) имеем $|z_n| < r$ при всех n. Это и означает ограниченность последовательности (z_n). ■

Определение 5. Последовательность (a_n) называется *возрастающей (неубывающей)*, если $a_n < a_{n+1}$ $(a_n \leqslant a_{n+1})$ для любого n, и *убывающей (невозрастающей)*, если $a_n > a_{n+1}$ $(a_n \geqslant a_{n+1})$ для любого n.

Примеры. 5. Последовательность 1/2, 2/3, 3/4, 4/5, ..., $n/n + 1$, ... является возрастающей.

6. Последовательность 1, 1/2, 1/3, 1/4, ..., $1/n$, ... является убывающей.

Возрастающие и убывающие последовательности называются *монотонными* последовательностями.

§ 13. Предел числовой последовательности

1. Бесконечно малые последовательности

Рассмотрим три последовательности:

$$x_n = \frac{1}{n},\ y_n = -\frac{2}{n^2},\ z_n = \frac{(-1)^n}{2n+1}.$$

Они обладают одним общим свойством. Чтобы выявить его, поступим следующим образом.

Возьмем число $\varepsilon_1 = 0{,}01$. Замечаем, что $|x_n| < 0{,}01$ при $n \geqslant 101$; $|y_n| < 0{,}01$ при $n \geqslant 15$ (в самом деле, $|y_n| < 0{,}01 \Leftrightarrow \left|-\frac{2}{n^2}\right| < 0{,}01 \Leftrightarrow \frac{2}{n^2} < 0{,}01 \Leftrightarrow n^2 > 200$; это верно при $n \geqslant 15$); $|z_n| < 0{,}01$ при $n \geqslant 50$ (действительно, $|z_n| < 0{,}01 \Leftrightarrow \left|\frac{(-1)^n}{2n+1}\right| < 0{,}01 \Leftrightarrow \frac{1}{2n+1} < 0{,}01 \Leftrightarrow 2n+1 > 100$; это верно при $n \geqslant 50$).

Возьмем число $\varepsilon_2 = 0{,}001$. Рассуждая, как и выше, замечаем, что $|x_n| < 0{,}001$ при $n \geqslant 1001$; $|y_n| < 0{,}001$ при $n \geqslant 45$; $|z_n| < 0{,}001$ при $n \geqslant 500$.

Вообще, какое бы число $\varepsilon > 0$ ни взять, окажется, что начиная с некоторого номера (для каждой последовательности своего) все члены последовательности по модулю меньше ε. Последовательности, обладающие указанным свойством, называют бесконечно малыми. Бесконечно малые последовательности мы будем обозначать греческими буквами: (α_n), (β_n), (γ_n).

Определение 1. Последовательность (α_n) называется *бесконечно малой,* если для любого $\varepsilon > 0$ существует номер N, начиная с которого каждый член последовательности (α_n) по модулю меньше ε. Короче:

$$(\forall \varepsilon > 0)(\exists N)(\forall n \geqslant N)|\alpha_n| < \varepsilon.$$

Пример 1. Доказать, что последовательность (α_n), где $\alpha_n = \frac{c}{an+b}$, является бесконечно малой (a, b, c — действительные числа, причем $c \neq 0$, $a > 0$).

○ Возьмем произвольное $\varepsilon > 0$ и рассмотрим неравенство

$$\left|\frac{c}{an+b}\right| < \varepsilon. \tag{1}$$

Далее, имеем

$$\frac{|an+b|}{|c|} > \frac{1}{\varepsilon},\ |an+b| > \frac{|c|}{\varepsilon}.$$

Так как $a > 0$, то начиная с некоторого номера N_1 выполняется неравенство $an + b > 0$ и потому $|an+b| = an+b$. Значит, $an + b > \frac{|c|}{\varepsilon}$, откуда

$$n > \frac{1}{a}\left(\frac{|c|}{\varepsilon} - b\right). \tag{2}$$

Если в качестве номера N взять наименьшее натуральное число, большее чем N_1 и $\frac{1}{a}\left(\frac{|c|}{\varepsilon} - b\right)$, то для $n \geqslant N$ тем более будет выполнено неравенство (2), а следовательно, и неравенство (1). Это и означает, что (α_n) — бесконечно малая последовательность. ●

Члены последовательности (α_n) можно изобразить точками на координатной прямой. Если некоторый член α_n последовательности удовлетворяет неравенству $|\alpha_n| < \varepsilon$ или, что то же самое, неравенству $-\varepsilon < \alpha_n < \varepsilon$, то он лежит внутри интервала $(-\varepsilon, \varepsilon)$; если неравенство $|\alpha_n| < \varepsilon$ выполняется для всех $n \geqslant N$, то член последовательности с номером N и все следующие за ним лежат внутри интервала $(-\varepsilon, \varepsilon)$. Таким образом, можно дать следующее геометрическое истолкование бесконечно малой последовательности: *какой бы интервал $(-\varepsilon, \varepsilon)$ мы ни взяли, вся последовательность начиная с номера N лежит внутри этого интервала.*

Рассмотрим свойства бесконечно малых последовательностей.

1^0. *Стационарная последовательность $c, c, c, \ldots, c, \ldots$ является бесконечно малой тогда и только тогда, когда $c = 0$.*

2^0. *Свойство последовательности быть бесконечно малой не нарушится, если отбросить конечное число членов последовательности или, напротив, приписать к данной последовательности конечное число членов.*

3^0. *Если (β_n) — бесконечно малая последовательность и для всех n выполняется неравенство $|\alpha_n| \leqslant |\beta_n|$, то и последовательность (α_n) является бесконечно малой.*

Эти свойства достаточно очевидны, их доказательства легко провести самостоятельно.

4^0. *Бесконечно малая последовательность является ограниченной.*

□ Возьмем произвольное $\varepsilon > 0$. Так как последовательность (α_n) — бесконечно малая, то $|\alpha_n| < \varepsilon$ для всех n начиная с некоторого номера N. Это означает (см. теорему 1 § 12), что если последовательность (α_n) рассматривать только начиная с номера N, то она является ограниченной. Тогда по теореме 2 § 12 и вся последовательность (α_n) ограничена. ■

5^0. *Сумма двух бесконечно малых последовательностей есть также бесконечно малая последовательность.*

□ Пусть (α_n), (β_n) — бесконечно малые последовательности. Возьмем произвольное $\varepsilon > 0$. Тогда существует номер N_1, начиная с которого выполняется неравенство $|\alpha_n| < \varepsilon/2$, и существует номер N_2, начиная с которого выполняется неравенство $|\beta_n| < \varepsilon/2$.

Из чисел N_1, N_2 возьмем наибольшее и обозначим его через N; тогда для всех $n \geqslant N$ выполняются неравенства $|\alpha_n| < \varepsilon/2$ и $|\beta_n| < \varepsilon/2$.

Рассмотрим последовательность (γ_n) такую, что $\gamma_n = \alpha_n + \beta_n$. Имеем $|\gamma_n| = |\alpha_n + \beta_n| \leqslant |\alpha_n| + |\beta_n|$ (мы воспользовались неравенством (3) из § 9). Тогда для $n \geqslant N$ получим

$$|\gamma_n| \leqslant |\alpha_n| + |\beta_n| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Итак, для $n \geqslant N$ выполняется неравенство $|\gamma_n| < \varepsilon$, а это и означает, что (γ_n) — бесконечно малая последовательность. ■

6^0. *Если (α_n) — бесконечно малая последовательность, а (y_n) — ограниченная последовательность, то их произведение является бесконечно малой последовательностью.*

□ Так как (y_n) — ограниченная последовательность, то существует $r > 0$ такое, что $|y_n| < r$ для всех n (см. теорему 1 § 12).

Возьмем произвольное $\varepsilon > 0$. Поскольку (α_n) — бесконечно малая последовательность, существует номер N, начиная с которого выполняется неравенство $|\alpha_n| < \varepsilon/r$.

Рассмотрим последовательность (x_n) такую, что $x_n = y_n \cdot \alpha_n$. Имеем

$$|x_n| = |y_n \alpha_n| = |y_n| \cdot |\alpha_n| < r \cdot \varepsilon/r = \varepsilon \text{ для } n \geqslant N.$$

Итак, $|x_n| < \varepsilon$ для всех $n \geqslant N$. Это и означает, что (x_n) — бесконечно малая последовательность. ■

Следствие 1. *Если (α_n) — бесконечно малая последовательность, то и последовательность $(c\alpha_n)$, где c — любое действительное число, также является бесконечно малой.*

Следствие 2. *Произведение двух и вообще любого конечного числа бесконечно малых последовательностей есть бесконечно малая последовательность.*

Следствие 3. *Если* (α_1), (α_2), (α_3), ..., (α_n) — *бесконечно малые последовательности, то и последовательность* (β_n) *такая, что* $\beta_n = c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 + \dots + c_n\alpha_n$, *где* $c_1, c_2, c_3, \dots, c_n$ — *действительные числа, также является бесконечно малой.*

Так, выше мы отмечали, что $\left(\frac{1}{n}\right)$ — бесконечно малая последовательность. Тогда бесконечно малыми являются и такие последовательности: $\left(\frac{5}{n}\right)$ (согласно следствию 1); $\left(\frac{1}{n^4}\right)$ (согласно следствию 2, так как $\frac{1}{n^4} = \frac{1}{n}\cdot\frac{1}{n}\cdot\frac{1}{n}\cdot\frac{1}{n}$); $\left(\frac{1}{n\sqrt{n}}\right)$ (в силу свойства 3^0, поскольку $\frac{1}{n\sqrt{n}} \leqslant \frac{1}{n}$ для всех n); $\left(\frac{1}{n} + \frac{2}{n^2} - \frac{3}{n^3} + \frac{\sqrt{2}}{n^4}\right)$ (согласно следствию 3).

Замечание. Если сумма, разность и произведение бесконечно малых последовательностей также являются бесконечно малыми, то о частном этого сказать нельзя. Так, $\alpha_n = \frac{3}{n^2}$, $\beta_n = \frac{1}{n^2}$ — бесконечно малые последовательности, но $x_n = \frac{\alpha_n}{\beta_n} = 3$, т. е. (x_n) — стационарная последовательность 3, 3, 3, ..., не являющаяся бесконечно малой.

2. Предел числовой последовательности и его свойства

Определение 2. Число b называют *пределом последовательности* (x_n), если $(x_n - b)$ — бесконечно малая последовательность, и пишут $\lim\limits_{n\to\infty} x_n = b$. Говорят также, что последовательность (x_n) *сходится*. Если последовательность не имеет предела, то ее называют *расходящейся*.

Из этого определения, в частности, следует, что

$$\lim_{n\to\infty} c = c, \quad \lim_{n\to\infty} \alpha_n = 0,$$

где (α_n) — бесконечно малая последовательность.

Пример 1. Доказать, что $\lim\limits_{n\to\infty} \frac{2n+1}{3n+2} = \frac{2}{3}$.

○ Рассмотрим последовательность (α_n) такую, что $\alpha_n = \frac{2n+1}{3n+2} - \frac{2}{3}$. Имеем $\alpha_n = \frac{6n+3-6n-4}{3(3n+2)} = -\frac{1}{9n+6}$. Поскольку $(\alpha_n) = \left(\frac{-1}{9n+6}\right)$ — бесконечно малая последовательность (см. пример 1 п. 1), это и означает, что $\lim\limits_{n\to\infty} \frac{2n+1}{3n+2} = \frac{2}{3}$. ●

Воспользовавшись для бесконечно малой последовательности $(x_n - b)$ определением 1, получим еще одно определение предела последовательности, эквивалентное предыдущему.

Определение 3. Число b называется *пределом последовательности* (x_n), если для любого $\varepsilon > 0$ существует номер N, начиная с которого выполняется неравенство $|x_n - b| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)(\exists N)(\forall n \geqslant N)|x_n - b| < \varepsilon.$$

Определению 3 можно дать геометрическое истолкование; для этого введем понятие ε-окрестности точки.

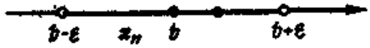


Рис. 9

Интервал $(b - \varepsilon,\ b + \varepsilon)$ называют *ε-окрестностью точки* b, а число ε — *радиусом* этой окрестности $(\varepsilon > 0)$.

Изобразим члены последовательности (x_n) точками на координатной прямой. Неравенство $|x_n - b| < \varepsilon$, или, что то же самое, $b - \varepsilon < x_n < b + \varepsilon$, означает, что x_n принадлежит интервалу $(b - \varepsilon,\ b + \varepsilon)$, т. е. ε-окрестности точки b (рис. 9). Поэтому определение 3 можно сформулировать следующим образом.

Определение 4. Число b называется *пределом последовательности* (x_n), если какую бы ε-окрестность точки b мы ни взяли, все члены последовательности начиная с некоторого номера N принадлежат этой окрестности.

На практике в различных ситуациях оказывается удобным то или другое из определений 2—4. Чтобы различать их, условимся определение 2 называть определением предела «на языке бесконечно малых», определение 3 — «на языке $\varepsilon - N$», определение 4 — «на языке окрестностей».

Теорема 1. *Если последовательность сходится, то только к одному пределу.*

□ Пусть последовательность (x_n) сходится. Предположим, что ее предел не является единственным, т. е. что одновременно верны равенства $\lim\limits_{n \to \infty} x_n = b$ и $\lim\limits_{n \to \infty} x_n = c$, где $b \neq c$.

Воспользуемся определением «на языке бесконечно малых». Равенство $\lim\limits_{n \to \infty} x_n = b$ означает, что $x_n = b + \alpha_n$, а равенство $\lim\limits_{n \to \infty} x_n = c$ — что $x_n = c + \beta_n$, где (α_n) и (β_n) — бесконечно малые последовательности. Тогда $b + \alpha_n = c + \beta_n$, откуда $\alpha_n - \beta_n = c - b$. Но последовательность $(\alpha_n - \beta_n)$ — бесконечно малая (см. следствие 3 в п. 1); значит, стационарная последовательность $(c - b)$ является бесконечно малой, а это возможно лишь в случае, когда $c = b$ (см. свойство 1^0 из п. 1), что противоречит условию. ■

Теорема 2. *Если последовательность сходится, то она ограничена.*

□ Пусть $\lim\limits_{n \to \infty} x_n = b$. Воспользуемся определением «на языке бесконечно малых». Имеем $x_n = b + \alpha_n$, где (α_n) — бесконечно малая последовательность. Стационарная последовательность (b) и бесконечно малая последовательность (α_n) являются ограниченными (см. свойство 4^0 из п. 1), тогда и их сумма (согласно свойству 3^0 из п. 1) также ограничена. ■

Теорема 3 (о предельном переходе в неравенствах). *Если последовательности* (x_n) *и* (y_n) *сходятся и* $x_n \leqslant y_n$ *для всех* n, *то* $\lim\limits_{n \to \infty} x_n \leqslant \lim\limits_{n \to \infty} y_n$.

□ Пусть $\lim\limits_{n \to \infty} x_n = b$, $\lim\limits_{n \to \infty} y_n = c$. Воспользуемся определением «на языке $\varepsilon - N$». Предположим противное, что $b > c$. Выберем $\varepsilon > 0$ так, чтобы выполнялось неравенство $c + \varepsilon < b - \varepsilon$.

Поскольку $\lim\limits_{n \to \infty} x_n = b$, существует номер N_1, начиная с которого выполняется неравенство $|x_n - b| < \varepsilon$, или, что то же самое,

$$b - \varepsilon < x_n < b + \varepsilon. \quad (3)$$

Аналогично, так как $\lim\limits_{n \to \infty} y_n = c$, то существует номер N_2, начиная с которого выполняется неравенство $|y_n - c| < \varepsilon$, или, что то же самое,

$$c - \varepsilon < y_n < c + \varepsilon. \quad (4)$$

Обозначим наибольшее из чисел N_1, N_2 через N. Тогда при $n \geqslant N$ будут выполнены неравенства (3) и (4). Поэтому $y_n < c + \varepsilon < b - \varepsilon < x_n$, т. е. $x_n > y_n$, что противоречит условию $x_n \leqslant y_n$ для всех n. Таким образом, сделанное предположение неверно и, значит, $b \leqslant c$. ■

Замечание. Теорема 3 верна только для нестрогих неравенств: из $x_n < y_n$ не следует, что $\lim\limits_{n\to\infty} x_n < \lim\limits_{n\to\infty} y_n$.

Рассмотрим, например, две последовательности: $x_n = -1/n$ и $y_n = 1/n$. Ясно, что $x_n < y_n$, но $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n = 0$. В то же время этот пример не противоречит теореме о предельном переходе под знаком нестрогого неравенства: вместо $x_n < y_n$ можно написать $x_n \leqslant y_n$, а вместо $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n$ можно написать $\lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$. Тогда получаем $x_n \leqslant y_n \Rightarrow \lim\limits_{n\to\infty} x_n \leqslant \lim\limits_{n\to\infty} y_n$.

Следующие два свойства непосредственно вытекают из теоремы о предельном переходе в неравенстве.

Следствие 1. *Если все члены сходящейся последовательности неотрицательны, то предел последовательности есть неотрицательное число.*

Следствие 2. *Если все члены сходящейся последовательности неположительны, то предел последовательности есть неположительное число.*

Отметим еще одно свойство сходящихся последовательностей.

Теорема 4. *Если $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} z_n = b$ и для всех n справедливо неравенство $x_n \leqslant y_n \leqslant z_n$,* то $\lim\limits_{n\to\infty} y_n = b$.

□ Воспользуемся определением «на языке $\varepsilon - N$». Возьмем произвольное $\varepsilon > 0$. Так как $\lim\limits_{n\to\infty} x_n = b$, то начиная с некоторого номера N_1 будет выполнено неравенство $|x_n - b| < \varepsilon$, или, что то же самое,

$$b - \varepsilon < x_n < b + \varepsilon. \quad (5)$$

Аналогично, поскольку $\lim\limits_{n\to\infty} z_n = b$, начиная с некоторого номера N_2 будет выполнено неравенство $|z_n - b| < \varepsilon$, или, что то же самое,

$$b - \varepsilon < z_n < b + \varepsilon. \quad (6)$$

Обозначив через N наибольший из номеров N_1, N_2, получим, что для всех $n \geqslant N$ будет выполнены неравенства (5) и (6). Воспользовавшись ими и заданным неравенством $x_n \leqslant y_n \leqslant z_n$, получим $b - \varepsilon < x_n \leqslant y_n \leqslant z_n < b + \varepsilon$, откуда $b - \varepsilon < y_n < b + \varepsilon$, или, что то же самое, $|y_n - b| < \varepsilon$.

Итак, мы доказали следующее:

$$(\forall \varepsilon > 0)(\exists N)(\forall n \geqslant N)\,|y_n - b| < \varepsilon,$$

а это и означает, что $\lim\limits_{n\to\infty} y_n = b$. ■

3. Вычисление пределов

Для вычисления пределов последовательностей часто используется теорема об арифметических операциях над пределами.

Теорема 5. *Пусть* $\lim\limits_{n\to\infty} x_n = a$, $\lim\limits_{n\to\infty} y_n = b$. *Тогда:*

1) $\lim\limits_{n\to\infty}(x_n + y_n) = a + b$; 2) $\lim\limits_{n\to\infty}(x_n y_n) = ab$;

3) $\lim\limits_{n\to\infty}\dfrac{x_n}{y_n} = \dfrac{a}{b}$ *(если* $b \neq 0$*).*

Иными словами, *предел суммы, произведения, частного равен соответственно сумме, произведению, частному пределов.*

□ Ограничимся доказательством второго утверждения. Так как $\lim\limits_{n\to\infty} x_n = a$, а $\lim\limits_{n\to\infty} y_n = b$, то $x_n = a + \alpha_n$, $y_n = b + \beta_n$, где (α_n) и (β_n) — бесконечно малые последовательности. Рассмотрим последовательность (z_n) такую, что $z_n = x_n y_n$. Имеем $z_n = x_n y_n = (a+\alpha_n)(b+\beta_n) = ab + (a\beta_n + \alpha_n b + \alpha_n\beta_n)$. Согласно свойствам бесконечно малых последовательностей (см. п. 1) последовательность $\gamma_n = a\beta_n + \alpha_n b + \alpha_n\beta_n$ является бесконечно малой. Значит, $z_n = ab + \gamma_n$, где γ_n — бесконечно малая последовательность, а это означает, что $\lim\limits_{n\to\infty} z_n = ab$, т. е.

$$\lim_{n\to\infty} x_n y_n = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n. \ \blacksquare$$

Примеры. 1. Вычислить $\lim\limits_{n\to\infty}\dfrac{n^2-5n+4}{2n^2-3n+7}$.

○ Разделив почленно числитель и знаменатель n-го члена заданной последовательности на n в наивысшей из имеющихся в числителе и знаменателе степеней, т. е. на n^2, получим

$$\frac{n^2-5n+4}{2n^2-3n+7} = \frac{1-\dfrac{5}{n}+\dfrac{4}{n^2}}{2-\dfrac{3}{n}+\dfrac{7}{n^2}}.$$

Следовательно,

$$\lim_{n\to\infty}\frac{n^2-5n+4}{2n^2-3n+7} = \lim_{n\to\infty}\frac{1-\dfrac{5}{n}+\dfrac{4}{n^2}}{2-\dfrac{3}{n}+\dfrac{7}{n^2}} =$$

$$= \frac{1-5\lim\limits_{n\to\infty}\dfrac{1}{n}+4\lim\limits_{n\to\infty}\dfrac{1}{n^2}}{2-3\lim\limits_{n\to\infty}\dfrac{1}{n}+7\lim\limits_{n\to\infty}\dfrac{1}{n^2}} = \frac{1-5\cdot 0+4\cdot 0}{2-3\cdot 0+7\cdot 0} = \frac{1}{2}. \ \bullet$$

2. Вычислить $\lim\limits_{n\to\infty}\dfrac{2n^2+n+1}{n^3-n-3}$.

○ Применим тот же прием, что в предыдущем примере; разделив числитель и знаменатель почленно на n^3, находим

$$\lim_{n\to\infty}\frac{\dfrac{2n^2}{n^3}+\dfrac{n}{n^3}+\dfrac{1}{n^3}}{\dfrac{n^3}{n^3}-\dfrac{n}{n^3}-\dfrac{3}{n^3}} = \lim_{n\to\infty}\frac{\dfrac{2}{n}+\dfrac{1}{n^2}+\dfrac{1}{n^3}}{1-\dfrac{1}{n^2}-\dfrac{3}{n^3}} = \frac{0+0+0}{1-0-0} = 0. \ \bullet$$

4. Признаки существования предела последовательности

Выше было доказано, что любая сходящаяся последовательность является ограниченной. Однако не всякая ограниченная последовательность имеет предел. Например, ограниченная последовательность 0, 1, 2, 3, 0, 1, 2, 3, 0, 1, ... не имеет предела.

Одним из условий, обеспечивающих существование предела, является монотонность ограниченной последовательности. Например, последовательность $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots, \frac{n}{n+1}, \ldots$ ограничена и возрастает. Эта последовательность сходится: $\lim\limits_{n\to\infty}\frac{n}{n+1}=1$.

Последовательность $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots$ ограничена и убывает. Эта последовательность также сходится: $\lim\limits_{n\to\infty}\frac{1}{n}=0$.

Оказывается, совокупность двух указанных признаков (ограниченность и монотонность) является достаточным условием сходимости последовательности. Именно, справедливы следующие теоремы.

Теорема 6. *Если последовательность возрастает (хотя бы в нестрогом смысле) и ограничена сверху, то она сходится.*

□ Согласно условию, последовательность (x_n) ограничена сверху; всякое ограниченное сверху множество имеет верхнюю грань (см. § 10, п. 2): $b=\sup(x_n)$. Рассмотрим произвольную ε — окрестность точки b (см. рис. 9).

Так как $b-\varepsilon$ уже не является верхней границей для множества (x_n), то найдется номер N такой, что $x_N>b-\varepsilon$. По условию, последовательность (x_n) возрастает: $x_N\leqslant x_{N+1}\leqslant x_{N+2}\leqslant x_{N+3}\leqslant\ldots$. Значит, все ее члены с номерами, бóльшими чем N, находятся между x_N и b (b — одна из верхних границ), т. е. во всяком случае все они принадлежат ε-окрестности точки b. На «языке окрестностей» это и означает, что $\lim\limits_{n\to\infty}x_n=b$. ■

Теорема 7. *Если последовательность убывает и ограничена снизу, то она сходится.*

Эта теорема доказывается аналогично.

Воспользуемся приведенными теоремами для получения важных результатов о связи между действительными числами и последовательностями их десятичных приближений по недостатку и по избытку.

Пусть a — положительное число и (x_n) — последовательность его десятичных приближений по недостатку. Эта последовательность возрастает (для любого n имеем $x_n\leqslant x_{n+1}$) и ограничена сверху (например, самим числом a). Значит, существует $\lim\limits_{n\to\infty}x_n$. Этот предел равен числу a, поскольку разность $a-x_n$ не превышает $1/10^n$ и может быть сделана сколь угодно малой. Точно так же доказывается, что предел убывающей последовательности десятичных приближений числа a по избытку равен a. Аналогичные результаты справедливы для отрицательных действительных чисел. Итак, любое действительное число является пределом последовательностей своих десятичных приближений по недостатку и по избытку.

Рассмотрим еще несколько примеров применения теорем 6 и 7.

Примеры. 1. Пусть S_n — площадь правильного 2^{n+1}-угольника, вписанного в окружность радиуса R. Доказать, что последовательность (S_n) имеет предел.

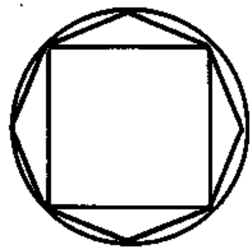


Рис. 10

○ Из рис. 10 видно, что при удвоении числа сторон площадь многоугольника увеличивается. Поэтому последовательность (S_n) возрастает. Все ее члены не превышают числа $4R^2$ — площади описанного квадрата. Значит, последовательность (S_n) ограничена сверху. Согласно теореме 6, она имеет предел, который равен площади круга радиуса R. ●

2. Доказать, что если $|q|<1$, то $\lim\limits_{n\to\infty} q^n=0$; если же $|q|>1$, то последовательность (q^n) расходится.

□ Ясно, что если $q=0$, то $\lim\limits_{n\to\infty} q^n=0$. Пусть $0<q<1$. Рассмотрим последовательность $q, q^2, q^3, \dots, q^n, \dots$. Она является убывающей: $q>q^2>q^3>\dots>q^n>\dots$ и ограниченной снизу (например, нулем). Значит, существует предел: $\lim\limits_{n\to\infty} q^n=a$. Так как $q^{n+1}=q\cdot q^n$, то $a=\lim\limits_{n\to\infty} q^{n+1}=q\cdot\lim\limits_{n\to\infty} q^n=qa$. Итак, $a=qa$, откуда (ввиду $q\neq 1$) имеем $a=0$.

Пусть $-1<q<0$. Тогда $0<|q|<1$ и потому $\lim\limits_{n\to\infty}|q|^n=0$, т. е. $\lim\limits_{n\to\infty}|q^n|=0$. Это означает, что $|q^n|$, а следовательно, и (q^n) — бесконечно малая последовательность, т. е. $\lim\limits_{n\to\infty} q^n=0$.

Таким образом, если $-1<q<1$, то $\lim\limits_{n\to\infty} q^n=0$.

Пусть $|q|>1$. Тогда последовательность (q^n) расходится. В самом деле, если бы она сходилась и имела своим пределом число b, то $b=0$ (это устанавливается так же, как было сделано выше для числа a). Однако это невозможно, поскольку при $|q|>1$ справедливо неравенство $|q^n|>1$, а потому при $\varepsilon<1$ не может выполняться неравенство $|q^n-0|<\varepsilon$. ●

В качестве еще одного примера рассмотрим последовательность (x_n), образующуюся при извлечении квадратного корня из положительного числа *методом последовательных приближений.*

Пусть нужно извлечь квадратный корень из положительного числа a. Возьмем какое-нибудь приближенное значение x_1 этого корня по избытку. Поскольку при $a\geqslant 1$ имеем $a^2\geqslant a$ и, значит, $a\geqslant\sqrt{a}$, а при $0<a<1$ имеем $1\geqslant\sqrt{a}$, в качестве x_1 можно взять большее из чисел a и 1, т. е. $x_1=\max(a,1)$. Тогда a/x_1 — приближенное значение корня по недостатку. Рассмотрим среднее арифметическое чисел x_1 и a/x_1, т. е. $\frac{1}{2}\left(x_1+\frac{a}{x_1}\right)$. Так как среднее арифметическое двух положительных чисел не меньше их среднего геометрического $\left(\frac{a+b}{2}\geqslant\sqrt{ab}\right)$, то

$$x_2=\frac{1}{2}\left(x_1+\frac{a}{x_1}\right)\geqslant\sqrt{x_1\cdot\frac{a}{x_1}}=\sqrt{a}\,.$$

Значит, x_2 — приближенное значение корня по избытку. Тогда a/x_2 — приближенное значение корня по недостатку. Рассуждая, как и выше, находим более точное приближение корня по избытку $x_3 = \frac{1}{2}\left(x_2 + \frac{a}{x_2}\right)$. Вообще, если уже найдено приближенное значение x_n или $\sqrt{a}$, то следующее приближение выражается формулой

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right). \tag{7}$$

Таким образом, мы пришли к рекуррентному заданию последовательности:

$$x_1 = \max(a, 1),\ x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Докажем, что *предел этой последовательности существует и равен* $\sqrt{a}$.

□ Предположим, что эта последовательность имеет предел: $\lim\limits_{n\to\infty} = b$. Найдем значение этого предела.

Так как значение предела не зависит от добавления или отбрасывания конечного числа членов последовательности, то $\lim\limits_{n\to\infty} x_{n+1}$ также равен b. Перейдем в равенстве $x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ к пределу при $n \to \infty$:

$$\lim_{n\to\infty} x_{n+1} = \frac{1}{2}\lim_{n\to\infty} x_n + \frac{a}{2\lim\limits_{n\to\infty} x_n}.$$

Значит, $b = \frac{1}{2}b + \frac{a}{2b}$, откуда находим $b = \sqrt{a}$.

Итак, $\lim\limits_{n\to\infty} x_n = \sqrt{a}$, т. е. если предел существует, то он равен $\sqrt{a}$. Докажем теперь, что предел действительно существует.

Предположим сначала, что на каком-то шаге выполняется равенство $x_n = \sqrt{a}$. Тогда из сооотношения (7) получаем, что и $x_{n+1} = \sqrt{a}$, $x_{n+2} = \sqrt{a}$ и т. д. Ясно, что в таком случае $\lim\limits_{n\to\infty} x_n = \sqrt{a}$.

Пусть равенство $x_n = \sqrt{a}$ не выполняется ни на каком шаге. Тогда x_n и a/x_n — неравные положительные числа, и их среднее арифметическое больше среднего геометрического. Поэтому для любого n имеем

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) > \sqrt{x_n \cdot \frac{a}{x_n}} = \sqrt{a}.$$

Это означает, что последовательность (x_n) ограничена снизу и $\sqrt{a}$ — одна из нижних границ.

Далее, для любого n находим

$$x_{n+1} - x_n = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - x_n = \frac{a - x_n^2}{2x_n}.$$

Так как $x_n > \sqrt{a}$, то $a - x_n^2 < 0$ и потому $x_{n+1} < x_n$, а это означает, что последовательность (x_n) убывает.

Итак, мы доказали, что последовательность (x_n) ограничена снизу и убывает. По теореме 7 она имеет предел, что и требовалось установить. ■

Вычисляя, например, методом последовательных приближений число $\sqrt{3}$, получаем $x_1 = 3$, $x_2 = 2$, $x_3 = 1{,}75$, $x_4 \approx 1{,}73214$, $x_5 \approx 1{,}73205$, ... Каждый следующий шаг процесса приводит ко все более точным приближениям для $\sqrt{3}$. Процесс останавливают, когда разность между x_{n+1} и x_n становится меньше, чем требуемая точность вычислений. Например, если требуется вычислить $\sqrt{3}$ с точностью 0,001, то достаточно взять пять приближений и положить $\sqrt{3} = 1{,}732$.

5. Сумма бесконечно убывающей геометрической прогрессии

Возьмем отрезок $[0, 1]$ и разобьем его пополам. Правую половину отрезка, т. е. отрезок $[1/2, 1]$, снова разобьем пополам. Затем разобьем пополам правую половину отрезка $[1/2, 1]$, т. е. отрезок $[3/4, 1]$, и т. д. Продолжая этот процесс до бесконечности, получим разбиение отрезка на бесконечное множество отрезков $[0, 1/2]$, $[1/2, 3/4]$, $[3/4, 7/8]$, ... (рис. 11). Естественно считать, что «сумма» длин всех отрезков, на которые разбит отрезок $[0, 1]$, равна длине разбиваемого отрезка, т. е. единице. Иными словами, естественно считать верным следующее «равенство»:

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots+\frac{1}{2^n}+\ldots=1. \tag{8}$$

Левая часть «равенства» (8) представляет собой «сумму бесконечного множества слагаемых». Кавычки здесь использованы потому, что такое выра-

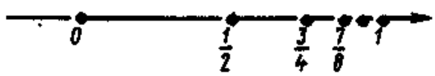


Рис. 11

жение нельзя понимать буквально. Речь идет не об обычной сумме (сумме конечного числа слагаемых), а о чем-то таком, что еще нужно правильно понять и истолковать.

При первом делении пополам отрезка $[0, 1]$ мы взяли отрезок $[0, 1/2]$; его длина $S_1 = 1/2$. На втором шаге к этому отрезку мы добавили отрезок $[1/2, 3/4]$; его длина равна $1/4$. В результате из двух отрезков $[0, 1/2]$ и $[1/2, 3/4]$ составился отрезок $[0, 3/4]$ с длиной S_2, равной $1/2+1/4=3/4$. На третьем шаге к первым двум отрезкам добавился отрезок $[3/4, 7/8]$, в результате из трех отрезков получился отрезок $[0, 7/8]$ с длиной S_3, равной $1/2+1/4+1/8=7/8$.

Вообще, на n-м шаге n отрезков составят отрезок длины S_n, где

$$S_n=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots+\frac{1}{2^n}.$$

Ясно, что чем больше n, тем ближе S_n подходит к единице — к длине отрезка $[0, 1]$, т. е. $\lim\limits_{n\to\infty} S_n = 1$. Именно такой смысл мы, видимо, и должны вложить в равенство (8).

Слагаемые в левой части равенства (8) образуют последовательность $1/2$, $1/4$, $1/8$, ..., $1/2^n$, ..., являющуюся геометрической прогрессией. Проведенное выше рассуждение позволяет говорить не только о сумме S_n первых n членов этой прогрессии, но и о сумме S всех членов прогрессии, понимая под этой суммой число $\lim\limits_{n\to\infty} S_n$.

Оказывается, понятие суммы можно ввести для любой бесконечной геометрической прогрессии (b_n), у которой $|q|<1$, где q — знаменатель прогрессии; такую прогрессию часто называют *бесконечно убывающей*. Дело в том, что для этой прогрессии существует $\lim\limits_{n\to\infty} S_n$ и этот предел естественно называть суммой прогрессии.

В самом деле, пусть (b_n) — геометрическая прогрессия и $|q|<1$. Тогда

$$S_n=\frac{b_1(q^n-1)}{q-1}.$$

Перейдем в этом равенстве к пределу при $n \to \infty$. Воспользуемся теоремами об арифметических операциях над пределами и тем, что $\lim_{n\to\infty} q^n = 0$, если $|q|<1$ (см. пример 2 п. 4); получим

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{b_1(q^n-1)}{q-1} = \frac{b_1}{q-1} \lim_{n\to\infty}(q^n-1) = \frac{b_1}{q-1}(0-1) = \frac{b_1}{1-q}.$$

Итак,

$$S = \lim_{n\to\infty} S_n = \frac{b_1}{1-q}. \tag{9}$$

Определение 5. Под *суммой бесконечно убывающей геометрической прогрессии* $b_1, b_2, b_3, \ldots, b_n, \ldots$ понимают предел последовательности (S_n), где $S_1 = b_1$, $S_2 = b_1 + b_2$, $S_3 = b_1 + b_2 + b_3$, ..., $S_n = b_1 + b_2 + b_3 + \ldots + b_n$, ..., т. е.

$$S = b_1 + b_2 + b_3 + \ldots + b_n + \ldots = \lim_{n\to\infty} S_n.$$

Как было показано выше, эта сумма находится по формуле (9).

Примеры. 1. Найти сумму геометрической прогрессии 15, 6, 12/5, 24/25, 48/125, ...

○ Так как $b_1 = 15$, $b_2 = 6$, то $q = b_2/b_1 = 6/15 = 2/5$, т. е. $|q|<1$. Используя формулу (9), находим

$$S = \frac{b_1}{1-q} = \frac{15}{1-\frac{2}{5}} = 25. \ ●$$

2. Сумма геометрической прогрессии, у которой $|q|<1$, равна 9, а сумма квадратов ее членов равна 40, 5. Найти первый член и знаменатель прогрессии.

○ Обозначим через b_1 и q соответственно первый член и знаменатель искомой прогрессии. Используя формулу суммы геометрической прогрессии, получим $\frac{b_1}{1-q} = 9$.

Рассмотрим последовательность $b_1^2, b_2^2, b_3^2, \ldots, b_n^2, \ldots$. Имеем $b_2 = b_1 q$, т. е. $b_2^2 = b_1^2 q^2$; $b_3 = b_2 q$, т. е. $b_3^2 = b_2^2 q^2$; $b_4 = b_3 q$, т. е. $b_4^2 = b_3^2 q^2$, и т. д. Замечаем, что каждый член последовательности (b_n^2) получается из предыдущего умножением на q^2. Это означает, что (b_n^2) — геометрическая прогрессия, первый член которой равен b_1^2, а знаменатель равен q^2. Из условия $|q|<1$ следует, что $q^2<1$. Тогда сумму S прогрессии (b_n^2) можно вычислить по формуле $S = \frac{b_1^2}{1-q^2}$, а по условию эта сумма равна 40,5.

Таким образом, задача сводится к решению системы уравнений

$$\begin{cases} \dfrac{b_1}{1-q} = 9, \\ \dfrac{b_1^2}{1-q^2} = 40{,}5, \end{cases}$$

из которой находим $b_1 = 6$, $q = 1/3$. ●

6. Число e

Теорема 8. *Последовательность* $x_n = \left(1+\frac{1}{n}\right)^n$ *сходится.*

□ Рассмотрим сначала вспомогательную последовательность $y_n = \left(1+\frac{1}{n}\right)^{n+1}$.

Докажем, что она убывает, т. е. что $y_n > y_{n+1}$ для любого n. Имеем $y_n = \left(1+\frac{1}{n}\right)^{n+1}$, $y_{n+1} = \left(1+\frac{1}{n+1}\right)^{n+2}$; поэтому

$$\frac{y_n}{y_{n+1}}=\left(\frac{n+1}{n}\right)^{n+1}:\left(\frac{n+2}{n+1}\right)^{n+2}=\frac{(n+1)^{2n+3}}{n^{n+1}\cdot(n+2)^{n+2}}=$$

$$=\left(\frac{(n+1)^2}{n(n+2)}\right)^{n+1}\frac{n+1}{n+2}=\left(\frac{n^2+2n+1}{n^2+2n}\right)^{n+1}\frac{n+1}{n+2}=$$

$$=\left(1+\frac{1}{n^2+2n}\right)^{n+1}\frac{n+1}{n+2}.$$

Применим к выражению $\left(1+\frac{1}{n^2+2n}\right)^{n+1}$ неравенство Бернулли $(1+x)^n>$ $>1+nx$ (см. пример 3 § 11) для случая, когда $x=\frac{1}{n^2+2n}$. Получим

$$\left(1+\frac{1}{n^2+2n}\right)^{n+1}\geqslant 1+(n+1)\frac{1}{n^2+2n}>1+(n+1)\frac{1}{n^2+2n+1}=$$

$$=1+\frac{1}{n+1}=\frac{n+2}{n+1}.$$

Тогда $\frac{y_n}{y_{n+1}}>\frac{n+2}{n+1}\frac{n+1}{n+2}$. Значит, $\frac{y_n}{y_{n+1}}>1$, т. е. $y_n>y_{n+1}$.

Докажем теперь, что последовательность (y_n) ограничена снизу. Снова воспользуемся неравенством Бернулли: $\left(1+\frac{1}{n}\right)^{n+1}\geqslant 1+(n+1)\cdot\frac{1}{n}>2$. Следовательно, для любого n имеем $y_n>2$; значит, последовательность (y_n) ограничена снизу.

Итак, последовательность (y_n) убывает и ограничена снизу; поэтому в силу теоремы 7 она сходится. Тогда сходится и последовательность (x_n). В самом деле,

$$\lim_{n\to\infty}x_n=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=\lim_{n\to\infty}\frac{\left(1+\frac{1}{n}\right)^{n+1}}{\left(1+\frac{1}{n}\right)}=$$

$$=\frac{\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}}{\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)}=\frac{\lim\limits_{n\to\infty}y_n}{1}=\lim_{n\to\infty}y_n. \blacksquare$$

Выпишем несколько членов последовательности $x_n=\left(1+\frac{1}{n}\right)^n$:

$$x_1=\left(1+\frac{1}{1}\right)^1=2;\quad x_2=\left(1+\frac{1}{2}\right)^2=2{,}25;\quad x_3=\left(1+\frac{1}{3}\right)^3=2{,}37;$$

$$x_4=\left(1+\frac{1}{4}\right)^4=2{,}44;\quad x_5=\left(1+\frac{1}{5}\right)^5=2{,}49;\ \ldots$$

Вычисляя x_n для бо́льших номеров n, мы будем ближе подходить к пределу последовательности (x_n). В математике этот предел обозначают буквой e; он вычислен с большой точностью; $e=2{,}718281284590\ldots$; доказано, что e — иррациональное число. Число e, как мы не раз убедимся в дальнейшем, играет в математическом анализе особую роль.

7. Бесконечно большие последовательности

В п. 1 было дано определение бесконечно малой последовательности как такой последовательности, члены которой с увеличением номера могут сделаться по модулю меньше любого наперед заданного положительного числа. Встречаются и последовательности, члены которых с увеличением номера могут сделаться по модулю больше любого наперед заданного числа; такие последовательности называют бесконечно большими.

Определение 6. Последовательность (x_n) называют *бесконечно большой*, если для любого $P>0$ существует номер N, начиная с которого каждый член последовательности (x_n) по модулю больше P. Короче:

$$(\forall P>0)(\exists N)(\forall n\geqslant N)|x_n|>P.$$

Для бесконечно большой последовательности (x_n) используется следующая условная запись: $\lim\limits_{n\to\infty} x_n=\infty$.

Из определения следует, что всякая бесконечно большая последовательность является неограниченной, а потому расходящейся (см. теорему 2).

Между бесконечно малой и бесконечно большой последовательностями существует естественная связь.

Теорема 9. *Для того чтобы последовательность (x_n) была бесконечно большой, необходимо и достаточно, чтобы последовательность $\alpha_n=1/x_n$ была бесконечно малой.*

□ Необходимость. Пусть (x_n) — бесконечно большая последовательность. Возьмем произвольное $\varepsilon>0$ и положим $P=1/\varepsilon$. Существует номер N, начиная с которого выполняется неравенство $|x_n|>P$. Тогда $1/|x_n|<1/P$, т. е. $|\alpha_n|<\varepsilon$. Это и означает, что (α_n) — бесконечно малая последовательность.

Достаточность. Пусть (α_n) — бесконечно малая последовательность. Возьмем произвольное $P>0$ и положим $\varepsilon=1/P$. Существует номер N, начиная с которого выполняется неравенство $|\alpha_n|<\varepsilon$. Тогда $1/|\alpha_n|>1/\varepsilon$, т. е. $|x_n|>P$. Это и означает, что (x_n) — бесконечно большая последовательность. ■

Из этой теоремы, например, следует, что поскольку последовательности $\left(\frac{1}{n}\right)$, $\left(\frac{1}{n^2}\right)$, $\left(\frac{c}{an+b}\right)$ — бесконечно малые (см. п. 1), последовательности (n), (n^2), $\left(\frac{an+b}{c}\right)$ являются бесконечно большими.

Отметим одно свойство бесконечно больших последовательностей, часто используемое в приложениях.

Теорема 10. *Если (x_n) — бесконечно большая последовательность, а (y_n) — сходящаяся последовательность, не являющаяся бесконечно малой, то их произведение есть бесконечно большая последовательность.*

□ Пусть $z_n=x_ny_n$. Положим $\alpha_n=1/z_n$, откуда имеем $\alpha_n=1/(x_ny_n)=1/(x_n)\cdot 1/(y_n)$.

По условию, (y_n) — сходящаяся (но не бесконечно малая) последовательность; значит, $\lim\limits_{n\to\infty} y_n=b\neq 0$. Тогда $\lim\limits_{n\to\infty}(1/y_n)=1/b$. Так как последовательность $(1/y_n)$ сходится, то по теореме 2 она ограничена; последовательность же $(1/x_n)$ является бесконечно малой (по теореме 9). Поэтому α_n как произведение бесконечно малой последовательности на ограниченную есть бесконечно малая (согласно свойству 6° из п. 1). Отсюда следует, что $z_n=1/\alpha_n$ — бесконечно большая последовательность. ■

Примеры. 1. Доказать, что $x_n = n^2 - 2n + 3$ — бесконечно большая последовательность.

○ Имеем $x_n = n^2\left(1 - \frac{2}{n} + \frac{3}{n^2}\right) = y_n z_n$, где $y_n = n^2$ — бесконечно большая последовательность, а $z_n = 1 - \frac{2}{n} + \frac{3}{n^2}$ — сходящаяся последовательность $(\lim\limits_{n\to\infty} z_n = 1)$. Значит, согласно теореме 10, (x_n) — бесконечно большая последовательность. ●

2. Вычислить $\lim\limits_{n\to\infty} \frac{n^3 - n - 3}{2n^2 + n + 1}$.

○ Выше (см. пример 2 п. 3) мы доказали, что $\left(\frac{2n^2+n+1}{n^3-n-3}\right)$ — бесконечно малая последовательность. Следовательно, $\left(\frac{n^3-n-3}{2n^2+n+1}\right)$ — бесконечно большая последовательность, т. е. $\lim\limits_{n\to\infty} \frac{n^3-n-3}{2n^2+n+1} = \infty$. ●

Замечание. В условиях теоремы 10 оговорка о том, что сходящаяся последовательность не является бесконечно малой, существенна: о произведении бесконечно большой и бесконечно малой последовательностей ничего определенного сказать нельзя. Например:

1) $\alpha_n = \frac{1}{n}$ — бесконечно малая последовательность, $x_n = n^2$ — бесконечно большая последовательность, а их произведение $\alpha_n x_n = n$ — бесконечно большая последовательность; $\lim\limits_{n\to\infty} y_n = \infty$.

2) $\alpha_n = \frac{1}{n^2}$ — бесконечно малая последовательность, $x_n = n$ — бесконечно большая последовательность, а их произведение $\alpha_n x_n = \frac{1}{n}$ — бесконечно малая последовательность; $\lim\limits_{n\to\infty} y_n = 0$;

3) $\alpha_n = \frac{1}{n^2}$ — бесконечно малая последовательность, $x_n = n^2$ — бесконечно большая последовательность, а их произведение $\alpha_n x_n = 1$ — стационарная последовательность; $\lim\limits_{n\to\infty} y_n = 1$;

4) $\alpha_n = \frac{(-1)^n}{n}$ — бесконечно малая последовательность, $x_n = n$ — бесконечно большая последовательность, а их произведение $\alpha_n x_n = (-1)^n$, т. е. последовательность $-1, 1, -1, 1, -1, 1, \ldots$ расходится, причем она не является бесконечно большой.

4

Функции одной переменной

§ 14. Свойства функций

1. Основные понятия

Напомним известное из школьного курса определение функции одной переменной.

Определение 1. Зависимость переменной y от переменной x называется *функцией*, если каждому значению x соответствует единственное значение y. Переменную x называют *независимой переменной* (или *аргументом*), а переменную y — *зависимой переменной*. Значение y, соответствующее заданному значению x, называют *значением функции*.

Если переменная y является функцией от переменной x, то пишут $y = f(x)$. Буквой f обозначают данную функцию, т. е. функциональную зависимость между переменными x и y; $f(x)$ есть значение функции, соответствующее значению аргумента x. Говорят также, что $f(x)$ есть *значение функции в точке* x.

Все значения, которые принимает независимая переменная, образуют *область определения функции*.

Все значения, которые принимает функция $f(x)$ (при x, принадлежащих области ее определения), образуют *область значений функции*.

Чтобы задать функцию, нужно указать способ, с помощью которого для каждого значения аргумента можно найти соответствующее значение функции. Наиболее употребительным является способ задания функции с помощью формулы $y = f(x)$, где $f(x)$ — некоторое выражение с переменной x. В таком случае говорят, что *функция задана формулой* или что *функция задана аналитически*.

Для аналитически заданной функции иногда не указывают явно область определения функции. В таком случае подразумевают, что область определения функции $y = f(x)$ совпадает с областью определения выражения $f(x)$, т. е. с множеством тех значений x, при которых выражение $f(x)$ имеет смысл.

Примеры. 1. Функция $f(x)$ задана аналитически формулой $f(x) = x^2 + 5x - 1$, где $x \geqslant 0$. Область определения функции — луч $[0, +\infty]$. Чтобы найти значение функции в любой точке $x \geqslant 0$, достаточно найти числовое значение выражения $x^2 + 5x - 1$ в выбранной точке. Например, $f(2) = 2^2 + 5\cdot 2 - 1 = 13$.

2. Функция $y = f(x)$ задана аналитически формулой $f(x) = \frac{x^2+x+1}{x^4+3}$. Найти: а) $f(-x)$; б) $f(kx)$; в) $f(x + a)$; г) $f(|x|)$.

○ а) Чтобы найти $f(-x)$, надо в $f(x)$ всюду вместо x подставить $(-x)$. Получим $f(-x)=\frac{(-x)^2+(-x)+1}{(-x)^4+3}=\frac{x^2-x+1}{x^4+3}$.

Аналогично находим:

б) $y=f(kx)=\frac{(kx)^2+(kx)+1}{(kx)^4+3}=\frac{k^2x^2+kx+1}{k^4x^4+3}$;

в) $y=f(x+a)=\frac{(x+a)^2+(x+a)+1}{(x+a)^4+3}$;

г) $y=f(|x|)=\frac{|x|^2+|x|+1}{|x|^4+3}=\frac{x^2+|x|+1}{x^4+3}$. ●

3. Найти область определения функции $y=\sqrt{x-1}$.

○ Выражение $\sqrt{x-1}$ определено при тех x, для которых $x-1\geqslant 0$, т. е. при $x\geqslant 1$. Значит, область определения функции — луч $[1, +\infty]$. ●

Иногда функция задается на различных промежутках различными формулами, например:

$$f(x)=\begin{cases} 2x+3 & \text{при } -1\leqslant x\leqslant 0, \\ x+2 & \text{при } \quad 0<x\leqslant 1. \end{cases}$$

Эта функция определена на отрезке $[-1, 1]$. Для вычисления значений функции нужно прежде всего определить, какой формулой следует воспользоваться для заданного конкретного значения аргумента. Например, если требуется вычислить $f(0,5)$, то, поскольку $0,5\in[0, 1]$, пользуемся равенством $f(x)=x+2$ и получаем $f(0,5)=2,5$. Если же требуется вычислить $f(-0,5)$, то пользуемся равенством $f(x)=2x+3$ и получаем $f(-0,5)=2$.

На практике часто используется *табличный способ* задания функции, когда задается таблица, содержащая значения функции для имеющихся значений аргумента. Примерами табличного задания функции являются таблица квадратов, таблица кубов, таблица квадратных корней. Во многих случаях табличное задание функции оказывается более удобным, чем другие способы задания, а иногда и единственно возможным. На практике зависимость одной величины от другой часто находят опытным путем. В этом случае одной величине придают определенные значения, а затем из опыта для каждого из таких значений находят значение (обычно приближенное) второй величины. В результате получают таблицу значений функции. Существуют методы, которые позволяют по такой таблице подбирать формулы, задающие функции с определенной точностью.

Весьма распространенным, особенно в экспериментальных науках, является *графический способ* задания функции. Рассмотрим множество F точек координатной плоскости xy, обладающее следующим свойством: любая прямая, параллельная оси ординат, пересекает это множество не более чем в одной точке (рис. 12). Пусть множество абсцисс всех точек множества F представляет собой отрезок $[a, b]$. Возьмем произвольное число x из отрезка $[a, b]$ и проведем через точку x прямую, параллельную оси ординат. Эта прямая пересекает F в точке M. Спроецировав точку M на ось ординат, найдем число $f(x)$, соответствующее числу x. Тем самым на отрезке $[a, b]$ задана функция $y=f(x)$.

Определение 2. Пусть функция задана аналитически формулой $y=f(x)$. Отметим на координатной плоскости все точки, обладающие следующим свойством: абсцисса точки принадлежит области определения функции, а ордината равна соответствующему значению функции. Множество точек $(x; f(x))$ называют *графиком функции* $y=f(x)$.

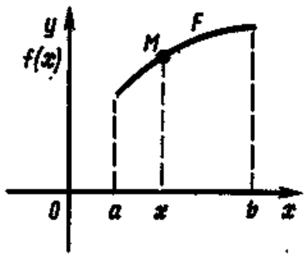


Рис. 12

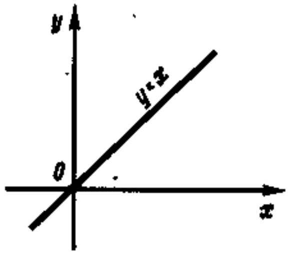


Рис. 13

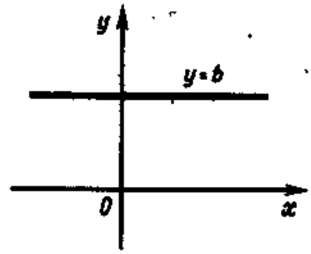


Рис. 14

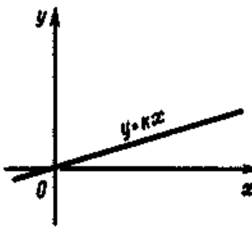


Рис. 15

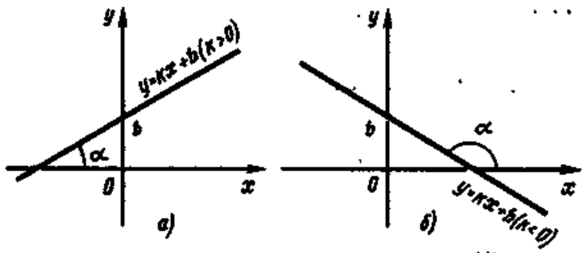


Рис. 16

Так, графиком функции $y = x$ является множество точек вида $(x;\ x)$, т. е. точек, имеющих одинаковые координаты. Это множество точек есть биссектриса I и III координатных углов (рис. 13).

На практике для построения графика функции составляют таблицу значений функции при некоторых значениях аргумента, наносят на плоскость соответствующие точки и соединяют их плавной линией.

Напомним некоторые известные из школьного курса функции и их графики.

1. *Линейная функция*, т. е. функция вида $y = kx + b$. Если $k = 0$, то получаем постоянную функцию $y = b$; ее график изображен на рис. 14. Если $b = 0$, то получаем прямую пропорциональность $y = kx$; ее графиком является прямая, проходящая через начало координат (рис. 15). Если $k \neq 0$, $b \neq 0$, то получаем линейную функцию общего вида, графиком которой является прямая; на рис. 16, *а* она изображена для случая $k > 0$, на рис. 16, *б* — для случая $k < 0$. Число k называется *угловым коэффициентом*; оно равно $\operatorname{tg}\alpha$, где α — угол между прямой и осью x.

2. *Обратная пропорциональность*, т. е. функция вида $y = k/x$, где $k \neq 0$. Ее графиком является *гипербола*, расположенная в I и III координатных углах при $k > 0$ (рис. 17, *а*) и во II и IV координатных углах при $k < 0$ (рис. 17, *б*).

3. Функция $y = ax^2$ $(a \neq 0)$; ее графиком является *парабола* с вершиной в точке (0; 0) и с ветвями, направленными вверх при $a > 0$ и вниз при $a < 0$ (рис. 18).

4. *Квадратичная функция*, т. е. функция вида $y = ax^2 + bx + c$, где $a \neq 0$. Ее графиком является парабола, полученная из параболы $y = ax^2$ параллельным переносом, при котором вершина параболы переходит в точку $\left(-\frac{b}{2a};\ \frac{4ac - b^2}{4a}\right)$ (рис. 19).

5. Функция $y = ax^3$ $(a \neq 0)$; ее графиком является *кубическая парабола* (рис. 20).

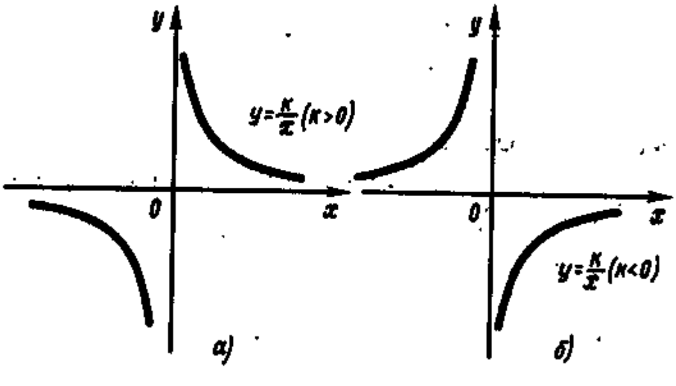


Рис. 17

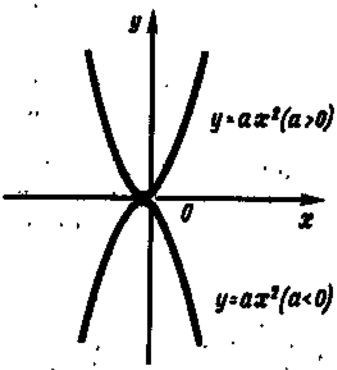


Рис. 18

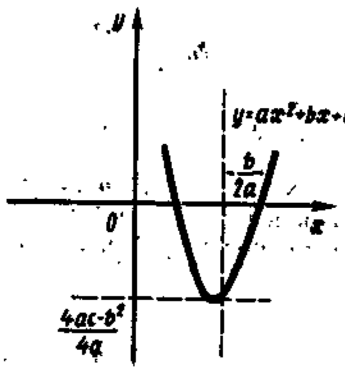


Рис. 19

Существуют функции, графики которых изобразить невозможно. Примером может служить функция

$$D(x)=\begin{cases} 1, \text{ если } x \text{ — рациональное число;} \\ 0, \text{ если } x \text{ — иррациональное число.} \end{cases}$$

Она называется *функцией Дирихле**. Ее график состоит из точек оси x с иррациональными абсциссами и точек прямой $y=1$ с рациональными абсциссами. Поскольку на любом сколь угодно малом отрезке числовой прямой есть как рациональные, так и иррациональные точки, нарисовать такой график невозможно.

Существуют и другие способы задания функций кроме аналитического, табличного и графического, с ними мы познакомимся в дальнейших разделах курса. Здесь упомянем еще лишь *словесный способ* задания, когда функция описывается правилом ее составления. Такова, например, функция *целая часть числа*: наибольшее целое число, не превосходящее числа x. Целую часть числа x обозначают $[x]$. Например, $[5{,}2]=5$; $[-3{,}756]=-4$; $[7]=7$. Наряду с функцией «целая часть числа», рассматривают и функцию *дробная часть числа* — это разность между числом x и его целой частью. Дроб-

* Л. Дирихле (1805—1859) — немецкий математик.

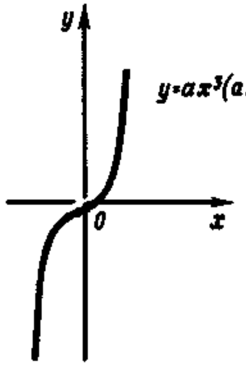


Рис. 20

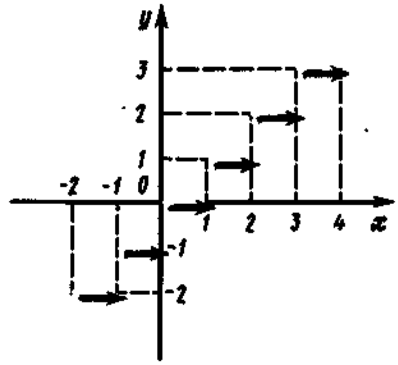


Рис. 21

ную часть числа обозначают $\{x\}$. Например, $\{5{,}2\}=0{,}2$; $\{-3{,}756\}= 0{,}244$ $(-3{,}756-(-4)=0{,}244)$; $\{7\}=0$.

Пример 4. Построить график функции $y=[x]$.

○ Если n — целое число, то на промежутке $[n,\ n+1)$ имеем $[x]=n$. Например, $[x]=0$ на промежутке $[0,\ 1)$; $[x]=1$ на промежутке $[1,\ 2)$; $[x]=-1$ на промежутке $[-1,\ 0)$. График функции (рис. 21) состоит из горизонтальных отрезков, у которых исключены их правые концы (отмеченные стрелками). ●

2. Четные и нечетные функции

Определение 3. Функция $y=f(x)$ называется *четной*, если для любого x из области определения функции выполняется равенство $f(-x)=f(x)$. Функция $y=f(x)$ называется *нечетной*, если для любого x из области определения функции выполняется равенство $f(-x)=-f(x)$.

Если функция $y=f(x)$ такова, что хотя бы для одной пары значений x и $-x$ оказалось, что $f(-x)\neq -f(x)$, и хотя бы для одной пары значений x и $-x$ оказалось, что $f(-x)\neq f(x)$, то функция не является ни четной, ни нечетной.

Из определения следует, что область определения X как четной, так и нечетной функции должна обладать следующим свойством: если $x\in X$, то и $-x\in X$ (т. е. X — симметричное относительно 0 множество).

Пример 1. Исследовать на четность функции: а) $y=x^2$; б) $y=x^3$; в) $y=\frac{x-4}{x^2-9}$.

○ а) Имеем $f(x)=x^2$, $f(-x)=(-x)^2=x^2$. Значит, $f(-x)=f(x)$ для всех x, т. е. функция является четной.

б) Имеем $f(x)=x^3$, $f(-x)=(-x)^3=-x^3$. Поэтому $f(-x)=-f(x)$ для всех x, т. е. функция является нечетной.

в) Имеем $f(x)=\frac{x-4}{x^2-9}$, $f(-x)=\frac{-x-4}{(-x)^2-9}=-\frac{x+4}{x^2-9}$. Так как $f(-x)\neq f(x)$ и $f(-x)\neq -f(x)$, то функция не является ни четной, ни нечетной. ●

Следующие теоремы (одна из которых доказывается, а доказательство другой предлагается провести самостоятельно) выявляют особенности графиков четных и нечетных функций.

Теорема 1. *Если функция $y=f(x)$, $x\in X$ является четной, то ее график симметричен относительно оси ординат.*

□ Пусть M $(x;\ f(x))$ — точка графика рассматриваемой функции. Так как по условию функция четна, то, во-первых, $(-x)\in X$, и во-вторых, $f(-x)=f(x)$.

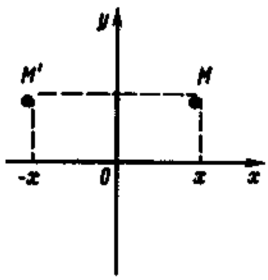


Рис. 22

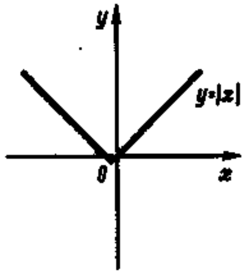


Рис. 23

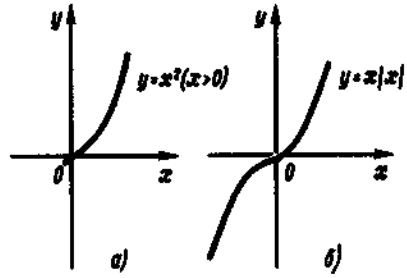


Рис. 24

Значит, точка $M'(-x;\ f(x))$ также принадлежит графику функции. Но точки M и M' симметричны относительно оси ординат (рис. 22). Таким образом, график четной функции вместе с каждой своей точкой содержит точку, симметричную с ней относительно оси ординат. Поэтому график четной функции симметричен относительно оси ординат. ■

Теорема 2. *Если функция $y=f(x)$, $x\in X$ является нечетной, то ее график симметричен относительно начала координат.*

Примеры. 2. Построить график функции $y=|x|$.

○ Данная функция четна, поэтому ее график симметричен относительно оси ординат. Если $x\geqslant 0$, то $|x|=x$, т. е. при $x\geqslant 0$ имеем $y=x$. Графиком функции $y=x$, $x\geqslant 0$ служит биссектриса I координатного угла. Добавив к ней симметричный относительно оси y луч, получим график функции $y=|x|$ (рис. 23). ●

3. Построить график функции $y=x|x|$.

○ Эта функция является нечетной, а потому ее график симметричен относительно начала координат. Если $x\geqslant 0$, то $|x|=x$, а $f(x)=x|x|=x\cdot x=x^2$. Значит, при $x\geqslant 0$ имеем $y=x^2$. Графиком является ветвь параболы, изображённая на рис. 24, *а*. Добавив к ней симметричную относительно начала координат ветвь, получим график функции $y=x|x|$ (рис. 24, *б*). ●

3. Периодические функции

Определение 4. Говорят, что функция $f(x)$ имеет *период T*, если для любого значения x, при котором она определена, выполняются равенства $f(x-T)=f(x)=f(x+T)$. Функция $f(x)$, имеющая отличный от нуля период T, называется *периодической*.

Из этого определения следует, что если функция $f(x)$ с периодом T определена в точке x, то она определена и в точках $x+T$, $x-T$.

Теорема 3. *Если функция $f(x)$ имеет период T, то любое число, кратное T (т. е. число вида kT, $k\in\mathbf{Z}$), также является ее периодом.*

□ Докажем, например, что $2T$ — период функции $f(x)$. Имеем

$$f(x)=f(x+T)=f((x+T)+T)=f(x+2T),$$
$$f(x)=f(x-T)=f((x-T)-T)=f(x-2T).$$

Аналогично доказывается, что $f(x)=f(x+3T)=f(x-3T)$ и вообще $f(x-kT)=$ $=f(x)=f(x+kT)$ для любого $k\in\mathbf{Z}$. Значит, все числа вида $kT(k\in\mathbf{Z})$ — периоды функции. ■

Таким образом, периодическая функция имеет бесконечное множество различных периодов. Если среди положительных периодов периодической

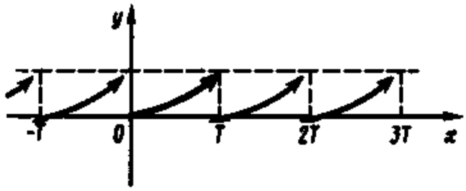


Рис. 25

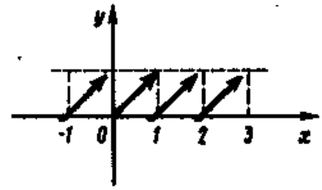


Рис. 26

функции существует наименьший, то его называют *основным периодом* этой функции; все остальные ее периоды кратны основному периоду.

Периодическая функция не всегда имеет основной период. Например, функция Дирихле (см. п. 1) периодическая: любое рациональное число является ее периодом. Однако среди положительных рациональных чисел нет наименьшего.

Графики периодических функций обладают следующей особенностью: если T — основной период функции $y=f(x)$, то для построения ее графика достаточно построить ветвь графика на одном из промежутков длины T, а затем произвести параллельный перенос этой ветви вдоль оси x на $\pm T$, $\pm 2T$, $\pm 3T$, ... (рис. 25). Чаще всего в качестве такого промежутка длины T выбирают промежуток с концами в точках $(-T/2;\ 0)$ и $(T/2;\ 0)$ или $(0;\ 0)$ и $(T;\ 0)$.

Пример. Доказать, что функция $y=\{x\}$ ($\{x\}$ — дробная часть числа x; см. п. 1) является периодической, и построить ее график.

○ Очевидно, что числа x и $x+k$ имеют одинаковую дробную часть, т. е. $\{x\}=\{x+k\}$. Значит, любое целое число k является периодом функции $\{x\}$, а основной период $T=1$.

Сначала построим график функции $y=\{x\}$ на промежутке $[0,\ 1]$, длина которого равна основному периоду. Если $0\leqslant x<1$, то $\{x\}=x$, т. е. на этом промежутке имеем $y=x$. График функции $y=\{x\}$ изображен на рис. 26. ●

4. Монотонные функции

Определение 5. Функция $y=f(x)$ называется *возрастающей* на промежутке X, если для любых двух точек x_1, x_2 этого промежутка таких, что $x_1<x_2$, выполняется неравенство $f(x_1)<f(x_2)$ (иными словами, если большему значению аргумента соответствует большее значение функции). Функция $y=f(x)$ называется *убывающей* на промежутке X, если для любых двух точек x_1, x_2 этого промежутка таких, что $x_1<x_2$, выполняется неравенство $f(x_1)>f(x_2)$ (иными словами, если большему значению аргумента соответствует меньшее значение функции).

При движении вдоль оси абсцисс слева направо ордината графика возрастающей функции увеличивается (рис. 27, *а*), а ордината графика убывающей функции уменьшается (рис. 27, *б*).

Возрастающие и убывающие функции объединяются термином *монотонные функции*.

Геометрические изображения позволяют убедиться в справедливости следующих утверждений:

1°. *Если функция $y=f(x)$ четна и возрастает при $x>0$, то она убывает при $x<0$* (рис. 28).

2°. *Если функция $y=f(x)$ четна и убывает при $x>0$, то она возрастает при $x<0$* (рис. 29).

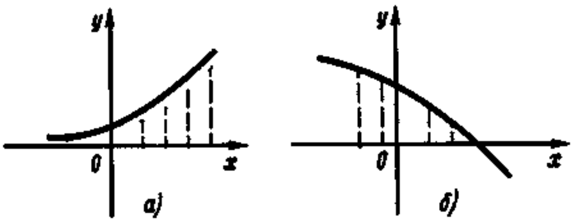


Рис. 27

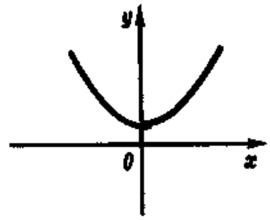


Рис. 28

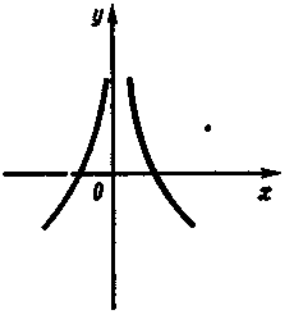


Рис. 29

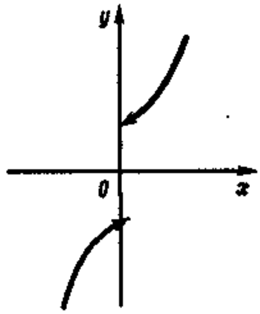


Рис. 30

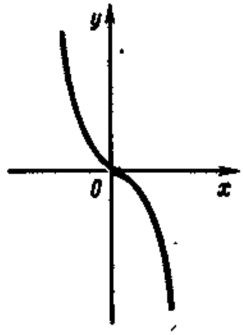


Рис. 31

3^0. *Если функция $y=f(x)$ нечетна и возрастает при $x>0$, то она возрастает и при $x<0$* (рис. 30).

4^0. *Если функция $y=f(x)$ нечетна и убывает при $x>0$, то она убывает и при $x<0$* (рис. 31).

5^0. *Монотонная функция не может быть периодической.*

Примеры. 1. Исследовать на монотонность функцию $y=x^3$.

○ Эта функция является нечетной; значит, достаточно провести исследование для промежутка $[0, +\infty]$. Если $0\leqslant x_1<x_2$, то $x_1^3<x_2^3$. Значит, из $x_1<x_2$ следует $f(x_1)<f(x_2)$, т. е. функция возрастает при $x\geqslant 0$. Тогда на основании свойства 3^0 она возрастает и при $x<0$. Наконец, так как для $x_1<0$ и $x_2>0$ очевидно, что $f(x_1)<f(x_2)$, то можно сделать вывод о том, что функция возрастает не только при $x<0$ и при $x>0$, но и на всей числовой прямой. ●

2. Исследовать на монотонность функцию $y=x^2+2$.

○ Данная функция четная. Если $0<x_1<x_2$, то используя свойства числовых неравенств, получим $x_1^2+2<x_2^2+2$, т. е. $f(x_1)<f(x_2)$. Значит, функция возрастает при $x>0$. Тогда, согласно свойству 1^0, она убывает при $x<0$. ●

5. Обратимые функции. Понятие обратной функции

Сравним две функции $y=f(x)$ и $y=g(x)$, графики которых изображены на рис. 32. Обе они определены на отрезке $[a, b]$, а множество их значений — отрезок $[c, d]$. Функция $y=f(x)$ обладает следующим свойством: какое бы число y_0 из множества значений функции ни взять, оно является значением функции в одной точке x_0: $y_0=f(x_0)$. Функция $y=g(x)$ этим свойством не обладает: так, для выбранного на рис. 32 значения y_0 имеем $y_0=g(x_1)$, $y_0=g(x_2)$ и $y_0=g(x_3)$. Иными словами, среди значений функции $y=g(x)$ имеются такие, которые функция принимает более чем в одной точке области

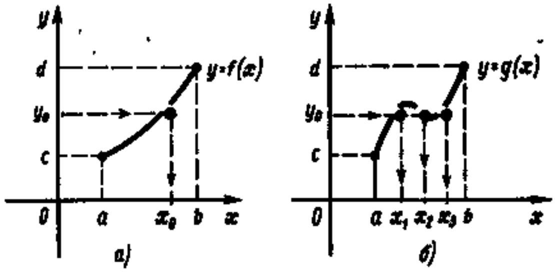


Рис. 32

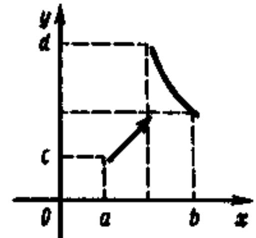


Рис. 33

определения. Говорят, что функция $y = f(x)$ обратима, а функция $y = g(x)$ необратима.

Определение 6. Функция $y = f(x)$, определенная на промежутке X, называется *обратимой*, если любое свое значение она принимает только в одной точке промежутка X; иными словами, если любым различным значениям аргумента соответствуют разные значения функции.

Теорема 4. *Если функция $y = f(x)$ монотонна на промежутке X, то она обратима.*

□ Пусть $y = f(x)$ возрастает на X. Тогда из $x_1 < x_2$ следует $f(x_1) < f(x_2)$. Таким образом, различным значениям аргумента соответствуют разные значения функции, т. е. функция обратима. ■

Наглядную иллюстрацию этой теоремы дает рис. 32: функция $y = f(x)$ монотонна и обратима, тогда как функция $y = g(x)$ немонотонна и необратима.

Определение 7. Пусть обратимая функция $y = f(x)$ определена на промежутке X, а область ее значений есть Y. Поставим в соответствие каждому y из Y то единственное значение x, при котором $f(x) = y$ (т. е. единственный корень уравнения $f(x) = y$ относительно переменной x). Тогда получим функцию, которая определена на Y, а область ее значений есть X. Эта функция обозначается $x = f^{-1}(y)$ и называется *обратной* по отношению к функции $y = f(x)$.

Из теоремы 4 следует, что для любой монотонной на X функции $y = f(x)$ существует обратная функция. Чтобы найти ее, нужно из уравнения $y = f(x)$ выразить x через y.

Теорема 5. *Если функция $y = f(x)$ возрастает (убывает) на промежутке X, а Y — область значений функции, то обратная функция $x = f^{-1}(y)$ также возрастает (убывает) на Y.*

□ Пусть $y = f(x)$ — возрастающая функция, y_1 и y_2 — два ее значения, причем $y_1 < y_2$. Так как функция обратима, то каждое из этих значений достигается в одной точке: $y_1 = f(x_1)$ и $y_2 = f(x_2)$. Значения x_1 и x_2 связаны неравенством $x_1 < x_2$. В самом деле, если предположить, что $x_1 \geqslant x_2$, то из возрастания функции $y = f(x)$ следовало бы $f(x_1) \geqslant f(x_2)$, т. е. $y_1 \geqslant y_2$, что противоречит условию. Таким образом, из $y_1 < y_2$ следует $x_1 < x_2$, т. е. $f^{-1}(y_1) < f^{-1}(y_2)$, а это и означает, что обратная функция $x = f^{-1}(y)$ возрастает на Y. ■

Монотонность функции является достаточным, но не необходимым условием существования обратной функции. Так, на рис. 33 изображен график немонотонной, но обратимой функции.

Переход от функции $y=f(x)$, $x\in X$ к обратной функции $x=f^{-1}(y)$, $y\in Y$ сводится лишь к изменению ролей множеств X и Y: в первом случае X отображается на Y, во втором — Y на X. Графики же функций $y=f(x)$ и $x=f^{-1}(y)$ представляют собой одно и то же множество точек координатной плоскости xy.

Обычно для обратной функции аргумент обозначают через x, а значение функции — через y, т. е. вместо $x=f^{-1}(y)$ пишут $y=f^{-1}(x)$. Таким образом, если пара чисел $(x;\ y)$ удовлетворяет уравнению $y=f(x)$ или эквивалентному уравнению $x=f^{-1}(y)$, то уравнению $y=f^{-1}(x)$ удовлетворяет пара чисел $(y;\ x)$.

Поэтому график функции $y=f^{-1}(x)$ получается из графика функции $y=f(x)$ с помощью преобразования плоскости xy, переводящего точку $(x;\ y)$

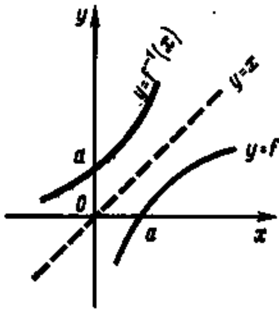


Рис. 34

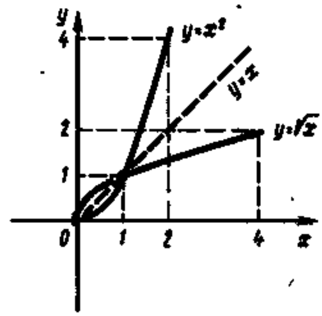


Рис. 35

в точку $(y;\ x)$. Этим преобразованием является симметрия относительно прямой $y=x$. Таким образом, *чтобы получить график функции $y=f^{-1}(x)$, обратной по отношению к функции $y=f(x)$, нужно график функции $y=f(x)$ преобразовать симметрично относительно прямой $y=x$* (рис. 34).

Пример 1. Дана функция $y=x^2$, $x\in[0,\ +\infty)$. Доказать, что для нее существует обратная функция, записать аналитическое выражение обратной функции в виде $y=f^{-1}(x)$ и построить график обратной функции.

○ Заданная функция возрастает на промежутке $[0,\ +\infty)$; значит, она имеет обратную функцию. Из уравнения $y=x^2$ находим $x=\sqrt{y}$ или $x=-\sqrt{y}$. Промежутку $[0,\ +\infty)$ принадлежат лишь значения функции $x=\sqrt{y}$. Это и есть обратная функция, которая определена на промежутке $[0,\ +\infty)$.

Поменяв местами x и y, получим $y=\sqrt{x}$, $x\in[0,\ +\infty)$. График этой функции получается из графика функции $y=x^2$, $x\in[0,\ +\infty)$ с помощью преобразования симметрии относительно прямой $y=x$ (рис. 35). ●

АЛГОРИТМ
составления обратной функции для функции $y=f(x)$, $x\in X$

1°. Убедиться в том, что функция $y=f(x)$ обратима на X.

2°. Из уравнения $y=f(x)$ выразить x через y, учтя при этом, что $x\in X$.

3°. В полученном равенстве поменять местами x и y.

Пример 2. Найти обратную функцию для функции $y=x^3+5$.

○ 1°. Эта функция возрастает на $(-\infty,\ \infty)$ и, значит, обратима.

2°. Из уравнения $y=x^3+5$ находим $x=\sqrt[3]{y-5}$.

3°. Поменяв x и y местами, получим $y=\sqrt[3]{x-5}$ — искомую обратную функцию. ●

6. Ограниченные функции

Определение 8. Функция $y = f(x)$ называется *ограниченной*, если область ее значений является ограниченным множеством. Иными словами, функция $y = f(x)$, $x \in X$ ограничена, если существует число $r > 0$ такое, что $|f(x)| < r$ для всех $x \in X$.

Геометрически это означает, что *график функции* $y = f(x)$ *целиком лежит внутри полосы, ограниченной прямыми* $y = -r$, $y = r$ (рис. 36). Например, ограниченной является функция $y = \frac{1}{1+x^2}$ (рис. 37).

Иногда понятие ограниченности рассматривают в несколько более узком смысле.

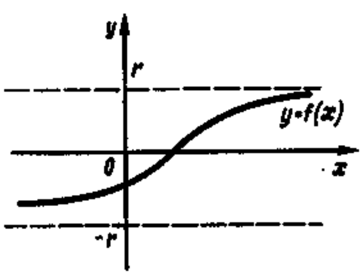


Рис. 36

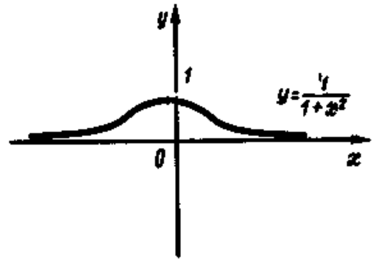


Рис. 37

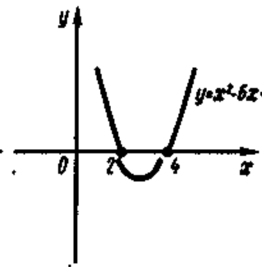


Рис. 38

Определение 9. Функция $y = f(x)$, $x \in X$ называется *ограниченной на промежутке* $X_1 \in X$, если множество значений функции на промежутке X_1 ограничено.

Например, функция $y = x^2 - 6x + 8$ (ее график изображен на рис. 38) ограничена на отрезке $[2, 4]$. Во всей же области определения эта функция не ограничена (она ограничена снизу, но не ограничена сверху).

§ 15. Предел функции на бесконечности

Понятие предела функции на бесконечности является в определенном смысле обобщением понятия предела последовательности, поэтому настоящий параграф мы построим по аналогии с § 13.

1. Бесконечно малые функции при $x \to +\infty$

Рассмотрим на луче $[1, +\infty]$ функции $f(x) = 1/x$ и $g(x) = -2/x^2$. Их графики обладают общей особенностью: неограниченно (в математике говорят «асимптотически») приближаются к положительному направлению оси x. Ось x называется горизонтальной асимптотой графика той и другой функции. Такие функции называют бесконечно малыми при $x \to +\infty$. Более строгое определение почти дословно повторяет определение бесконечно малой последовательности.

Определение 1. Функцию $\alpha(x)$, определенную на некотором луче $[Q, +\infty)$, называют *бесконечно малой при* $x \to +\infty$, если для любого $\varepsilon > 0$ существует число $M > 0$ такое, что при всех $x > M$ выполняется неравенство $|\alpha(x)| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)(\exists M > 0)(\forall x > M)\ |\alpha(x)| < \varepsilon.$$

Пример. Доказать, что $\alpha(x)=\frac{c}{ax+b}$ $(c\neq 0,\ a>0)$ — бесконечно малая функция при $x\to+\infty$.

○ Возьмем произвольное $\varepsilon>0$ и рассмотрим неравенство $\left|\frac{c}{ax+b}\right|<\varepsilon$. Рассуждая, как в примере 1 п. 1 § 13, придем к равносильному неравенству $x>$ $>\frac{1}{a}\left(\frac{|c|}{\varepsilon}-b\right)$. Полагая $M=\frac{1}{a}\left(\frac{|c|}{\varepsilon}-b\right)$, получим, что из $x>M$ следует $\left|\frac{c}{ax+b}\right|<\varepsilon$, т. е. $\frac{c}{ax+b}$ — бесконечно малая функция при $x\to+\infty$. ●

Геометрический смысл бесконечно малой при $x\to+\infty$ функции $y=$ $=\alpha(x)$ заключается в том, что график этой функции асимптотически приближается к положительному направлению оси x (рис. 39).

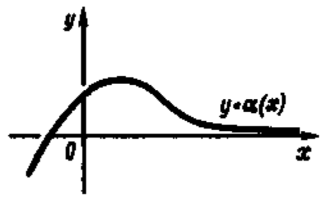


Рис. 39

Свойства бесконечно малых при $x\to+\infty$ функций аналогичны свойствам бесконечно малых последовательностей.

Теорема 1. *Постоянная функция $y=c$ является бесконечно малой при $x\to+\infty$ тогда и только тогда, когда $c=0$.*

Теорема 2. *Если $\beta(x)$ — бесконечно малая функция при $x\to+\infty$ и для всех x из некоторого луча $[Q,+\infty)$ выполняется неравенство $|\alpha(x)|\leqslant|\beta(x)|$, то и $\alpha(x)$ есть бесконечно малая функция при $x\to+\infty$.*

Теорема 3. *Если $\alpha(x)$ — бесконечно малая функция при $x\to+\infty$, то она является ограниченной на некотором луче $[M,+\infty)$.*

Эти теоремы мы приводим без доказательств, так как они легко следуют из определения 1.

Теорема 4. *Сумма двух бесконечно малых при $x\to+\infty$ функций также является бесконечно малой при $x\to+\infty$ функцией.*

□ Пусть $\alpha(x)$, $\beta(x)$ — бесконечно малые при $x\to+\infty$ функции. Возьмем произвольное $\varepsilon>0$. Тогда существуют числа $M_1>0$ и $M_2>0$ такие, что при $x>M_1$, выполняется неравенство $|\alpha(x)|<\varepsilon/2$, а при $x>M_2$ — неравенство $|\beta(x)|<\varepsilon/2$.

Из чисел M_1 и M_2 возьмем наибольшее, которое обозначим через M; тогда для всех $x>M$ одновременно выполняются неравенства $|\alpha(x)|<\varepsilon/2$ и $|\beta(x)|<\varepsilon/2$.

Рассмотрим функцию $\gamma(x) = \alpha(x) + \beta(x)$. Имеем $|\gamma(x)| = |\alpha(x) + \beta(x)| \leqslant \leqslant |\alpha(x)| + |\beta(x)|$. Поэтому при $x > M$ получим $|\gamma(x)| \leqslant |\alpha(x)| + |\beta(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$, т. е. $|\gamma(x)| < \varepsilon$, а это и означает, что $\gamma(x)$ — бесконечно малая функция при $x \to +\infty$. ■

Заметим, что проведенное доказательство почти дословно повторяет доказательство свойства 5^0 из § 13. То же относится и к следующей теореме (сравните ее со свойством 6^0 из § 13), поэтому приводим ее без доказательства.

Теорема 5. *Если $\alpha(x)$ — бесконечно малая при $x \to +\infty$ функция, а $y = f(x)$ — ограниченная функция на некотором луче $[a, +\infty)$, то их произведение является бесконечно малой при $x \to +\infty$ функцией.*

Следствие 1. *Если $\alpha(x)$ — беконечно малая при $x \to +\infty$ функция, то и $c\alpha(x)$, где c — любое действительное число, также является бесконечно малой при $x \to +\infty$ функцией.*

Следствие 2. *Произведение двух (и вообще любого конечного числа) бесконечно малых при $x \to +\infty$ функций есть бесконечно малая при $x \to +\infty$ функция.*

Следствие 3. Если $\alpha_1(x)$, $\alpha_2(x)$, ..., $\alpha_n(x)$ — *бесконечно малые при $x \to +\infty$ функции, то и $c_1\alpha_1(x) + c_2\alpha_2(x) + ... + c_n\alpha_n(x)$ ($c_1, c_2, ..., c_n$ — действительные числа) также является бесконечно малой при $x \to +\infty$ функцией.*

Например, функция $y = \frac{1}{x} + \frac{2}{x^2} - \frac{3}{x} + \frac{\sqrt{2}}{x^4}$ есть бесконечно малая при $x \to +\infty$.

2. Предел функции при $x \to +\infty$ и его свойства

Определение 2. Число b называют *пределом функции $f(x)$ при $x \to +\infty$*, если $f(x) - b$ есть бесконечно малая функция при $x \to +\infty$; пишут $\lim\limits_{x \to +\infty} f(x) = b$.

В частности, из этого определения следует, что

$$\lim_{x \to +\infty} c = c, \quad \lim_{x \to +\infty} \alpha(x) = 0,$$

где $\alpha(x)$ — бесконечно малая при $x \to +\infty$ функция.

Воспользовавшись для бесконечно малой функции $f(x) - b$ определением 1, получим другое определение предела, эквивалентное предыдущему.

Определение 3. Число b называют *пределом функции $f(x)$ при $x \to +\infty$*, если для любого $\varepsilon > 0$ существует $M > 0$ такое, что при всех $x > M$ выполняется неравенство $|f(x) - b| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)\ (\exists M > 0)\ (\forall x > M)\ |f(x) - b| < \varepsilon.$$

Определение 2 условимся называть определением предела «на языке бесконечно малых», а определение 3 — «на языке $\varepsilon - M$».

Пусть $\lim\limits_{x \to +\infty} f(x) = b$, т. е. $f(x) = b + \alpha(x)$, где $\alpha(x)$ — бесконечно малая при $x \to +\infty$ функция. График функции $b + \alpha(x)$ получается из графика функции $\alpha(x)$ сдвигом на $|b|$ вверх, если $b > 0$, и вниз, если $b < 0$. Так как график бесконечно малой функции $\alpha(x)$ имеет горизонтальную асимптоту (ось x), то для графика функции $b + \alpha(x)$ горизонтальной асимптотой служит прямая $y = b$ (рис. 40).

Итак, геометрический смысл равенства $\lim\limits_{x \to +\infty} f(x) = b$ состоит в следую-

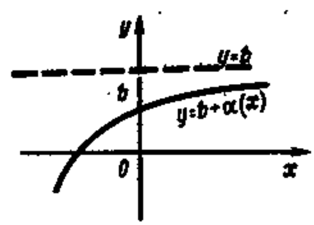


Рис. 40

щем: *прямая $y = b$ является горизонтальной асимптотой графика функции $y = f(x)$* (при $x \to +\infty$).

Рассмотрим некоторые свойства предела на бесконечности. В ряде случаев доказательства почти дословно повторяют доказательства аналогичных свойств предела последовательности и мы их не приводим.

Теорема 6. *Если функция имеет предел при $x \to +\infty$, то только один.*

Теорема 7. *Если функция имеет предел при $x \to +\infty$, то она ограничена на некотором открытом* луче $(M, +\infty)$.*

□ Пусть $\lim\limits_{x \to +\infty} f(x) = b$. Возьмем произвольное $\varepsilon > 0$; тогда существует такое $M > 0$, что при $x > M$ выполняется неравенство $|f(x) - b| < \varepsilon$. Иными словами, на луче $(M, +\infty)$ выполняется неравенство $b - \varepsilon < f(x) < b + \varepsilon$, а это и означает ограниченность функции на этом луче. ■

Теорема 8 (о предельном переходе в неравенствах). *Если $\lim\limits_{x \to +\infty} f(x) = b$, $\lim\limits_{x \to +\infty} g(x) = c$ и на некотором луче $(a, +\infty)$ выполняется неравенство $f(x) \leqslant g(x)$, то $b \leqslant c$.*

Следствие. *Если $f(x) \geqslant 0$ ($f(x) \leqslant 0$) на некотором луче $(a, +\infty)$ и $\lim\limits_{x \to +\infty} f(x)$ существует, то он неотрицателен (неположителен).*

Теорема 9 (о пределе промежуточной функции). *Если $\lim\limits_{x \to +\infty} f(x) = \lim\limits_{x \to +\infty} h(x) = b$ и на некотором луче $(0, +\infty)$ справедливо двойное неравенство $f(x) \leqslant g(x) \leqslant h(x)$, то и $\lim\limits_{x \to +\infty} g(x) = b$.*

Теорема 10. *Пусть $\lim\limits_{x \to +\infty} f(x) = b$, $\lim\limits_{x \to +\infty} g(x) = c$. Тогда:*

1) $\lim\limits_{x \to +\infty} (f(x) + g(x)) = b + c$; 2) $\lim\limits_{x \to +\infty} f(x)\, g(x) = bc$;

3) $\lim\limits_{x \to +\infty} \dfrac{f(x)}{g(x)} = \dfrac{b}{c}$ (*если* $c \neq 0$).

3. Предел функции при $x \to -\infty$, при $x \to \infty$

Наряду с функциями, бесконечно малыми при $x \to +\infty$, рассматривают и функции, бесконечно малые при $x \to -\infty$. Таковыми являются, например, функции $f(x) = 1/x$, $g(x) = -2/x^2$, рассмотренные на луче $(-\infty, 0)$. Ось x является горизонтальной асимптотой графика каждой из этих функций (см. левые ветви кривых на рис. 17, *а* и *б*).

---

* В дальнейшем, говоря об открытом луче, мы будем опускать термин «открытый», т. е. называть лучом и множество $[M, +\infty)$ и множество $(M, +\infty)$. По обозначению легко понять, где речь идет о луче, а где — об открытом луче.

Определение 4. Функцию $\alpha(x)$, определенную на некотором луче $(-\infty, a]$, называют *бесконечно малой при* $x \to -\infty$, если для любого $\varepsilon > 0$ существует число $M > 0$ такое, что при всех $x < -M$ выполняется неравенство $|\alpha(x)| < \varepsilon$.

Существуют функции, бесконечно малые как при $x \to +\infty$, так и при $x \to -\infty$. Таковы обе рассмотренные выше функции $1/x$ и $-2/x^2$ (ось x является горизонтальной асимптотой графика обеих функций как в положительном, так и в отрицательном направлении). Такова и функция $y = \frac{1}{1+x^2}$ (см. рис. 37).

Поскольку вместо двух неравенств $x > M$, $x < -M$ можно написать одно: $|x| > M$, получаем следующее определение.

Определение 5. Функцию $\alpha(x)$, определенную на объединении двух лучей $(-\infty, a_1] \cup [a_2, +\infty)$, называют *бесконечно малой при* $x \to \infty$, если для любого $\varepsilon > 0$ существует число $M > 0$ такое, что для всех x таких, что $|x| > M$, выполняется неравенство $|\alpha(x)| < \varepsilon$.

Для функций, бесконечно малых при $x \to -\infty$ или при $x \to \infty$, справедливы теоремы 1—5 (с соответствующими изменениями в формулировках и доказательствах).

Определение 6. Число b называют *пределом функции* $f(x)$ *при* $x \to \infty$ $(x \to -\infty)$, если $f(x) - b$ есть бесконечно малая функция при $x \to \infty (x \to -\infty)$; пишут $\lim\limits_{x \to \infty} f(x) = b$ $(\lim\limits_{x \to -\infty} f(x) = b)$.

В частности, из определения 6 следует, что $\lim\limits_{x \to \infty} c = c$, $\lim\limits_{x \to \infty} \alpha(x) = 0$, где $\alpha(x)$ — бесконечно малая функция при $x \to \infty$.

Определение 6 «на языке бесконечно малых» можно дать в эквивалентной форме «на языке $\varepsilon - M$»:

Определение 7. Число b называют *пределом функции* $f(x)$ *при* $x \to \infty$, если для любого $\varepsilon > 0$ существует $M > 0$ такое, что при всех x таких, что $|x| > M$, выполняется неравенство $|f(x) - b| < \varepsilon$.

Геометрический смысл равенства $\lim\limits_{x \to \infty} f(x) = b$ состоит в следующем: прямая $y = b$ является горизонтальной асимптотой графика функции $y = f(x)$ (и в положительном, и в отрицательном направлении).

Для предела функции при $x \to \infty$ справедливы теоремы 6—10 (с соответствующими изменениями в формулировках и доказательствах).

Примеры. 1. Вычислить $\lim\limits_{x \to \infty} \frac{3x^2 + x + 1}{2x^3 - x^2 - 3}$.

○ Находим

$$\lim_{x \to \infty} \frac{\frac{3x^2}{x^3} + \frac{x}{x^3} + \frac{1}{x^3}}{\frac{2x^3}{x^3} - \frac{x^2}{x^3} - \frac{3}{x^3}} = \lim_{x \to \infty} \frac{3 + \frac{1}{x^2} + \frac{1}{x^3}}{2 - \frac{1}{x} - \frac{3}{x^3}} =$$

$$= \frac{\lim\limits_{x \to \infty}\left(3 + \frac{1}{x} \cdot \frac{1}{x} + \frac{1}{x} \cdot \frac{1}{x} \cdot \frac{1}{x}\right)}{\lim\limits_{x \to \infty}\left(2 - \frac{1}{x} - 3 \cdot \frac{1}{x} \cdot \frac{1}{x} \cdot \frac{1}{x}\right)} = \frac{3 + 0 \cdot 0 + 0 \cdot 0 \cdot 0}{2 - 0 - 3 \cdot 0 \cdot 0 \cdot 0} = \frac{3}{2}.$$

Мы воспользовались теоремой 10 об арифметических операциях над пределами и тем, что функция $1/x$ является бесконечно малой при $x \to \infty$. ●

2. Вычислить $\lim\limits_{x \to \infty} \frac{x + 2}{x^4 - 3x - 1}$.

○ Имеем

$$\lim_{x\to\infty}\frac{\frac{x}{x^4}+\frac{2}{x^4}}{\frac{x^4}{x^4}-\frac{3x}{x^4}-\frac{1}{x^4}}=\lim_{x\to\infty}\frac{\frac{1}{x^3}+\frac{2}{x^4}}{1-\frac{3}{x^3}-\frac{1}{x^4}}=\frac{0+0}{1-0-0}=0.\ \bullet$$

В § 13 мы доказали, что $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n=e$. Аналогичное равенство имеет место и для предела функции на бесконечности:

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

(доказательство мы не приводим).

4. Бесконечно большие функции при $x\to\infty$

Выше (см. п. 7 § 13) мы ввели понятие бесконечно большой последовательности и рассмотрели ее свойства. Определение и свойства бесконечно большой при $x\to\infty$ функции аналогичны.

Определение 8. Функция $y=f(x)$, определенная на объединении двух лучей $(-\infty, a_1)\cup(a_2, +\infty)$, называется *бесконечно большой при* $x\to\infty$, если для любого $P>0$ существует число $M>0$ такое, что при всех x таких, что $|x|>M$, выполняется неравенство $|f(x)|>P$.

Для бесконечно большой функции используется запись $\lim\limits_{x\to\infty}f(x)=\infty$.

Так, функции $y=ax^3$ и $y=x^2-6x+8$ — бесконечно большие при $x\to\infty$ (см. соответственно рис. 20 и 38); функция, график которой изображен на рис. 27, *а*, является бесконечно большой при $x\to+\infty$ и бесконечно малой при $x\to-\infty$; функция, график которой изображен на рис. 40, есть бесконечно большая при $x\to-\infty$ и, кроме того, для нее $\lim\limits_{x\to+\infty}f(x)=b$.

Теорема 11. *Для того чтобы функция $f(x)$ была бесконечно большой при $x\to\infty$, необходимо и достаточно, чтобы функция $1/f(x)$ была бесконечно малой при $x\to\infty$.*

Теорема 12. *Если $f(x)$ — бесконечно большая функция при $x\to\infty$, а функция $g(x)$ такова, что $\lim\limits_{x\to\infty}g(x)=b\neq0$, то $f(x)g(x)$ является бесконечно большой функцией при $x\to\infty$.*

Пример. Вычислить $\lim\limits_{x\to\infty}\frac{x^4-3x-1}{x+2}$.

○ Выше (см. п. 3) мы доказали, что $\frac{x+2}{x^4-3x-1}$ — бесконечно малая функция при $x\to\infty$. Значит, согласно теореме 11, $\frac{x^4-3x-1}{x+2}$ является бесконечно большой функцией при $x\to\infty$, т. е. $\lim\limits_{x\to\infty}\frac{x^4-3x-1}{x+2}=\infty$. ●

5. Предел отношения двух многочленов при $x\to\infty$

Выше мы вычисляли предел отношения двух многочленов при $x\to\infty$. В примере 1 п. 3 степени числителя и знаменателя равны, а предел дроби оказался равным отношению старших коэффициентов числителя и знаменателя. В примере 2 п. 3 степень числителя меньше степени знаменателя, а предел дроби оказался равным нулю. Наконец, в примере п. 4 степень числителя больше степени знаменателя, а дробь оказалась бесконечно большой. Аналогичный результат (он получается с помощью тех же рассуждений, что были проведены в указанных примерах) справедлив для отношения любых многочленов. Например, не производя детальных выкладок, можно утверждать, что

$$\lim_{x\to\infty}\frac{5x^4+x-1}{3x^4+x^3-x^2-1000}=\frac{5}{3},$$

$$\lim_{x\to\infty}\frac{2x^2-x+5}{x^7-x^5+4}=0,\quad \lim_{x\to\infty}\frac{3x^2-x+5}{200x+41}=\infty.$$

6. Асимптоты

Определение 9. Прямая $y=kx+b$ называется *асимптотой* графика функции $y=f(x)$ при $x\to\infty$, если $\lim\limits_{x\to\infty}(f(x)-(kx+b))=0$; иными словами, если отклонение графика функции $f(x)$ от прямой $y=kx+b$ неограниченно уменьшается при $x\to\infty$ (рис. 41).

Аналогично можно определить асимптоту при $x\to+\infty$ или при $x\to-\infty$.

Мы знаем (см. п. 3), что если $\lim\limits_{x\to\infty}f(x)=b$, то прямая $y=b$ является асимптотой (горизонтальной) графика функции $y=f(x)$ (что находится в полном соответствии с определением 9). Верно и обратное: если $y=b$ — горизонтальная асимптота графика функции $y=f(x)$, то $\lim f(x)=b$.

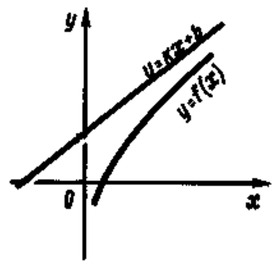


Рис. 41

В том случае, когда $\lim\limits_{x\to\infty}f(x)=\infty$, график функции $y=f(x)$ не имеет горизонтальной асимптоты, но может иметь наклонную асимптоту $y=kx+b$, где $k\neq 0$.

Так как при этом $\lim\limits_{x\to\infty}(f(x)-kx-b)=0$, то тем более $\lim\limits_{x\to\infty}\frac{f(x)-kx-b}{x}=0$ или $\lim\limits_{x\to\infty}\left(\frac{f(x)}{x}-k-\frac{b}{x}\right)=0$. Поэтому

$$\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\left(k+\frac{b}{x}\right)=k.$$

Далее, из равенства $\lim\limits_{x\to\infty}(f(x)-kx-b)=0$ следует, что $\lim\limits_{x\to\infty}(f(x)-kx)=b$. Значит, числа k и b можно найти из равенств

$$k=\lim_{x\to\infty}\frac{f(x)}{x},\quad b=\lim_{x\to\infty}(f(x)-kx).$$

АЛГОРИТМ

отыскания асимптоты графика функции $y = f(x)$.

1°. Вычислить $\lim\limits_{x\to\infty} f(x)$. Если этот предел существует и равен b, то $y = b$ — горизонтальная асимптота; если $\lim\limits_{x\to\infty} f(x) = \infty$, то перейти ко второму шагу.

2°. Вычислить $\lim\limits_{x\to\infty} \frac{f(x)}{x}$. Если этот предел не существует, то асимптоты нет; если он существует и равен $\lim\limits_{x\to\infty} \frac{f(x)}{x} = k$, то перейти к третьему шагу.

3°. Вычислить $\lim\limits_{x\to\infty} (f(x) - kx)$. Если этот предел не существует, то асимптоты нет; если он существует и равен b, то перейти к четвертому шагу.

4°. Записать уравнение наклонной асимптоты в виде $y = kx + b$.

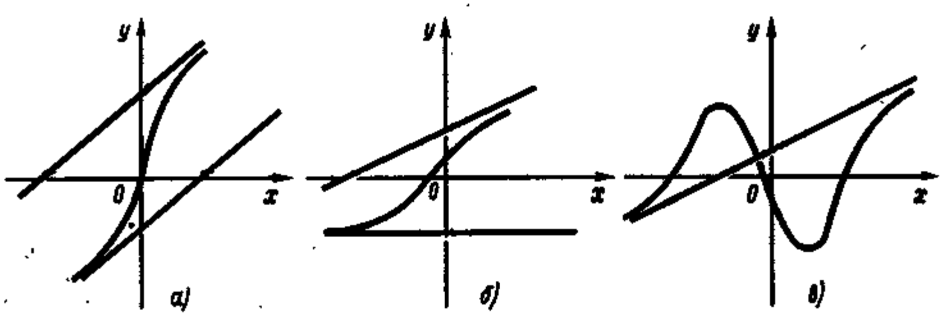


Рис. 42

Пример. Найти асимптоту графика функции $y = \frac{x^3 - 6x^2 + 3}{2x^2 + 5}$.

○ 1°. $\lim\limits_{x\to\infty} \frac{x^3 - 6x^2 + 3}{2x^2 + 5} = \infty$ (см. п. 5).

2°. $\lim\limits_{x\to\infty} \frac{x^3 - 6x^2 + 3}{x(2x^2 + 5)} = \lim\limits_{x\to\infty} \frac{x^3 - 6x^2 + 3}{2x^3 + 5x} = \frac{1}{2}$ (см. п. 5); значит, $k = \frac{1}{2}$.

3°. $\lim\limits_{x\to\infty} \left(\frac{x^3 - 6x^2 + 3}{2x^2 + 5} - \frac{1}{2}x \right) = \lim\limits_{x\to\infty} \frac{-12x^2 - 5x + 6}{4x^2 + 10} = \frac{-12}{4} = -3$, т. е. $b = -3$.

4°. Уравнение асимптоты имеет вид $y = \frac{1}{2}x - 3$. ●

Заметим, что на практике могут встретиться различные случаи: например, на рис. 41 изображен график функции, имеющий наклонную асимптоту только при $x \to +\infty$; на рис. 42, *а* — две наклонные асимптоты, различные при $x \to +\infty$ и $x \to -\infty$; на рис. 42, *б* — горизонтальную асимптоту при $x \to -\infty$ и наклонную асимптоту при $x \to +\infty$; на рис. 42, *в* — наклонную асимптоту при $x \to \infty$.

§ 16. Предел функции в точке

1. Бесконечно малые функции при $x \to a$

В § 15 мы рассматривали бесконечно малые функции при $x \to \infty$. Каждая такая функция обладает следующим свойством: значения функции по модулю могут стать меньше наперед заданного числа $\varepsilon > 0$, если значения аргумента выбирать достаточно большими по модулю. Имеются функции, обладающие аналогичным свойством, но не при $x \to \infty$, а при $x \to a$: чем ближе к a выбирать значения аргумента, тем меньше (по модулю) становятся значения функции, причем они могут стать меньше любого наперед заданного числа $\varepsilon > 0$. Графики нескольких таких функций (бесконечно малых при $x \to a$) изображены на рис. 43, *а — в*.

Прежде чем сформулировать строгое определение, заметим, что условие $x \to \infty$ описывается неравенством $|x| > M$; условие же $x \to a$ описывается неравенством $|x - a| < \varepsilon$. Кроме того (это отчетливо видно на рис 43), когда

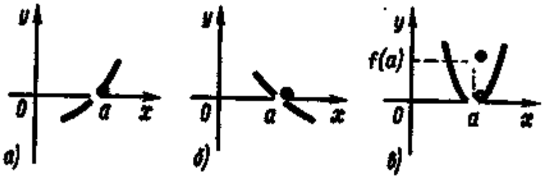


Рис. 43

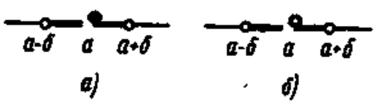


Рис. 44

изучают поведение функции при $x \to a$, саму точку $x = a$ не принимают во внимание, т. е. считают, что $x \neq a$, и потому $|x - a| > 0$.

Определение 1. Функцию $\alpha(x)$ называют *бесконечно малой при* $x \to a$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при всех x из неравенства $0 < |x - a| < \delta$ следует неравенство $|\alpha(x)| < \delta$. Короче:

$$(\forall \varepsilon > 0)\,(\exists \delta > 0)\,(\forall x: 0 < |x - a| < \delta)\ |\alpha(x)| < \varepsilon.$$

Простейшими бесконечно малыми при $x \to a$ являются функции $x - a$, $(x - a)^2$, $(x - a)^3$ и т. д.

Геометрически неравенство $|x - a| < \delta$ описывает δ-окрестность точки a (рис. 44, *а*). Неравенство $0 < |x - a| < \delta$ описывает ту же окрестность, из которой выброшена точка a (рис. 44, *б*); такую окрестность иногда называют *проколотой*. Тогда определение 1 можно сформулировать так:

Определение 2. Функцию $\alpha(x)$ называют *бесконечно малой при* $x \to a$, если для любого $\varepsilon > 0$ можно указать проколотую δ-окрестность точки a, во всех точках x которой выполняется неравенство $|\alpha(x)| < \varepsilon$.

Свойства бесконечно малых функций при $x \to a$ аналогичны свойствам бесконечно малых функций при $x \to \infty$ (см. § 15).

Теорема 1. *Постоянная функция $y = c$ является бесконечно малой при $x \to a$ тогда и только тогда, когда $c = 0$.*

Теорема 2. *Если $\beta(x)$ — бесконечно малая функция при $x \to a$ и для всех x из некоторой проколотой окрестности точки a выполняется неравенство $|\alpha(x)| \leqslant |\beta(x)|$, то и $\alpha(x)$ — бесконечно малая функция при $x \to a$.*

Теорема 3. *Если $\alpha(x)$ — бесконечно малая функция при $x \to a$, то она является ограниченной в некоторой проколотой окрестности точки a.*

Эти теоремы мы приводим без доказательства, поскольку они легко вытекают из определений 1 или 2.

Теорема 4. *Сумма двух бесконечно малых при $x \to a$ функций является бесконечно малой при $x \to a$ функцией.*

□ Пусть $\alpha(x)$, $\beta(x)$ — бесконечно малые функции при $x \to a$. Возьмем произвольное $\varepsilon > 0$. Тогда существуют числа $\delta_1 > 0$ и $\delta_2 > 0$ такие, что в проколотой δ_1-окрестности точки a выполняется неравенство $|\alpha(x)| < \varepsilon/2$, а в проколотой δ_2-окрестности точки a — неравенство $|\beta(x)| < \varepsilon/2$.

Из чисел δ_1 и δ_2 возьмем наименьшее, которое обозначим через δ; тогда в проколотой δ-окрестности точки a выполняются оба неравенства $|\alpha(x)| < \varepsilon/2$ и $|\beta(x)| < \varepsilon/2$.

Рассмотрим функцию $\gamma(x) = \alpha(x) + \beta(x)$. Имеем $|\gamma(x)| = |\alpha(x) + \beta(x)| \leqslant \leqslant |\alpha(x)| + |\beta(x)|$. Следовательно, в проколотой δ-окрестности точки a получим $|\gamma(x)| \leqslant |\alpha(x)| + |\beta(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$, т. е. $|\gamma(x)| < \varepsilon$, а это и означает, что $\gamma(x)$ — бесконечно малая функция при $x \to a$. ■

Заметим, что проведенное доказательство почти дословно повторяет доказательство теоремы 4 § 15, непринципиальное отличие было лишь во втором шаге. То же относится и к следующей теореме (ср. ее с теоремой 5 § 15), поэтому мы приводим ее без доказательства.

Теорема 5. *Если $\alpha(x)$ — бесконечно малая функция при $x \to a$, а $y = f(x)$ — функция, ограниченная в некоторой проколотой окрестности точки a, то их произведение является бесконечно малой функцией при $x \to a$.*

Следствие 1. *Если $\alpha(x)$ — бесконечно малая функция при $x \to a$, то и $c\alpha(x)$ (c — действительное число) является бесконечно малой функцией при $x \to a$.*

Следствие 2. *Произведение двух (и вообще любого конечного числа) бесконечно малых при $x \to a$ функций есть бесконечно малая при $x \to a$ функция.*

Следствие 3. *Если $\alpha_1(x)$, $\alpha_2(x)$, ..., $\alpha_n(x)$ — бесконечно малые функции при $x \to a$, то и $c_1\alpha_1(x) + c_2\alpha_2(x) + \ldots + c_n\alpha_n(x)$ (c_1, c_2, ..., c_n — действительные числа) есть бесконечно малая функция при $x \to a$.*

Например, $y = x$ — бесконечно малая функция при $x \to 0$. Тогда бесконечно малыми при $x \to 0$ являются такие функции: $2x$, $0{,}5x$, x^2, x^3, $3x^2$, $x + x^2 + x^3$, $2x^4 - 3x^3 + 0{,}5x^2 - x$.

2. Предел функции в точке и его свойства

Определение 3. Число b называют *пределом функции $f(x)$ при $x \to a$*, если $f(x) - b$ является бесконечно малой функцией при $x \to a$; пишут $\lim\limits_{x \to a} f(x) = b$.

В частности, из этого определения следует, что: $\lim\limits_{x \to a} c = c$; $\lim\limits_{x \to a} \alpha(x) = 0$, где $\alpha(x)$ — бесконечно малая функция при $x \to a$; $\lim\limits_{x \to a} x = a$.

Воспользовавшись для бесконечно малой функции $f(x) - b$ определениями 1 и 2, получим еще два определения, эквивалентные предыдущему.

Определение 4. Число b называют *пределом функции* $f(x)$ *при* $x \to a$, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что из неравенства $0 < |x-a| < \delta$ следует неравенство $|f(x)-b| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x : 0 < |x-a| < \delta)|f(x)-b| < \varepsilon.$$

Определение 5. Число b называют *пределом функции* $f(x)$ *при* $x \to a$, если для любого $\varepsilon > 0$ можно указать такую проколотую δ-окрестность точки a, в которой выполняется неравенство $|f(x)-b| < \varepsilon$.

Определение 3 условимся называть определением предела функции в точке «на языке бесконечно малых», определение 4 — «на языке $\varepsilon - \delta$», определение 5 — «на языке окрестностей».

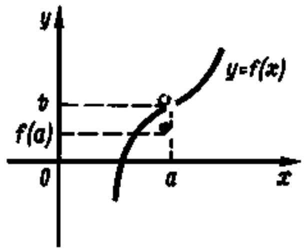


Рис. 45

Геометрический смысл равенства $\lim\limits_{x \to a} f(x) = b$ иллюстрирует рис. 45.

Перечислим свойства предела функции в точке. Их доказательств мы не приводим, поскольку они по структуре и идеям практически те же, что доказательства теорем 6—10 из § 15; их легко воспроизвести самостоятельно, ориентируясь, например, на доказательство теоремы 4.

Теорема 6. *Если функция имеет предел при $x \to a$, то только один.*

Теорема 7. *Если функция имеет предел при $x \to a$, то она ограничена в некоторой проколотой окрестности точки a.*

Теорема 8 (о предельном переходе в неравенствах). *Если $\lim\limits_{x \to a} f(x) = b$, $\lim\limits_{x \to a} g(x) = c$ и в некоторой проколотой окрестности точки a выполняется неравенство $f(x) \leqslant g(x)$, то $b \leqslant c$.*

Следствие. *Если $f(x) \geqslant 0$ ($f(x) \leqslant 0$) в некоторой проколотой окрестности точки a и $\lim\limits_{x \to a} f(x)$ существует, то он неотрицателен (неположителен).*

Теорема 9 (о пределе промежуточной функции). *Если $\lim\limits_{x \to a} f(x) = \lim\limits_{x \to a} h(x) = b$ и в некоторой проколотой окрестности точки a справедливо неравенство $f(x) \leqslant g(x) \leqslant h(x)$, то и $\lim\limits_{x \to a} g(x) = b$.*

Теорема 10. *Пусть $\lim\limits_{x \to a} f(x) = b$, $\lim\limits_{x \to a} g(x) = c$. Тогда: 1) $\lim\limits_{x \to a} (f(x) + g(x)) = b + c$; 2) $\lim\limits_{x \to a} f(x)g(x) = bc$; 3) $\lim\limits_{x \to a} \frac{f(x)}{g(x)} = \frac{b}{c}$ (если $c \neq 0$).*

Примеры. 1. Вычислить $\lim\limits_{x\to 2}\frac{x^2+4}{3x+5}$.

○ Используя теорему 10 об арифметических операциях над пределами и учитывая, что $\lim\limits_{x\to 2}x=2$ и $\lim\limits_{x\to 2}c=c$, получаем

$$\lim_{x\to 2}\frac{x^2+4}{3x+5}=\frac{\lim\limits_{x\to 2}(x^2+4)}{\lim\limits_{x\to 2}(3x+5)}=\frac{\lim\limits_{x\to 2}x\cdot x+\lim\limits_{x\to 2}4}{\lim\limits_{x\to 2}3x+\lim\limits_{x\to 2}5x}=\frac{2\cdot 2+4}{3\cdot 2+5}=\frac{8}{11}.$$

Короче:

$$\lim_{x\to 2}\frac{x^2+4}{3x+5}=\frac{2^2+4}{3\cdot 2+5}=\frac{8}{11}. \bullet$$

2. Вычислить $\lim\limits_{x\to 3}\frac{x^2-5x+6}{x^2-9}$.

○ Здесь нельзя применить теорему 10 о пределе частного, так как $\lim\limits_{x\to 3}(x^2-9)=0$. Выполним некоторые преобразования:

$$\frac{x^2-5x+6}{x^2-9}=\frac{(x-3)(x-2)}{(x-3)(x+3)}=\frac{x-2}{x+3}.$$

Сокращение на $x-3$ возможно, поскольку, как мы отмечали выше, условие $x\to 3$ предполагает, что $x\neq 3$. Тогда получим

$$\lim_{x\to 3}\frac{x^2-5x+6}{x^2-9}=\lim_{x\to 3}\frac{x-2}{x+3}=\frac{3-2}{3+3}=\frac{1}{6}. \bullet$$

3. Односторонние пределы

Изучая в § 15 предел функции на бесконечности, мы отмечали, что понятие предела функции $f(x)$ при $x\to\infty$ фактически является объединением двух понятий: предела при $x\to+\infty$ и предела при $x\to-\infty$: если $\lim\limits_{x\to\infty}f(x)=b$, то и $\lim\limits_{x\to+\infty}f(x)=b$, и $\lim\limits_{x\to-\infty}f(x)=b$; верно и обратное: если $\lim\limits_{x\to+\infty}f(x)=b$ и $\lim\limits_{x\to-\infty}f(x)=b$, то и $\lim\limits_{x\to\infty}f(x)=b$. Аналогично, при изучении предела $\lim\limits_{x\to a}f(x)$ можно разложить это понятие на две составляющие части: $\lim f(x)$, когда $x\to a$, оставаясь все время меньше a, т. е. $x\to a$ слева, и $\lim f(x)$, когда $x\to a$, оставаясь все время больше a, т. е. $x\to a$ справа. В этом случае говорят об односторонних пределах, обозначая их соответственно $\lim\limits_{x\to a-0}f(x)$ или $f(a-0)$ (предел слева) и $\lim\limits_{x\to a+0}f(x)$ или $f(a+0)$ (предел справа). На рис. 46, *а* дана геометрическая иллюстрация для $\lim\limits_{x\to a-0}f(x)=b$, на рис. 46, *б* — для $\lim\limits_{x\to a+0}f(x)=b$.

Строгие определения односторонних пределов мы получим, опираясь на определение предела функции в точке «на языке $\varepsilon-\delta$» (определение 4). Заметим лишь, что если $x\to a$ слева, т. е. $x<a$, то $|x-a|=a-x$, а если $x\to a$ справа, т. е. $x>a$, то $|x-a|=x-a$.

Определение 6. Число b называется *левосторонним пределом* функции $f(x)$ при $x\to a$, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что при всех x из неравенства $0<a-x<\delta$ следует неравенство $|f(x)-b|<\varepsilon$.

Число b называется *правосторонним пределом* функции $f(x)$ при $x\to a$, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что при всех x из неравенства $0<x-a<\delta$ следует неравенство $|f(x)-b|<\varepsilon$.

Достаточно очевидно следующее утверждение.

Теорема 11. *Для того чтобы существовал* $\lim\limits_{x\to a} f(x)$, *необходимо и достаточно, чтобы существовали односторонние пределы* $\lim\limits_{x\to a-0} f(x)$ *и* $\lim\limits_{x\to a+0} f(x)$ *и были равны между собой:* $\lim\limits_{x\to a-0} f(x) = \lim\limits_{x\to a+0} f(x)$.

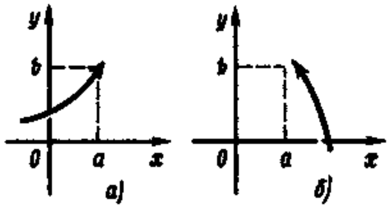


Рис. 46

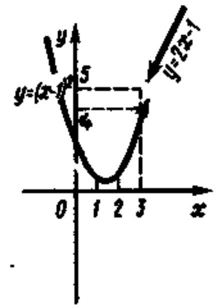


Рис. 47

Пример. Найти односторонние пределы при $x\to 3$ функции

$$f(x) = \begin{cases} (x-1)^2, & x<3; \\ 2x-1, & x>3. \end{cases}$$

Существует ли $\lim\limits_{x\to 3} f(x)$?

○ При $x<3$ имеем $f(x) = (x-1)^2$; поэтому $f(3-0) = \lim\limits_{x\to 3-0} (x-1)^2 = (3-1)^2 = 2^2 = 4$.

При $x>3$ имеем $f(x) = 2x-1$; значит, $f(3+0) = \lim\limits_{x\to 3+0} (2x-1) = 2\cdot 3 - 1 = 5$. Так как $f(3-0) \neq f(3+0)$, то $\lim\limits_{x\to 3} f(x)$ не существует. Геометрическая иллюстрация представлена на рис. 47.●

4. Бесконечно большие функции при $x\to a$

Выше (см. п. 4 § 15) мы ввели понятие бесконечно большой функции при $x\to\infty$. Аналогично формулируются определение и свойства бесконечно большой функции при $x\to a$.

Определение 7. Функция $y=f(x)$ называется *бесконечно большой при* $x\to a$, если для любого $P>0$ можно указать проколотую δ-окрестность точки a, во всех точках x которой выполняется неравенство $|f(x)|>P$.

Для бесконечно большой функции используется запись $\lim\limits_{x\to a} f(x) = \infty$.

Так как условие $x\to a$ распадается на два ($x\to a$ слева и $x\to a$ справа), а символ ∞ включает два случая ($-\infty$ и $+\infty$), то равенство $\lim\limits_{x\to a} f(x) = \infty$ объединяет четыре случая:

1) $\lim\limits_{x\to a-0} f(x) = +\infty$ (рис. 48, *а*); 2) $\lim\limits_{x\to a+0} f(x) = +\infty$ (рис. 48, *б*);

3) $\lim\limits_{x\to a-0} f(x) = -\infty$ (рис. 49, *а*); 4) $\lim\limits_{x\to a+0} f(x) = -\infty$ (рис. 49, *б*).

Во всех случаях прямая $x=a$ называется *вертикальной асимптотой* графика функции $y=f(x)$.

Если на практике возникает необходимость применить определение 8 конкретно к одному из случаев 1—4, то учитывают, что в случаях 1 и 2 можно писать $f(x)>P$ вместо $|f(x)|>P$, а в случаях 3 и 4 $f(x)<-P$ вместо $|f(x)|>P$.

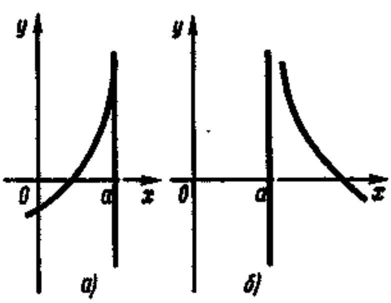


Рис. 48

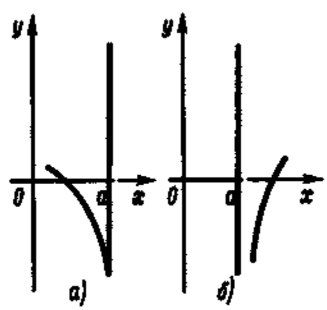


Рис. 49

Так, равенство $\lim\limits_{x\to a+0} f(x) = -\infty$ означает (в сокращенной кванторной записи), что

$$(\forall P>0)(\exists\delta>0)(\forall x: 0<x-a<\delta) f(x)<-P.$$

Справедливы теоремы о бесконечно больших функциях при $x\to a$, аналогичные соответствующим теоремам для бесконечно больших функций при $x\to\infty$, поэтому мы приводим их без доказательств.

Теорема 12. *Для того чтобы функция $f(x)$ была бесконечно большой при $x\to a$, необходимо и достаточно, чтобы функция $1/f(x)$ была бесконечно малой при $x\to a$.*

Например, функция $(x-3)^2$ — бесконечно малая при $x\to 3$, а функция $1/(x-3)^2$ — бесконечно большая при $x\to 3$.

Теорема 13. *Если $f(x)$ — бесконечно большая функция при $x\to a$, а функция $g(x)$ такова, что $\lim\limits_{x\to a} g(x) = b \neq 0$, то и $f(x)g(x)$ является бесконечно большой функцией при $x\to a$.*

Пример. Вычислить $\lim\limits_{x\to 1}\dfrac{2x+1}{x^2-1}$.

○ Здесь $\lim\limits_{x\to 1}(2x+1)=3$, а $\lim\limits_{x\to 1}(x^2-1)=0$; значит, теорему о пределе частного применить нельзя. Заметим, что $\lim\limits_{x\to 1}\dfrac{x^2-1}{2x+1}=\dfrac{0}{3}=0$, т. е. функция $\dfrac{x^2-1}{2x+1}$ есть бесконечно малая при $x\to 1$. Тогда по теореме 12 функция $\dfrac{2x+1}{x^2-1}$ является бесконечно большой при $x\to 1$, т. е. $\lim\limits_{x\to 1}\dfrac{2x+1}{x^2-1}=\infty$. ●

Замечание. Применив рассуждения, проведенные в этом примере, к общему случаю вычисления предела $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$, где $\lim\limits_{x\to a} f(x) = b \neq 0$, а $\lim\limits_{x\to a} g(x) = 0$, заключаем, что при указанных условиях $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}=\infty$.

Например, $\lim\limits_{x\to 5}\dfrac{x^2}{3x-15}=\infty$, так как $\lim\limits_{x\to 5}x^2=25$, а $\lim\limits_{x\to 5}(3x-15)=0$.

5. Глобальные и локальные свойства функций

Свойства функций разделяются на глобальные и локальные. *Глобальным* называют такое свойство функции, которое выполняется во всей области определения или на некотором заданном промежутке. К глобальным свойствам

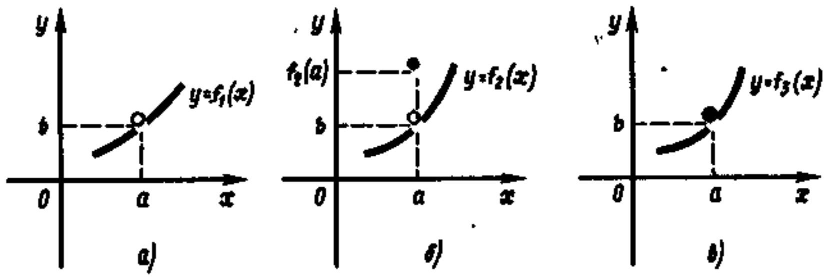


Рис. 50

относятся, например, четность, нечетность, периодичность, ограниченность, монотонность функции. *Локальным* называют такое свойство функции, которое выполняется в некоторой окрестности рассматриваемой точки. К локальным свойствам относятся, например, такие, как существование предела функции в точке $x=a$ или односторонних пределов, локальная ограниченность функции, имеющей предел в точке (см. теорему 7), свойство функции быть в данной точке бесконечно малой или бесконечно большой. В основном математический анализ изучает локальные свойства функций; об этом напоминает само слово «анализ», т. е. локальное исследование.

В этом пункте мы рассмотрим еще одно локальное свойство функции, имеющей предел в точке a.

Теорема 14. *Если* $\lim\limits_{x\to a} f(x)=b$ *и* $b>0$ $(b<0)$, *то в некоторой проколотой окрестности точки a выполняется неравенство* $f(x)>0$ $(f(x)<0)$.

□ Положим $\varepsilon=b/2$. Так как $\lim\limits_{x\to a} f(x)=b$, то существует $\delta>0$ такое, что в проколотой δ-окрестности точки a выполняется неравенство $|f(x)-b|<\varepsilon$ или что то же самое, $b-\varepsilon<f(x)<b+\varepsilon$. Поскольку $\varepsilon=b/2$, неравенство $f(x)>b-\varepsilon$ означает, что $f(x)>b-b/2=b/2>0$. Итак, $f(x)>0$ в проколотой δ-окрестности точки a. ■

§ 17. Непрерывные функции

1. Непрерывность функции в точке

На рис. 50 изображены графики трех различных функций, обладающих одним и тем же свойством: $\lim\limits_{x\to a} f_n(x)=b$ $(n=1, 2, 3)$. Эти функции отличаются друг от друга своим поведением в точке a: функция $f_1(x)$ не определена в этой точке (рис. 50, *а*); функция $f_2(x)$ определена в этой точке, но $f_2(x)\neq b$ (рис. 50, *б*); функция $f_3(x)$ определена в точке a, причем $f_3(a)=b$ (рис. 50, *в*).

Если поставить вопрос, какая из трех функций непрерывна в точке a, то ответ на него интуитивно ясен: функция $f_3(x)$. Условием, обеспечивающим непрерывность функции $f_3(x)$ в точке a, является, видимо, равенство $f_3(a)=b$, т. е. $f_3(a)=\lim\limits_{x\to a} f_3(x)$. Именно оно и кладется в основу соответствующего определения.

Определение 1. Функция $y=f(x)$, определенная в некоторой окрестности точки a, называется *непрерывной в точке a*, если $\lim\limits_{x\to a} f(x)=f(a)$.

Простейшими примерами непрерывных в любой точке a функций являются постоянная функция $y = c$ (поскольку $\lim\limits_{x \to a} c = c$) и функция $y = x$ (так как $\lim\limits_{x \to a} x = a$).

Непрерывность в точке является локальным свойством функций (см. п. 5 § 16).

Использовав для предела функции определения «на языке бесконечно малых», «на языке $\varepsilon - M$», «на языке окрестностей», получим соответствующие определения непрерывности функции, эквивалентные друг другу и определению 1. Наличие разных определений одного и того же понятия удобно тем, что в различных ситуациях полезным оказывается то или иное определение.

Чтобы сформулировать определение непрерывности «на языке бесконечно малых», введем понятие приращения функции: разность значений функции $f(x)$ в точках x и a, т. е. $f(x) - f(a)$ называют *приращением функции* и обозначают Δf. Тогда получим следующее определение:

Определение 2. Функция $y = f(x)$, определенная в некоторой окрестности точки a, называется *непрерывной в точке* a, если приращение Δf функции бесконечно мало при $x \to a$.

Прежде чем сформулировать определения непрерывности «на языке $\varepsilon - \delta$» и на «языке окрестностей», заметим, что в случае непрерывности функции в точке a нельзя исключать из рассмотрения саму точку a. Поэтому соответствующие определения примут следующий вид (ср. с определениями 4 и 5 § 16):

Определение 3. Функция $y = f(x)$, определенная в некоторой окрестности точки a, называется *непрерывной в точке* a, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при всех x из неравенства $|x - a| < \delta$ следует неравенство $|f(x) - f(a)| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x : |x - a| < \delta) |f(x) - f(a)| < \varepsilon.$$

Определение 4. Функция $y = f(x)$, определенная в некоторой окрестности точки a, называется *непрерывной в точке* a, если для любого $\varepsilon > 0$ можно указать такую δ-окрестность точки a, в которой выполняется неравенство $|f(x) - f(a)| < \varepsilon$.

Используя свойства пределов (см. теоремы 11, 7, 14 и 10 § 16), сформулируем соответствующие свойства непрерывных функций (их доказательства легко провести самостоятельно).

Теорема 1 (необходимое и достаточное условие непрерывности функции в точке). *Для того чтобы функция $y = f(x)$, определенная в некоторой окрестности точки a, была непрерывна в точке a, необходимо и достаточно, чтобы существовали односторонние пределы $\lim\limits_{x \to a - 0} f(x)$ и $\lim\limits_{x \to a + 0} f(x)$, равные друг другу и значению функции в точке a:*

$$f(a - 0) = f(a + 0) = f(a).$$

Теорема 2 (о локальной ограниченности непрерывной функции). *Если функция непрерывна в точке, то она ограничена в некоторой окрестности этой точки.*

Теорема 3 (о сохранении непрерывной функцией постоянного знака в окрестности точки). *Если функция $y = f(x)$ непрерывна в точке a и $f(a) > 0$ ($f(a) < 0$), то в некоторой окрестности точки a выполняется неравенство $f(x) > 0$ ($f(x) < 0$).*

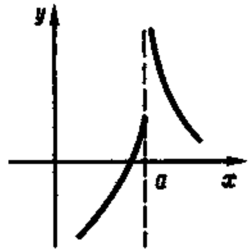


Рис. 51

Теорема 4 **(об арифметических операциях над непрерывными функциями).** *Если функции $y = f(x)$ и $y = g(x)$ непрерывны в точке a, то их сумма, произведение и частное (в случае $g(a) \neq 0$) также непрерывны в точке a.*

Используя последнюю теорему, нетрудно доказать такое утверждение.

Теорема 5. *Всякая рациональная функция, т. е. функция вида $y = P(x)/Q(x)$, где $P(x)$ и $Q(x)$ — многочлены, непрерывна в любой точке a, в которой $Q(a) \neq 0$.*

Доказательство, в котором, кроме теоремы 4 используется непрерывность функций $y = c$ и $y = x$, рекомендуем провести самостоятельно.

Пример. Вычислить $\lim\limits_{x\to-1} \dfrac{2x^3+x^2-x-1}{3x^2+4x+3}$.

○ Рассмотрим рациональную функцию $f(x) = \dfrac{2x^3+x^2-x-1}{3x^2+4x+3}$. Ее знаменатель не обращается в нуль при $x = -1$; значит, по теореме 5 функция $f(x)$ непрерывна в точке $x = -1$. Тогда в силу определения 1 имеем $\lim\limits_{x\to-1} f(x) = f(-1)$. Таким образом, находим $\lim\limits_{x\to-1} \dfrac{2x^3+x^2-x-1}{3x^2+4x+3} = \dfrac{2\cdot(-1)^3+(-1)^2-(-1)-1}{3\cdot(-1)^2+4\cdot(-1)+3} = -\dfrac{1}{2}$. ●

2. Точки разрыва

Определение 5. Если функция определена в некоторой проколотой окрестности точки a и не является непрерывной в точке a, то говорят, что $x = a$ — *точка разрыва* функции $f(x)$.

В силу теоремы 1 точка $x = a$ является точкой разрыва функции $f(x)$ в одном из следующих случаев:

1. Односторонние пределы функции $f(x)$ при $x \to a$ существуют, равны между собой, но не равны $f(a)$ (см. рис. 50, *а* и *б*). В этом случае точку a называют *точкой устранимого разрыва* (такое название объясняется тем, что достаточно переопределить или доопределить функцию в точке a, полагая $f(a) = f(a-0) = f(a+0)$, чтобы она стала непрерывной в точке a).

2. Односторонние пределы существуют, но не равны: $f(a-0) \neq f(a+0)$ (см. рис. 47). В этом случае точку a называют *точкой скачка* функции.

Точки устранимого разрыва и точки скачка объединяют общим термином: *точки разрыва первого рода*.

3. Хотя бы один из односторонних пределов не существует (рис. 51). В этом случае точку a называют *точкой разрыва второго рода*.

Теорема 6. *Пусть функция $y = f(x)$ определена и монотонна на промежутке X и a — внутренняя точка этого промежутка. Если функция имеет разрыв в точке a, то только скачок.*

□ Будем считать для определенности, что $X = (c, d)$, $c < a < d$ и функция возрастает на (c, d). Рассмотрим функцию на (c, a). В силу ее возрастания для любого x из (c, a) имеем $f(x) < f(a)$. Значит, множество Y значений функции на интервале (c, a) ограничено сверху и потому оно имеет верхнюю грань (см. § 10, п. 2). Положим $b = \sup Y$.

Возьмем произвольное $\varepsilon > 0$. Тогда число $b - \varepsilon$ не является верхней границей для множества Y и, следовательно, в интервале (c, a) найдется точка x_1 такая, что $f(x_1) > b - \varepsilon$.

Положим $x_1 = a - \delta$. Если $x_1 < x < a$, то $f(x_1) < f(x)$, а значит, $f(x) > b - \varepsilon$. С другой стороны, если $x_1 < x < a$, то $f(x) < b + \varepsilon$. Итак, из неравенства $a - \delta < x < a$ или, что то же самое, $0 < a - x < \delta$, следует неравенство $b - \varepsilon < f(x) < b + \varepsilon$ или, что то же самое, $|f(x) - b| < \varepsilon$. Но это и означает (см. п. 3 § 16), что $\lim\limits_{x \to a-0} f(x) = b$, причем

$$\lim_{x \to a-0} f(x) \leqslant f(a). \tag{1}$$

Аналогично доказывается, что существует $\lim\limits_{x \to a+0} f(x)$, причем

$$\lim_{x \to a+0} f(x) \geqslant f(a). \tag{2}$$

Так как $x = a$ есть точка разрыва, то хотя бы одно из неравенств (1), (2) является строгим. Значит, $f(a-0)$, $f(a+0)$ существуют; но $f(a-0) < f(a+0)$, т. е. $x = a$ — точка скачка. ■

3. Непрерывность функции на промежутке

Определение 6. Если функция $y = f(x)$ непрерывна в каждой точке интервала (a, b), то она называется *непрерывной на интервале* (a, b).

Это определение распространяется и на случай бесконечных интервалов, т. е. промежутков вида $(-\infty, b)$, $(a, +\infty)$, $(-\infty, +\infty)$. Например, функция $y = x^2$ непрерывна на $(-\infty, +\infty)$, а функция $y = 1/x$ непрерывна на каждом из двух промежутков: $(-\infty, 0)$, $(0, +\infty)$.

Определение 7. Говорят, что функция $y = f(x)$ *непрерывна в точке* $x = a$ *справа (слева)*, если $\lim\limits_{x \to a+0} f(x) = f(a)$ $(\lim\limits_{x \to a-0} f(x) = f(a))$.

Определение 8. Функция $y = f(x)$ называется *непрерывной на отрезке* $[a, b]$, если она непрерывна на интервале (a, b) и непрерывна в точке a справа и в точке b слева.

Геометрическую иллюстрацию функции, непрерывной на отрезке $[a, b]$, дает рис. 32.

§ 18. Свойства функций, непрерывных на промежутках

1. Теорема о промежуточном значении

Теорема 1 (о нуле непрерывной функции). *Пусть функция $y = f(x)$ непрерывна на отрезке $[a, b]$ и принимает на концах этого отрезка значения разных знаков. Тогда найдется точка отрезка $[a, b]$, в которой эта функция обращается в нуль.*

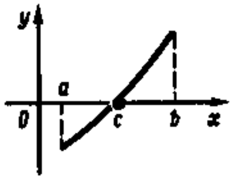


Рис. 52

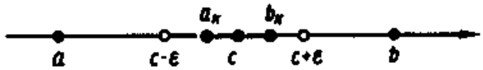


Рис. 53

□ Предположим для определенности, что $f(a)<0$ и $f(b)>0$ (рис. 52). Разделим отрезок $[a, b]$ пополам. Если в точке $\frac{a+b}{2}$ функция $f(x)$ обращается в нуль, т. е. $f\left(\frac{a+b}{2}\right)=0$, то теорема доказана. В противном случае на концах одной из половин отрезка функция принимает значения разных знаков. Снова разделим эту половину на два равных отрезка и продолжим указанный процесс. Возможны два случая:

1) в ходе последовательных делений получаем точку, в которой функция обращается в нуль. Тогда теорема доказана;

2) получаем бесконечную последовательность отрезков $[a, b]$, $[a_1, b_1]$, ..., $[a_n, b_n]$, ..., вложенных друг в друга и таких, что $f(a_n)<0$, $f(b_n)>0$. Согласно теореме о стягивающейся системе отрезков (см. теорему 4 § 10), найдется точка c, принадлежащая всем этим отрезкам.

Докажем, что $f(c)=0$. Предположим противное, т. е. что, например, $f(c)>0$. Тогда в силу теоремы 3 § 17 найдется такая окрестность $(c-\varepsilon, c+\varepsilon)$ точки c, во всех точках которой $f(x)>0$. Однако на некотором этапе деления отрезка $[a, b]$ получится отрезок $[a_k, b_k]$, целиком лежащий внутри $(c-\varepsilon, c+\varepsilon)$ (рис. 53), и тогда $f(a_k)<0$. Значит, в окрестности $(c-\varepsilon, c+\varepsilon)$ имеется точка, в которой значение функции отрицательно. Полученное противоречие означает, что предположение неверно, а потому $f(c)=0$. ■

Доказанное утверждение геометрически очевидно: ось абсцисс делит координатную плоскость на две полуплоскости; если $f(a)<0$, а $f(b)>0$, то точки $A(a; f(a))$ и $B(b; f(b))$ лежат в разных полуплоскостях, и поэтому непрерывная линия, соединяющая точки A и B, обязательно пересечет ось абсцисс хотя бы один раз (см. рис. 52).

При выполнении условий теоремы 1 функция $f(x)$ может обратиться в нуль несколько раз. Однако если функция $f(x)$ монотонна на $[a, b]$, то существует не более одной точки c такой, что $f(c)=0$ (если $f(x)$ возрастает, то при $x<c$ получаем, что $f(x)<f(c)=0$, а при $x>c$ — что $f(x)>f(c)=0$. Итак, справедливо такое следствие из теоремы 1.

Следствие. *Если функция $f(x)$ непрерывна и монотонна на отрезке $[a, b]$, причем знаки $f(a)$ и $f(b)$ различны, то на $[a, b]$ существует одна и только одна точка c такая, что $f(c)=0$.*

Доказанное следствие используют для приближенного вычисления корней уравнений. Пусть требуется решить уравнение $f(x)=0$. Если удается найти отрезок $[a, b]$, на котором функция $f(x)$ непрерывна, монотонна и на его концах принимает значения разных знаков, то внутри этого отрезка содержится ровно один корень уравнения. Чтобы уточнить значение искомого корня, отрезок $[a, b]$ делят пополам, выбирают половину, на концах которой значения функции различны по знаку, и т. д. Таким путем получают значение корня уравнения $f(x)=0$ с любой наперед заданной точностью.

Часто поступают так: ищут отрезок вида $[n, n+1]$, на котором лежит корень уравнения $f(x)=0$, делят этот отрезок на 10 равных частей, вычисляют значения функции в точках деления и устанавливают, на каком из отрезков функция меняет знак. Этот отрезок снова делят на 10 равных частей и т. д. Указанный процесс позволяет находить один за другим десятичные знаки искомого корня.

Пример 1. Доказать, что уравнение $x^3+4x+1=0$ имеет корень на отрезке $[-1, 0]$, и найти приближенное значение этого корня с точностью до 0,01.

○ Функция $f(x)=x^3+4x+1$ непрерывна на отрезке $[-1, 0]$, возрастает (как сумма двух возрастающих функций x^3 и $4x+1$) и принимает на концах отрезка значения разных знаков: $f(-1)=-4<0$, $f(0)=1>0$. Значит, на отрезке $[-1, 0]$ заданное уравнение имеет один корень.

Для вычисления этого корня применим метод последовательного деления. Разделим отрезок $[-1, 0]$ на 10 равных частей и вычислим значения функции в точках деления. Имеем $f(0)=1$; $f(-0,1)=0,599$; $f(-0,2)=0,192$; $f(-0,3)=-0,227$. Таким образом $f(-0,3)<0$, $f(-0,2)>0$; следовательно, искомый корень лежит на отрезке $[-0,3; -0,2]$. Разделим этот отрезок снова на 10 равных частей и вычислим значения функции в точках деления. Находим $f(-0,21)\approx 0,151$; ...; $f(-0,24)\approx 0,026$; $f(-0,25)\approx$ $\approx -0,016$. Значит, искомый корень лежит между $-0,24$ и 0,25. С точностью до 0,01 можно положить $x\approx -0,245$. ●

Теорема 1 играет важную роль и при решении неравенств.

Предположим, что требуется решить неравенство $f(x)>0$, где $f(x)$ — функция, непрерывная на всей числовой прямой. Сначала решим уравнение $f(x)=0$. Пусть x_1, x_2, ..., x_n — корни уравнения (множество этих корней может быть как конечным, так и бесконечным). Эти точки разбивают числовую прямую на такие промежутки, что на концах каждого промежутка функция $f(x)$ обращается в нуль, а внутри отлична от нуля. Тогда внутри каждого промежутка знак функции $f(x)$ не меняется. В самом деле, если предположить, что точки a и b лежат внутри одного промежутка, а знаки $f(a)$ и $f(b)$ различны, то непрерывная функция $f(x)$ обратится в нуль между точками a и b, что невозможно.

Итак, мы доказали, что корни уравнения $f(x)=0$ разбивают числовую прямую на промежутки, внутри которых функция $f(x)$ сохраняет постоянный знак. Теперь достаточно взять на каждом из полученных интервалов «пробную» точку и установить знак функции в этой точке. Тот же знак она будет иметь и на всем интервале. Это позволит отобрать те интервалы, на которых выполняется неравенство $f(x)>0$.

Если функция $f(x)$ имеет точки разрыва, то при решении неравенства $f(x)>0$ следует учесть не только корни уравнения $f(x)=0$, но и точки разрыва функции, и установить, какой знак имеет функция на каждом из полученных интервалов.

Указанный метод решения неравенств называется ***методом интервалов***.

Пример 2. Решить неравенство $\frac{(x+1)(x-1)^3(x+2)^2}{x-4}>0$.

○ Функция $\frac{(x+1)(x-1)^3(x+2)^2}{x-4}$ обращается в нуль в точках -2, -1, 1 и разрывна в точке 4. Эти точки разбивают числовую прямую на промежутки $(-\infty, -2]$, $[-2, -1]$, $[-1, 1]$, $[1, 4)$, $(4, +\infty)$. Внутри каждого из этих промежутков функция непрерывна и сохраняет постоянный знак. Взяв пробные точки -3; $-1,5$; 0; 2; 5, находим: $f(-3)<0$, $f(-1,5)<0$, $f(0)>0$, $f(2)<0$, $f(5)>0$. Значит, $f(x)>0$ на промежутках $(-1; 1)$ и $(4; +\infty)$. Их объединение и является решением заданного неравенства. ●

Теорема 1 является частным случаем более общего утверждения.

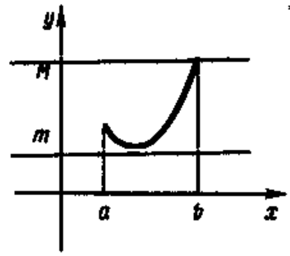


Рис. 54

Теорема 2 (о промежуточном значении). *Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и $f(a) \neq f(b)$. Тогда она принимает на этом отрезке любое значение μ, лежащее между $f(a)$ и $f(b)$, т. е. существует такая точка $c \in [a, b]$, что $f(c) = \mu$.*

□ Положим для определенности, что $f(a) < f(b)$. Рассмотрим вспомогательную функцию $F(x) = f(x) - \mu$. Эта функция непрерывна на отрезке $[a, b]$ как разность двух непрерывных функций $f(x)$ и μ (постоянная функция). При этом так как $f(a) < \mu$, то $F(a) = f(a) - \mu < 0$, а так как $f(b) > \mu$, то $F(b) = f(b) - \mu > 0$. Значит, функция $F(x)$ на концах отрезка $[a, b]$ принимает значения различных знаков. Следовательно, на этом отрезке найдется такая точка c, что $F(c) = 0$. Это означает, что $f(c) - \mu = 0$, т. е. что $f(c) = \mu$. ■

Следствие. ***Если функция $f(x)$ непрерывна на промежутке с концами a и b и монотонна на нем, то область ее значений представляет собой промежуток с концами $f(a)$ и $f(b)$.***

Справедливо обратное утверждение.

Теорема 3. *Если функция $f(x)$ возрастает (убывает) на промежутке X и областью ее значений является промежуток Y, то эта функция непрерывна на X.*

□ Если бы возрастающая функция $f(x)$ имела на X точку разрыва c, то последняя была бы точкой разрыва первого рода — скачка (см. теорему 6 § 17), т. е. получили бы, что $f(c-0) < f(c)$ или $f(c) < f(c+0)$. В этом случае промежуток $(f(c-0), f(c))$ или $(f(c), f(c+0))$, представляющий собой часть Y, не содержал бы ни одной точки, являющейся значением функции $f(x)$, вопреки тому, что по условию $f(X) = Y$. Полученное противоречие доказывает непрерывность функции $f(x)$ на X. ■

2. Ограниченность функции, непрерывной на отрезке

На рис. 54 изображен график функции $f(x)$, непрерывной на отрезке $[a, b]$. Мы видим, что этот график целиком расположен между двумя прямыми, параллельными оси абсцисс. Иными словами, функция $f(x)$ ограничена. Докажем, что это утверждение справедливо для любой функции, непрерывной на отрезке.

Теорема 4. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она ограничена на нем (т. е. существует такое число A, что $|f(x)| \leqslant A$ для всех $x \in [a, b]$).*

□ Доказательство проведем методом от противного. Предположим, что функция $f(x)$ не ограничена на отрезке $[a, b]$. Разобьем этот отрезок пополам. Тогда хотя бы на одной из половин функция окажется неограниченной:

если бы она была ограничена на обеих половинах, то она была бы ограниченной и на всем отрезке $[a, b]$. Ту часть отрезка $[a, b]$, на которой функция не ограничена, обозначим $[a_1, b_1]$ и снова разделим пополам. Выберем ту половину отрезка $[a_1, b_1]$, на которой функция не ограничена (если функция не ограничена на обеих половинах отрезка, то можно выбирать любую из них), и обозначим выбранный отрезок $[a_2, b_2]$. Продолжая описанный процесс последовательного деления и выбора отрезков, получим стягивающуюся систему отрезков $[a, b]$, $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$, ..., на каждом из которых функция $f(x)$ не ограничена. Эти отрезки стягиваются к некоторой точке c отрезка $[a, b]$, в которой, как и во всех точках $[a, b]$, функция $f(x)$ непрерывна. Тогда существует окрестность $(c-\varepsilon, c+\varepsilon)$ точки c, в которой функция $f(x)$ ограничена (см. теорему 2 § 17). Вследствие того, что длины отрезков $[a_n, b_n]$ стремятся к нулю, при некотором n длина отрезка $[a_n, b_n]$ станет меньше ε и, значит, этот отрезок целиком окажется внутри указанной окрестности, то она ограничена и на $[a_n, b_n]$. Однако в процессе построения мы брали отрезки так, что на каждом из них функция $f(x)$ была не ограничена. Полученное противоречие показывает, что сделанное предположение о неограниченности функции $f(x)$ на $[a, b]$ неверно. Итак, функция $f(x)$ ограничена на $[a, b]$. ■

Заметим, что эта теорема неверна для функций, непрерывных на интервале (a, b), — такие функции могут оказаться и неограниченными. Например, функция $1/x$ непрерывна на интервале $(0, 1)$, но не ограничена на нем: $\lim\limits_{x\to+\infty} 1/x = +\infty$.

Пусть Y — множество значений функции $f(x)$, непрерывной на $[a, b]$. Из теоремы 4 следует, что множество Y ограничено. Поэтому оно имеет нижнюю и верхнюю грани. Обозначим их так: $m = \inf Y$, $M = \sup Y$. Функция $f(x)$ не может принять значений, бо́льших чем M, или меньших чем m. Покажем, что на отрезке $[a, b]$ функция обязательно принимает значения m и M, т. е. что среди ее значений имеются и наименьшее, и наибольшее; они соответственно обозначаются $y_{\text{наим}}$ и $y_{\text{наиб}}$.

Теорема 5. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то среди ее значений на этом отрезке имеются наименьшее значение m и наибольшее значение M.*

□Предположим, что функция $f(x)$ ни в одной точке отрезка $[a, b]$ не принимает значения M. Рассмотрим вспомогательную функцию $\varphi(x) = \frac{1}{M-f(x)}$. Согласно предположению, ни в одной точке отрезка $[a, b]$ функция $f(x)$ не принимает значения M, а потому знаменатель выражения $\frac{1}{M-f(x)}$ не обращается в нуль, причем $M-f(x)>0$ при всех $x\in[a, b]$. Так как функция $\varphi(x)$ непрерывна во всех точках отрезка $[a, b]$, то по теореме 4 она ограничена на этом отрезке, т. е. существует такое число B, что для всех $x\in[a, b]$ выполняется неравенство $0<\frac{1}{M-f(x)}<B$. Но тогда $M-f(x)>\frac{1}{B}$ и, значит, $f(x)<M-\frac{1}{B}$. Последнее неравенство показывает, что число $M-\frac{1}{B}$ является одной из верхних границ для области Y значений функции $f(x)$ на $[a, b]$. Однако это невозможно, так как M — наименьшая из верхних границ для Y.

Полученное противоречие показывает, что сделанное предположение неверно, т. е. существует точка $c_1\in[a, b]$ такая, что $f(c_1)=M$. Точно так же доказывается существование такой точки $c_2\in[a, b]$, что $f(c_2)=m$.■

Из доказанной теоремы и теоремы о промежуточном значении вытекает такое утверждение.

Следствие. *Если функция $f(x)$ непрерывна на $[a, b]$, а m и M — ее наименьшее и наибольшее значения на этом отрезке, то областью значений функции является отрезок $[m, M]$* (см. рис. 54).

3. Непрерывность обратной функции

Теорема 6. *Пусть функция $y = f(x)$ монотонна и непрерывна на промежутке X, а Y — область значений функции. Тогда существует обратная функция $x = f^{-1}(y)$, монотонная и непрерывная на Y.*

□ Существование и монотонность обратной функции $x = f^{-1}(y)$ следует из теорем 4 и 5 § 14.

В силу того, что функция $y = f(x)$ монотонна и непрерывна на промежутке X, на основании следствия из теоремы 2 заключаем, что область ее значений есть промежуток Y.

Наконец, так как обратная функция определена и монотонна на промежутке Y, а областью ее значений является промежуток X, по теореме 3 обратная функция $x = f^{-1}(y)$ непрерывна на Y. ■

§ 19. Степенная функция с рациональным показателем

Степенной функцией называется функция вида $y = x^r$, где r — рациональное число.

1. Степенная функция с натуральным показателем

Функция $y = x^n$, где n — натуральное число, обладает следующими свойствами:

1^0. *Область определения:* $(-\infty, +\infty)$.

2^0. *Функция является четной при четном n и нечетной при нечетном n.*

3^0. *Если n нечетно, то функция $y = x^n$ возрастает на* $(-\infty, +\infty)$; *если n четно, то функция $y = x^n$ возрастает на луче* $[0, +\infty]$ *и убывает на луче* $(-\infty, 0]$.

4^0. *Функция непрерывна на* $(-\infty, +\infty)$.

График функции $y = x^n$ в частных случаях $n = 1, 2, 3$ построен в § 14: графиком функции $y = x$ является прямая, графиком функции $y = x^2$ — парабола, графиком функции $y = x^3$ — кубическая парабола.

Если n — четное число, большее двух, то график функции $y = x^n$ напоминает параболу; ветви графика тем круче направлены вверх, чем больше n (рис. 55). Если n — нечетное число, большее трех, то график функции $y = x^n$ напоминает кубическую параболу (рис. 56).

2. Степенная функция с целым отрицательным показателем

Функция $y = x^{-n}$, где n — натуральное число, обладает следующими свойствами:

1^0. *Область определения*: $(-\infty; 0) \cup (0; +\infty)$.

2^0. *Функция является четной при четном n и нечетной при нечетном n.*

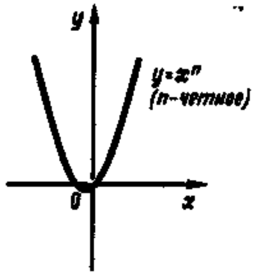


Рис. 55

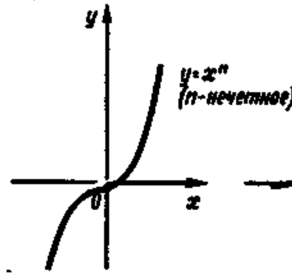


Рис. 56

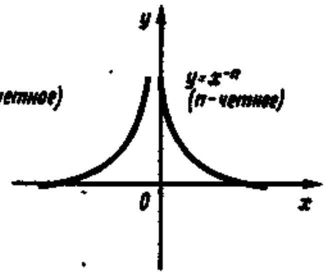


Рис. 57

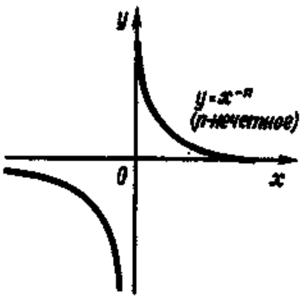


Рис. 58

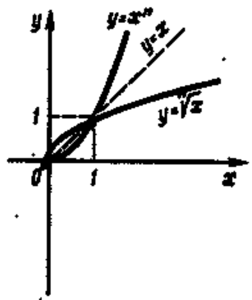


Рис. 59

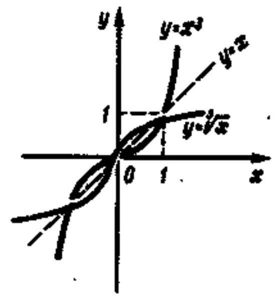


Рис. 60

3^0. *Функция убывает при* $x>0$; *она убывает при* $x<0$, *если* n *нечетно, и возрастает при* $x<0$, *если* n *четно* (см. п. 4 § 14).

4^0. *Функция непрерывна на* $(-\infty, 0)$ *и на* $(0, +\infty)$; $x=0$ — *точка разрыва второго рода.*

Графиком функции $y=x^{-1}$, т. е. $y=1/x$, является гипербола; аналогичный вид имеет график функции $y=x^{-n}$ при нечетном n (рис. 57). В случае четного n график функции $y=x^{-n}$ имеет вид, изображенный на рис. 58. Оси x и y являются соответственно горизонтальной и вертикальной асимптотами графика функции $y=x^{-n}$.

3. Функция $y=\sqrt[n]{x}$

Рассмотрим функцию $y=x^n$, $x\in[0, +\infty)$, где n — натуральное число, большее единицы. Эта функция возрастает и непрерывна на луче $[0, +\infty)$, а областью ее значений является луч $[0, +\infty)$. По теореме 6 § 18 для нее существует обратная функция, которая обозначается $x=\sqrt[n]{y}$. Поменяв местами x и y, получим $y=\sqrt[n]{x}$. Эта функция определена, возрастает и непрерывна на луче $[0, +\infty)$. Отобразив график функции $y=x^n$ симметрично относительно прямой $y=x$, получим график функции $y=\sqrt[n]{x}$ (рис. 59).

Заметим, что если n — нечетное число, то функция $y=x^n$ возрастает и непрерывна на множестве действительных чисел. Значит, в этом случае можно говорить о функции, обратной к функции $y=x^n$, $x\in\mathbb{R}$. Она также обозначается $y=\sqrt[n]{x}$.

На рис. 60 изображены графики функций $y=x^3$ и $y=\sqrt[3]{x}$.

4. Степенная функция с любым рациональным показателем

Здесь могут представиться три случая: $r>1$, $0<r<1$, $r<0$.

I случай: $r>1$. Пусть $r=m/n$, где $m>n$.

Отметим свойства функции $y=x^{m/n}$.

1°. *Область определения — луч* $[0, +\infty)$.

2°. *На луче* $[0, +\infty)$ *функция возрастает.*

□ Пусть $0<x_1<x_2$. В предыдущем пункте мы установили, что функция $y=\sqrt[n]{x}$ возрастает. Значит, из $x_1<x_2$ следует $\sqrt[n]{x_1}<\sqrt[n]{x_2}$, т. е. $x_1^{1/n}<x_2^{1/n}$.

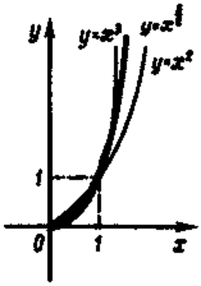


Рис. 61

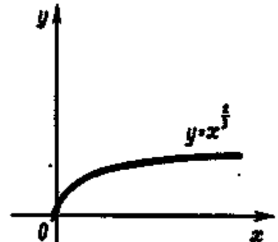


Рис. 62

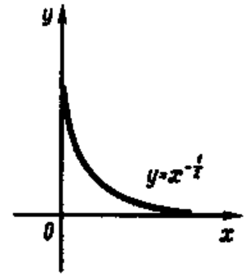


Рис. 63

Положим $t_1=x_1^{1/n}$, $t_2=x_2^{1/n}$ и рассмотрим функцию $y=t^m$. Она является возрастающей, т. е. из $t_1<t_2$ следует $t_1^m<t_2^m$. Таким образом, $(x_1^{1/n})^m<(x_2^{1/n})^m$, т. е. $x_1^{m/n}<x_2^{m/n}$. Итак, из $x_1<x_2$ следует $x_1^{m/n}<x_2^{m/n}$, а это и означает, что степенная функция $y=x^{m/n}$ на промежутке $[0, +\infty)$ возрастает. ■

3°. *Функция непрерывна* (как произведение m непрерывных функций вида $x^{1/n}$).

На рис. 61 изображен график функции $y=x^{5/2}$ (он заключен между графиками функций $y=x^2$ и $y=x^3$). Подобный вид имеет график любой функции вида $y=x^r$, $x\in[0, +\infty)$, где $r>1$. Чем больше r, тем круче вверх идет ветвь графика.

II случай: $0<r<1$. Как и в предыдущем случае, функция $y=x^r$ определена на луче $[0, +\infty)$, возрастает и непрерывна на нем. На рис. 62 изображен график функции $y=x^{2/3}$. Подобный вид имеет график любой степенной функции $y=x^r$, $x\in[0, +\infty)$, где $0<r<1$.

III случай: $r<0$. Пусть $r=-m/n$, где m и n — взаимно простые натуральные числа, причем $n\neq 1$.

Отметим свойства функции $y=x^{-m/n}$.

1°. *Область определения — луч* $(0, +\infty)$.

2°. *Функция убывает на луче* $(0, +\infty)$.

□ Пусть $0<x_1<x_2$. Тогда, как мы установили выше, $x_1^{m/n}<x_2^{m/n}$. На основании свойства числовых неравенств имеем $1/x_1^{m/n}>1/x_2^{m/n}$, т. е. $x_1^r>x_2^r$. Итак, из $x_1<x_2$ следует $x_1^r>x_2^r$, а это и означает, что функция на луче $(0, +\infty)$ убывает. ■

3°. *Функция непрерывна на* $(0, +\infty)$ (как частное двух непрерывных функций 1 и $x^{m/n}$).

Построим, например, график функции $y=x^{-1/2}$. Составим таблицу значений функции:

x	1	4	9	1/4	1/9
y	1	1/2	1/3	2	3

Нанесем полученные точки на координатную плоскость и соединим их плавной кривой (рис. 63). Подобный вид имеет график любой функции $y=x^r$, $x\in(0, +\infty)$, где $r<0$.

Оси x и y являются соответственно горизонтальной и вертикальной асимптотами графика функции $y=x^r$, где $r<0$.

§ 20. Показательная функция

1. Показательная функция на множестве рациональных чисел

Пусть a — положительное и отличное от 1 действительное число. Какое бы рациональное число r ни взять, можно вычислить a^r. Таким образом, на множестве **Q** всех рациональных чисел определена функция $y=a^x$ $(a>0, a\neq 1)$, которая называется *показательной функцией* с основанием a.

Отметим свойства показательной функции сначала для случая, когда $a>1$, а затем для случая, когда $0<a<1$.

I случай: $a>1$.

1^0. Если $x>0$, то $a^x>1$.

□ Пусть $x=p/q$, где $p, q\in \mathbf{N}$, $q\neq 1$; тогда $a^x=a^{p/q}=\sqrt[q]{a^p}$. Так как при $x>0$ функции $y=x^p$ и $y=\sqrt[q]{x}$ являются возрастающими, то $a>1 \Rightarrow a^p>$ $>1\Rightarrow \sqrt[q]{a^p}>\sqrt[q]{1^p}$. Но $\sqrt[q]{1^p}=1$; значит, $a^x>1$. ■

2^0. *Если $x<0$, то $0<a^x<1$.*

□ Пусть $x=-r$: тогда $r>0$. Согласно свойству 1^0, имеем $a^r>1$. Поэтому $0<1/a^r<1$, т. е. $0<a^{-r}<1$. Это и означает, что $0<a^x<1$. ■

3^0. *Функция принимает только положительные значения.*

Это непосредственно следует из свойств 1^0 и 2^0, а также из того, что $a=1$ при $x=0$.

4^0. *Функция возрастает на множестве* **Q**.

□ Пусть $x_1, x_2\in \mathbf{Q}$ и $x_1<x_2$; тогда $f(x_1)=a^{x_1}$, $f(x_2)=a^{x_2}$. Покажем, что $f(x_1)<f(x_2)$.

Имеем $f(x_2)-f(x_1)=a^{x_2}-a^{x_1}=a^{x_1}(a^{x_2-x_1}-1)$. Так как $x_2-x_1>0$, то $a^{x_2-x_1}>1$ (в силу свойства 1^0), а потому $a^{x_2-x_1}-1>0$. Кроме того, согласно свойству 3^0, имеем $a^{x_1}>0$. Значит, $a^{x_1}(a^{x_2-x_1}-1)>0$, т. е. $f(x_2)-f(x_1)>0$, откуда $f(x_1)<f(x_2)$. ■

Отметим, что возрастание функции происходит плавно, без скачков. Иными словами, если точки x_1 и x_2 выбираются достаточно близко друг к другу, то a^{x_1} и a^{x_2} незначительно отличаются друг от друга.

Точнее этот результат сформулируем в виде следующего свойства (доказательство которого мы не приводим).

5^0. *Для любого $\varepsilon>0$ существует $\delta>0$ такое, что* $|x_1-x_2|<\delta\Rightarrow$ $\Rightarrow|a^{x_1}-a^{x_2}|<\varepsilon$.

II случай: $0<a<1$. Положим $b=1/a$; тогда $b>1$. Для функции b^x имеем: $x>0\Rightarrow b^x>1$; $x<0\Rightarrow 0<b^x<1$; $x=0\Rightarrow b^x=1$. Но из $b^x>1$ следует $(1/a)^x>1$, т. е. $0<a^x<1$; из $0<b^x<1$ следует $0<(1/a)^x<1$, т. е. $a^x>1$; если же $x=0$, то $a^x=1$.

Таким образом, мы получаем следующие свойства показательной функции с основанием $0<a<1$:

1°. *Если $x>0$, то $0<a^x<1$.*

2°. *Если $x<0$, то $a^x>1$.*

3°. *Функция принимает только положительные значения.*

4°. *Функция убывает на множестве* $\mathbf{Q}$.

5°. *Для любого $\varepsilon>0$ существует $\delta>0$ такое, что $|x_1-x_2|<\delta \Rightarrow |a^{x_1}-a^{x_2}|<\varepsilon$.*

Замечание. Мы рассмотрели показательную функцию $y=a^x$, $x\in\mathbf{Q}$ в двух случаях: $a>1$ и $0<a<1$. Случай $a=1$ обычно не рассматривают, поскольку при этом получается функция 1^x, принимающая только одно значение 1, — такая функция не представляет интереса. Случай $a\leqslant 0$ также не рассматривают, поскольку для неположительных чисел возведение в рациональную степень не всегда возможно (например, $(-1)^{1/2}$ не является действительным числом).

2. Степень с иррациональным показателем

Пусть $a>0$ и α — иррациональное число. Определим, что мы будем понимать под записью a^α и какой смысл следует вкладывать в понятие степени положительного числа с иррациональным показателем.

Если $a=1$, то положим $1^\alpha=1$.

Пусть $a>1$. Рассмотрим два числовых множества A и B: множество A состоит из всех чисел вида a^r, где r — рациональное число такое, что $r<\alpha$; множество B состоит из всех чисел вида $a^{\tilde r}$, где $\tilde r$ — рациональное число такое, что $\tilde r>\alpha$. Так как $r<\tilde r$, то $a^r<a^{\tilde r}$ (это следует из возрастания функции $y=a^x$, $x\in\mathbf{Q}$, $a>1$). Значит, любой элемент множества A меньше любого элемента множества B, т. е. множество A расположено левее множества B, и, следовательно, (см. § 10, п. 3) существует число c, разделяющее эти множества: $a^r<c<a^{\tilde r}$ для любых r и $\tilde r$. Согласно свойству 5° из п. 1, числа $a^{\tilde r}$ и a^r можно выбрать сколь угодно близко друг к другу; тогда число c, разделяющее множества A и B, является единственным. Оно и принимается за a^α.

Итак, мы пришли к следующему определению.

Определение. **Под *степенью с иррациональным показателем* a^α ($a>1$, α — иррациональное число) понимается единственное число, разделяющее два множества: множество A, состоящее из чисел вида a^r, где $r<\alpha$, r — рациональное число, и множество B, состоящее из чисел вида $a^{\tilde r}$, где $\tilde r>\alpha$, $\tilde r$ — рациональное число.**

Из этого определения следует, что если $a>1$, то

$$r_1<\alpha<r_2 \Rightarrow a^{r_1}<a^\alpha<a^{r_2}. \tag{1}$$

Аналогично определяется степень с иррациональным показателем при $0<a<1$, только в этом случае, вследствие убывания функции $y=a^x$, $x\in\mathbf{Q}$ меняются обозначения для множеств A и B.

Отсюда следует, что если $0<a<1$, то

$$r_1<\alpha<r_2 \Rightarrow a^{r_2}<a^\alpha<a^{r_1}. \tag{2}$$

Можно доказать, что при таком определении понятия степени с иррациональным показателем для любых действительных чисел x_1, x_2 и для любых положительных чисел a, b справедливы следующие равенства, известные в случае рациональных показателей:

1^0. $a^{x_1}\cdot a^{x_2}=a^{x_1+x_2}$.
2^0. $a^{x_1}:a^{x_2}=a^{x_1-x_2}$.
3^0. $(a^{x_1})^{x_2}=a^{x_1x_2}$.
4^0. $a^{x_1}b^{x_1}=(ab)^{x_1}$.
5^0. $\frac{a^{x_1}}{b^{x_1}}=\left(\frac{a}{b}\right)^{x_1}$.

3. Показательная функция

Выше (см. п. 1) мы рассмотрели показательную функцию на множестве $\mathbb{Q}$ рациональных чисел. Теперь, после того как введено понятие степени с иррациональным показателем, мы можем изучить функцию $y=a^x$, определенную на множестве $\mathbb{R}$ действительных чисел.

Рассмотрим сначала случай, когда $a>1$.

1^0. *Область определения:* $(-\infty, +\infty)$.

2^0. *Функция $y=a^x$ — возрастающая.*

□ Пусть $x_1<x_2$. Докажем, что $a^{x_1}<a^{x_2}$.

Если x_1 и x_2 — рациональные числа, то $a^{x_1}<a^{x_2}$ (это уже было доказано в п. 1). Если же одно из чисел рационально, а другое иррационально, то неравенство $a^{x_1}<a^{x_2}$ следует из полученного в п. 2 неравенства (1). Пусть, наконец, x_1 и x_2 — иррациональные числа. Тогда, выбрав между ними рациональное число r, т. е. $x_1<r<x_2$, получим $a^{x_1}<a^r$, $a^r<a^{x_2}$, откуда $a^{x_1}<a^{x_2}$.

Итак, в любом случае $x_1<x_2 \Rightarrow a^{x_1}<a^{x_2}$. ■

Следствие. *Если $a^{x_1}=a^{x_2}$, то $x_1=x_2$.*

3^0. *Если $x\to+\infty$, то $a^x\to+\infty$.*

□ Возьмем произвольное число $P>0$. Нам надо доказать, что существует число $M>0$ такое, что из $x>M$ следует $a^x>P$.

Положим $a=1+h$. Так как $a>1$, то $h>0$. Воспользуемся неравенством Бернулли (см. § 11)

$$(1+h)^n>1+nh, \quad n\in\mathbb{N}.$$

Ясно, что при некотором достаточно большом $n=M$ выполняется неравенство $1+Mh>P$; тогда $a^M=(1+h)^M>P$. В силу возрастания показательной функции $y=a^x$ из $x>M$ следует $a^x>a^M>P$. Таким образом, $x>M\Rightarrow$ $\Rightarrow a^x>P$. ■

4^0. *Если $x\to-\infty$, то $a^x\to 0$.*

□ Положим $t=-x$. Согласно свойству 3^0, если $t\to+\infty$, то $a^t\to+\infty$. Далее, имеем $a^x=a^{-t}=1/a^t$. Если $x\to-\infty$, то $t\to+\infty$, $a^t\to+\infty$, $1/a^t\to 0$, т. е. $a^x\to 0$. ■

Мы доказали, что $\lim\limits_{x\to-\infty}a^x=0$. Это означает, что прямая $y=0$ (ось x) является горизонтальной асимптотой графика показательной функции $y=a^x$ при $x\to-\infty$.

5^0. *Функция непрерывна на множестве $\mathbb{R}$ всех действительных чисел.*

Доказательство этого свойства мы не приводим.

6^0. *Областью значений показательной функции является множество всех положительных чисел.*

Это сразу следует из свойств 3^0, 4^0 и 5^0, а также из теоремы о промежуточном значении непрерывной функции (см. теорему 2 из § 18).

Воспользовавшись полученными результатами, построим график функции $y=a^x$ при $a>1$. Пусть, например, $a=2$. Составим таблицу значений функции $y=2^x$:

x	-2	-1	0	1	2	3
y	1/4	1/2	1	2	4	8

С помощью найденных контрольных точек строим график функции $y = 2^x$ (рис. 64).

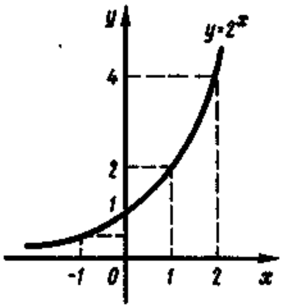


Рис. 64

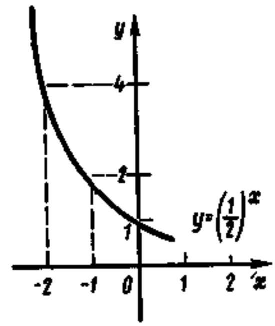


Рис. 65

Рассмотрим теперь случай, когда $0<a<1$. Пусть $b = 1/a$. Тогда $0<a<1 \Rightarrow b>1$ и из свойств функции $y = b^x$, $b>1$ легко получить интересующие нас свойства функции $y = a^x$. Приведем эти свойства.

1^0. *Область определения:* $(-\infty, +\infty)$.

2^0. *Функция убывающая.*

3^0. *Если* $x \to +\infty$, *то* $a^x \to 0$.

Это значит, что прямая $y = 0$ (ось x) является горизонтальной асимптотой графика функции $y = a^x$ при $x \to +\infty$.

4^0. *Если* $x \to -\infty$, *то* $a^x \to +\infty$.

5^0. *Функция непрерывна на множестве* **R** *всех действительных чисел.*

6^0. *Область значений:* $(0, +\infty)$.

Воспользовавшись полученными результатами, построим график функции $y = a^x$ при $0<a<1$. Пусть, например, $a = 1/2$. Составим таблицу значений функции $y = \left(\frac{1}{2}\right)^x$:

x	-2	-1	0	1	2	3
y	4	2	1	1/2	1/4	1/8

С помощью найденных контрольных точек строим график функции $y = \left(\frac{1}{2}\right)^x$ (рис. 65).

4. Показательные уравнения

Показательными называются уравнения вида $a^{f(x)} = a^{g(x)}$, где $a>0$ и $a \neq 1$, и уравнения, сводящиеся к уравнению указанного вида. В основе решения показательных уравнений лежит следующая теорема.

Теорема 1. *Показательное уравнение*

$$a^{f(x)}=a^{g(x)} \tag{3}$$

равносильно уравнению

$$f(x)=g(x). \tag{4}$$

□ Пусть x_0 — корень уравнения (3), т. е. $a^{f(x_0)}=a^{g(x_0)}$. Тогда $f(x_0)=g(x_0)$, т. е. x_0 — корень уравнения (4).

Обратно, если x_1 — корень уравнения (4), т. е. $f(x_1)=g(x_1)$, то $a^{f(x_1)}=a^{g(x_1)}$, а это означает, что x_1 — корень уравнения (3).

Таким образом, уравнения (3) и (4) равносильны. ■

Примеры. 1. Решить уравнение $2^{x^2-2x}=2^{3x-6}$.

○ Заданное уравнение равносильно уравнению $x^2-2x=3x-6$, корни которого $x_1=2$, $x_2=3$ являются корнями исходного уравнения. ●

2. Решить уравнение $\frac{(0{,}2)^{x-0{,}5}}{\sqrt{5}}=5\cdot 0{,}04^{x-1}$.

○ Приведем все степени к основанию $1/5$:

$$\left(\frac{1}{5}\right)^{x-0{,}5}\cdot\left(\frac{1}{5}\right)^{0{,}5}=\left(\frac{1}{5}\right)^{-1}\cdot\left(\left(\frac{1}{5}\right)^{2}\right)^{x-1}.$$

Далее, имеем $\left(\frac{1}{5}\right)^{x}=\left(\frac{1}{5}\right)^{2x-3}$. Последнее уравнение равносильно уравнению $x=2x-3$, откуда $x=3$. ●

3. Решить уравнение $4^x+2^{x+1}-24=0$.

○ Так как $4^x=(2^x)^2$ и $2^{x+1}=2\cdot 2^x$, то данное уравнение можно переписать следующим образом:

$$(2^x)^2+2\cdot 2^x-24=0.$$

Полагая $2^x=z$, получаем квадратное уравнение $z^2+2z-24=0$ с корнями $z_1=4$, $z_2=-6$. Поэтому задача сводится к решению совокупности уравнений $2^x=4$; $2^x=-6$. Из первого уравнения этой совокупности находим $x=2$. Второе уравнение совокупности не имеет решений, так как $2^x>0$ при любых значениях x. Итак, получаем ответ: $x=2$. ●

4. Решить уравнение $\left(\frac{3}{5}\right)^{x}+\frac{7}{5}=2^x$.

○ Это уравнение не решается ни одним из рассмотренных в предыдущих примерах приемов. Попробуем найти какое-либо решение уравнения методом подбора. В данном случае это сделать нетрудно: $x_1=1$. Разумеется, пока еще нельзя считать, что уравнение решено, так как оно может иметь и другие корни. Докажем, что других корней нет.

Пусть $x>1$. Тогда вследствие убывания функции $y=\left(\frac{3}{5}\right)^{x}$ имеем $\left(\frac{3}{5}\right)^{x}+\frac{7}{5}<2$, а вследствие возрастания функции $y=2^x$ имеем $2^x>2$. Значит, уравнение не имеет корней, бóльших чем 1.

Пусть $x<1$. Тогда в силу убывания функции $y=\left(\frac{3}{5}\right)^{x}$ имеем $\left(\frac{3}{5}\right)^{x}+\frac{7}{5}>2$, а в силу возрастания функции $y=2^x$ имеем $2^x<2$. Значит, уравнение не имеет корней, меньших чем 1. Итак, $x=1$ — единственный корень уравнения. ●

Вообще, в том случае, когда одна часть уравнения содержит убывающую функцию, а другая часть — возрастающую и уравнение имеет корень, то он единственный. Это наглядно подтверждает рис. 66.

Любое показательное неравенство сводится в конечном счете к неравенству вида

$$a^{f(x)}>a^{g(x)}, \tag{5}$$

где, как обычно, $a>0$ и $a\neq 1$. При переходе от неравенства (5) к неравенству, связывающему показатели $f(x)$ и $g(x)$, нужно учитывать, что показательная функция $y=a^x$ возрастает при $a>1$ и убывает при $0<x<1$. Значит, если $a>1$, то, сравнивая показатели $f(x)$ и $g(x)$, нужно сохранить знак неравенства (5), т. е. записать $f(x)>g(x)$. В случае же, когда $0<a<1$, сравнивая

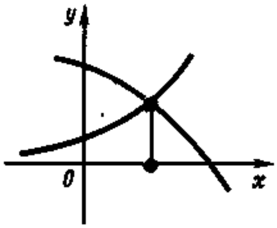


Рис. 66

Рис. 67

показатели $f(x)$ и $g(x)$, нужно использовать знак неравенства противоположного смысла, т. е. от неравенства (5) перейти к неравенству $f(x)<g(x)$.

Таким образом, справедливы следующие две теоремы.

Теорема 2. *Если* $a>1$, *то неравенство* (5) *равносильно неравенству* $f(x)>g(x)$.

Теорема 3. *Если* $0<a<1$, *то неравенство* (5) *равносильно неравенству* $f(x)<g(x)$.

Примеры. 1. Решить неравенство $(0{,}04)^{5x-x^2-8}<625$.

○ Замечая, что $625=(0{,}04)^{-2}$, перепишем заданное неравенство в виде $(0{,}04)^{5x-x^2-8}<(0{,}04)^{-2}$. Так как $0<0{,}04<1$, то по теореме 3 это неравенство равносильно неравенству $5x-x^2-8>-2$. Решив последнее неравенство, находим $2<x<3$. Итак, получаем ответ: $(2,\ 3)$. ●

2. Решить неравенство $\dfrac{3}{(0{,}5)^x-1}+\dfrac{1}{1-(0{,}5)^{x+1}}\geqslant 0$.

○ Положим $(0{,}5)^x=y$; тогда заданное неравенство примет вид $\dfrac{3}{y-1}+\dfrac{1}{1-0{,}5y}\geqslant$ $\geqslant 0$, откуда после преобразований получаем

$$\frac{y-4}{(y-1)\cdot(y-2)}\geqslant 0.$$

С помощью метода интервалов (рис. 67) находим $1<y<2$; $y\geqslant 4$. Таким образом, задача сводится к решению следующей совокупности неравенств: $1<(0{,}5)^x<2$; $(0{,}5)^x\geqslant 4$ или $(0{,}5)^0<(0{,}5)^x<(0{,}5)^{-1}$; $(0{,}5)^x\geqslant(0{,}5)^{-2}$.

Учитывая, что $0<0{,}5<1$, от последней совокупности переходим к совокупности $0>x>-1$; $x\leqslant -2$. Итак, получаем ответ: $(-\infty,\ -2]\cup(-1,\ 0)$. ●

3. Решить неравенство

$$3\cdot 2^{2x+1}-13\cdot 6^x+2\cdot 3^{2x+1}\leqslant 0.$$

○ Преобразуем неравенство к виду

$$6\cdot 2^{2x}-13\cdot 2^x\cdot 3^x+6\cdot 3^{2x}\leqslant 0.$$

Разделив обе части неравенства на одно и то же положительное выражение 3^{2x}, получим

$$6\cdot\frac{2^{2x}}{3^{2x}}-13\cdot\frac{2^x\cdot 3^x}{3^{2x}}+\frac{6\cdot 3^{2x}}{3^{2x}}\leqslant\frac{0}{3^{2x}}$$

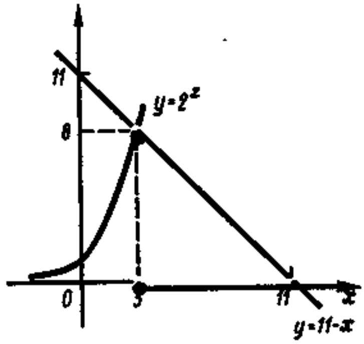


Рис. 68

и далее $6\cdot\left(\frac{2}{3}\right)^{2x}-13\cdot\left(\frac{2}{3}\right)^{x}+6\leqslant 0$. Полагая $y=\left(\frac{2}{3}\right)^{x}$, получим $6y^2-13y+6\leqslant$ $\leqslant 0$, откуда $\frac{2}{3}\leqslant y\leqslant\frac{3}{2}$.

Остается решить систему неравенств $\frac{2}{3}\leqslant\left(\frac{2}{3}\right)^{x}\leqslant\frac{3}{2}$, откуда находим $1\geqslant x\geqslant$ $\geqslant -1$. Итак, получаем ответ: $[-1, 1]$. ●

4. Решить неравенство $2^x\geqslant 11-x$.

○ Построим графики функций $y=2^x$ и $y=11-x$ (рис. 68). Они пересекаются в точке с абсциссой $x=3$. Из рисунка наглядно видно, что множеством решений неравенства $2^x\geqslant 11-x$ служит промежуток $[3, +\infty)$. ●

§ 21. Логарифмическая функция

1. Логарифмическая функция

Рассмотрим показательную функцию $y=a^x$, где $a>0$ и $a\neq 1$. В предыдущем параграфе мы отметили ряд свойств этой функции, из которых выделим сейчас следующие:

1^0. *Область определения*: $X=(-\infty, +\infty)$.

2^0. *Область значений*: $Y=(0, +\infty)$.

3^0. $y=a^x$ — *монотонная функция (возрастающая при* $a>1$ *и убывающая при* $0<a<1$).

4^0. $y=a^x$ — *непрерывная функция.*

По теореме 6 из § 18 существует обратная функция, определенная на $Y=(0, +\infty)$, имеющая область значений $X=(-\infty, +\infty)$, монотонная и непрерывная на Y.

Эта обратная функция называется *логарифмической функцией* и обозначается так: $x=\log_a y$ (читается «логарифм числа y по основанию a»). Поменяв x и y местами, получим $y=\log_a x$.

Согласно сказанному выше, логарифмическая функция $y=\log_a x$ обладает следующими свойствами:

1^0. *Область определения:* $(0, +\infty)$.

2^0. *Область значений:* $(-\infty, +\infty)$.

3^0. $y=\log_a x$ — *монотонная функция, причем возрастающая, если* $a>1$, *и убывающая, если* $0<a<1$.

4^0. $y=\log_a x$ — *непрерывная на* $(0, +\infty)$ *функция.*

Из монотонности логарифмической функции $y=\log_a x$ следует, в частности, что если $\log_a x_1=\log_a x_2$, то $x_1=x_2$.

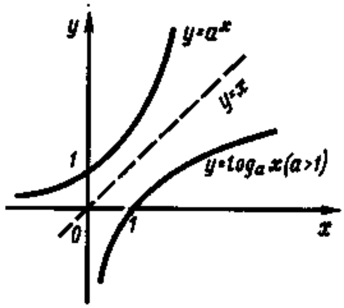


Рис. 69

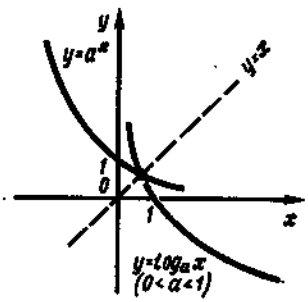


Рис. 70

График функции $y = \log_a x$ можно получить из графика функции $y = a^x$ с помощью преобразования симметрии последнего относительно прямой $y = x$. На рис. 69 построен график логарифмической функции для случая, когда $a > 1$, а на рис. 70 — для случая, когда $0 < a < 1$. Замечаем, что прямая $x = 0$ (ось y) является вертикальной асимптотой графика логарифмической функции. Кроме того, можно отметить, что график логарифмической функции пересекает ось x в одной точке $(1; 0)$, т. е. уравнение $\log_a x = 0$ имеет единственное решение $x = 1$.

2. Свойства логарифмов

В п. 1 было отмечено, что $y = \log_a x$ и $y = a^x$ — две взаимно обратные функции. Это означает, в частности, что равенства $y = \log_a x$ и $x = a^y$ выражают одну и ту же зависимость между x и y подобно тому, как это имеет место, например, для равенств $y = x^n$ и $x = \sqrt[n]{y}$. Итак, $y = \log_a x \Leftrightarrow x = a^y$.

Определение. *Логарифмом* положительного числа x по основанию a называется показатель степени, в которую нужно возвести число a, чтобы получить число x (здесь, как и всюду в этой главе, предполагается, что $a > 0$ и $a \neq 1$):

$$a^{\log_a x} = x. \qquad (1)$$

Например, $\log_3 81 = 4$, так как $3^4 = 81$; $\log_{10} 0{,}001 = -3$, так как $10^{-3} = 0{,}001$; $\log_{1/2}\sqrt{2} = -1/2$, так как $(1/2)^{-1/2} = 2^{1/2} = \sqrt{2}$.

Из определения логарифма вытекают следующие важные равенства:

$$\log_a 1 = 0, \ \log_a a = 1.$$

Первое следует из того, что $a^0 = 1$, а второе — из того, что $a^1 = a$.

Отметим некоторые свойства логарифмов.

1^0. *Если $x_1 > 0$ и $x_2 > 0$, то*

$$\log_a(x_1x_2) = \log_a x_1 + \log_a x_2$$

(логарифм произведения двух положительных чисел равен сумме логарифмов сомножителей).

□ Положим $y = \log_a(x_1x_2)$, $y_1 = \log_a x_1$, $y_2 = \log_a x_2$. Нам надо доказать, что $y = y_1 + y_2$.

Из равенства $y = \log_a(x_1x_2)$ следует, что $a^y = x_1x_2$. Аналогично из равенства $y_1 = \log_a x_1$ следует, что $a^{y_1} = x_1$, а из равенства $y_2 = \log_a x_2$ — что $a^{y_2} = x_2$. Тогда имеем

$$a^y = x_1x_2 = a^{y_1}a^{y_2} = a^{y_1+y_2}.$$

Но из равенства $a^y = a^{y_1+y_2}$ вытекает равенство $y = y_1 + y_2$, что и требовалось доказать. ■

Свойство 1^0 распространяется на произведение любого числа положительных сомножителей.

Дадим краткую запись приведенного доказательства:

$$\begin{array}{l|l}
\log_a(x_1x_2) = y, & \left.\begin{array}{l} a^y = x_1x_2, \\ a^{y_1} = x_1, \\ a^{y_2} = x_2 \end{array}\right\} \begin{array}{l} a^y = a^{y_1}\cdot a^{y_2} \Rightarrow a^y = a^{y_1+y_2} \Rightarrow \\ \Rightarrow y = y_1 + y_2 \end{array} \\
\log_a x_1 = y_1, & \\
\log_a x_2 = y_2. & \\
\hline
\text{Доказать: } y = y_1 + y_2 &
\end{array}$$

Аналогично доказываются остальные свойства.

2^0. *Если $x_1 > 0$ и $x_2 > 0$, то*

$$\log_a \frac{x_1}{x_2} = \log_a x_1 - \log_a x_2$$

(логарифм частного двух положительных чисел равен разности логарифмов делимого и делителя).

3^0. *Если $x > 0$, то*

$$\log_a x^k = k \log_a x, \text{ где } k \in \mathbf{R}$$

(логарифм степени положительного числа равен произведению показателя степени на логарифм основания степени).

4^0. *Если $x > 0$, то*

$$\log_a x = \frac{\log_b x}{\log_b a}$$

(формула перехода к новому основанию; здесь как обычно $b > 0$ и $b \neq 1$).

□

$$\begin{array}{l|l}
\log_a x = y, & \left.\begin{array}{l} a^y = x, \\ b^{y_1} = x, \\ b^{y_2} = a \end{array}\right\} a^y = x \Rightarrow (b^{y_2})^y = b^{y_1} \Rightarrow y_2 y = y_1 \Rightarrow y = \dfrac{y_1}{y_2}. \ ■ \\
\log_b x = y_1, & \\
\log_b a = y_2. & \\
\hline
\text{Доказать: } y = \dfrac{y_1}{y_2}. &
\end{array}$$

5^0. *Если $x > 0$, то*

$$\log_a x = \log_{a^k} x^k$$

(значение логарифма не изменится, если основание логарифма и логарифмируемое число возвести в одну и ту же степень).

□ Преобразуем правую часть к логарифму по основанию a. Используя свойство 4^0, получим

$$\log_{a^k} x^k = \frac{\log_a x^k}{\log_a a^k}.$$

Согласно свойству 3^0, имеем

$$\log_a x^k = k \log_a x, \ \log_a a^k = k \log_a a = k.$$

Следовательно,

$$\log_{a^k} x^k = \frac{k \log_a x}{k} = \log_a x. \ ■$$

Примеры. 1. Вычислить $\log_3 6$, если известно, что $\log_3 2 = a$.

○ Используя свойство 1°, имеем

$$\log_3 6 = \log_3(2\cdot 3) = \log_3 2 + \log_3 3 = a+1. \bullet$$

2. Вычислить $\log_2 \sqrt[3]{\frac{3}{4}}$, если $\log_2 3 = a$.

○ Используя свойства 2° и 3°, находим

$$\log_2 \sqrt[3]{\frac{3}{4}} = \log_2 \left(\frac{3}{4}\right)^{1/3} = \frac{1}{3}\log_2 \frac{3}{4} = \frac{1}{3}(\log_2 3 - \log_2 4) = \frac{1}{3}(a-2). \bullet$$

3. Вычислить $\log_5 6$, если $\log_2 3 = a$, $\log_2 10 = b$.

○ Перейдем в $\log_5 6$ к основанию 2. Воспользовавшись свойством 4°, а затем свойствами 1° и 2°, получим

$$\log_5 6 = \frac{\log_2 6}{\log_2 5} = \frac{\log_2(2\cdot 3)}{\log_2(10/2)} = \frac{\log_2 2 + \log_2 3}{\log_2 10 - \log_2 2} = \frac{1+a}{b-1}. \bullet$$

4. Вычислить $\log_{3\sqrt{2}} \sqrt[5]{32}$.

○ Согласно свойству 5°, основание логарифма и логарифмируемое число можно возвести в одну и ту же степень. Следовательно,

$$\log_{\sqrt[3]{2}} \sqrt[5]{32} = \log_{(\sqrt[3]{2})^3} (\sqrt[5]{32})^3 = \log_2 \sqrt{32} = \log_2 2^{5/2} = \frac{5}{2}. \bullet$$

5. Вычислить $49^{1-\frac{1}{4}\log_7 25}$

○ Имеем

$$49^{1-\frac{1}{4}\log_7 25} = (7^2)^{1-\frac{1}{4}\log_7 25} = 7^{2-\frac{1}{2}\log_7 25}$$

Преобразуем показатель полученной степени:

$$2-\frac{1}{2}\log_7 25 = 2\log_7 7 - \log_7 25^{1/2} = \log_7 7^2 - \log_7 5 = \log_7 \frac{49}{5}.$$

Таким образом, $7^{2-\frac{1}{2}\log_7 25} = 7^{\log_7 \frac{49}{5}}$. Согласно определению логарифма (1), получим

$$7^{\log_7 \frac{49}{5}} = \frac{49}{5} \bullet$$

3. Логарифмирование и потенцирование

Если некоторое выражение A составлено из положительных чисел с помощью операций умножения, деления и возведения в степень, то, используя свойства логарифмов, можно выразить $\log_a A$ через логарифмы входящих в выражение A чисел. Такое преобразование называется *логарифмированием*.

Пример 1. Прологарифмировать по основанию 5 выражение $\frac{125a^3b^2}{\sqrt{c}}$, где a, b, c — положительные числа.

○ Используя свойства 1°—3° логарифмов, получим

$$\log_5 \frac{125a^3b^2}{\sqrt{c}} = \log_5(125a^3b^2) - \log_5\sqrt{c} = \log_5 125 + \log_5 a^3 + \log_5 b^2 - \log_5 c^{1/2} =$$

$$= 3 + 3\log_5 a + 2\log_5 b - \frac{1}{2}\log_5 c. \bullet$$

Часто приходится решать обратную задачу: находить выражение, логарифм которого представлен через логарифмы некоторых чисел. Такое преобразование называется *потенцированием*.

Пример 2. Найти x, если

$$\log_3 x = 2\log_3 5 + \frac{1}{2}\log_3 8 - 3\log_3 10.$$

○ Имеем

$$2\log_3 5 + \frac{1}{2}\log_3 8 - 3\log_3 10 = \log_3 5^2 + \log_3 8^{1/2} - \log_3 10^3 = \log_3 \frac{25\cdot 2\sqrt{2}}{1000} = \log_3 \frac{\sqrt{2}}{20}.$$

Из равенства $\log_3 x = \log_3(\sqrt{2}/20)$ находим $x = \sqrt{2}/20$. ●

4. Дополнительные замечания о свойствах логарифмов

Выше мы рассмотрели свойства логарифмов, справедливые для положительных значений переменных, содержащихся под знаком логарифма. Однако при решении примеров нужно учитывать и возможность появления в ряде случаев отрицательных значений переменных, а тогда свойства логарифмов следует использовать с осторожностью.

Например, если $x_1 < 0$ и $x_2 < 0$, то нельзя написать

$$\log_a(x_1x_2) = \log_a x_1 + \log_a x_2,$$

так как правая часть такого «равенства» не имеет смысла (логарифм отрицательного числа не существует). Здесь можно рассуждать так: x_1 и x_2 — отрицательные числа; следовательно, $x_1x_2 > 0$. Но тогда $x_1x_2 = |x_1x_2| = |x_1|\cdot|x_2|$. Значит,

$$\log_a(x_1x_2) = \log_a|x_1x_2| = \log_a(|x_1|\cdot|x_2|).$$

Так как $|x_1| > 0$ и $|x_2| > 0$, то, применив свойство 1^0 из п. 2, получим

$$\log_a(|x_1|\cdot|x_2|) = \log_a|x_1|\cdot\log_a|x_2|.$$

Итак, если $x_1x_2 > 0$, то

$$\log_a(x_1x_2) = \log_a|x_1| + \log_a|x_2|$$

и аналогично

$$\log_a\frac{x_1}{x_2} = \log_a|x_1| - \log_a|x_2|$$

(нетрудно заметить, что оба равенства справедливы и в случае, когда $x_1 > 0$, $x_2 > 0$).

На случай отрицательного значения переменной x распространяется и свойство 3^0, когда k — четное число. Именно, справедливо следующее утверждение: если k — четное число, то

$$\log_a x^k = k\log_a|x|$$

для любого $x \neq 0$.

5. Десятичные логарифмы

На практике часто используются логарифмы по основанию 10, или *десятичные логарифмы*. Для десятичного логарифма $\log_{10} x$ принята сокращенная запись $\lg x$. В этом пункте мы выясним мотивы преимущественного использования для вычислений десятичных логарифмов.

Прежде всего введем понятие стандартного вида положительного числа. Любое положительное число можно представить в виде $a_1\cdot 10^n$, где $1 \leqslant a_1 < 10$, а $n \in \mathbf{Z}$.

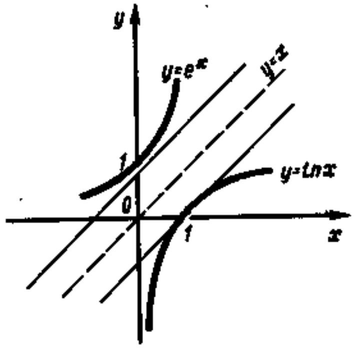


Рис. 71

Примеры. 1. Пусть $a=283$; тогда $a=2{,}83\cdot10^2$; здесь $a_1=2{,}83$, $1\leqslant a_1<10$ и $n=2$.

2. Пусть $a=3{,}76$; тогда $a=3{,}76\cdot10^0$; здесь $a_1=3{,}76$, $1\leqslant a_1<10$ и $n=0$.

3. Пусть $a=0{,}00075$; тогда $a=7{,}5\cdot10^{-4}$; здесь $a_1=7{,}5$, $1\leqslant a_1<10$ и $n=-4$.

Если положительное число a представлено в виде $a_1\cdot10^n$, где $1\leqslant a_1<10$, а $n\in\mathbf{Z}$, то говорят, что число a записано в *стандартном виде*. При этом показатель n называется *порядком* числа.

Пусть $\lg a$ — десятичный логарифм положительного числа a. Записав это число в стандартном виде и используя свойства логарифмов, находим

$$\lg a=\lg(a_1\cdot10^n)=\lg a_1+\lg 10^n=\lg a_1+n\lg 10.$$

Так как $\lg 10=\log_{10}10=1$, то

$$\lg a=n+\lg a_1; \qquad (2)$$

поскольку $1\leqslant a_1<10$, имеем $\lg 1\leqslant\lg a_1<\lg 10$, т. е. $0\leqslant\lg a_1<1$.

Из равенства (2) следует, что n есть наибольшее целое число, не превосходящее число $\lg a$; иными словами, n есть целая часть числа $\lg a$, т. е. $n=[\lg a]$, а $\lg a_1$ есть дробная часть числа $\lg a$, т. е. $\lg a_1=\{\lg a\}$. Целая часть числа $\lg a$, т. е. порядок числа a, называется *характеристикой* $\lg a$, а дробная часть числа $\lg a$ — его *мантиссой*.

Имеет место следующее утверждение: *если число $a>0$ умножить на 10^k, где $k\in\mathbf{Z}$, то мантисса логарифма не изменится, т. е. мантиссы чисел $\lg a$ и $\lg(a\cdot10^k)$ одинаковы.*

□ В самом деле, имеем

$$\lg(a\cdot10^k)=\lg a+\lg 10^k=\lg(a_1\cdot10^n)+k\lg 10=\lg a_1+\lg 10^n+k=n+k+\lg a_1.$$

Мантиссой числа $\lg(a\cdot10^k)$ является $\lg a_1$, т. е. то же число, которое служит мантиссой для $\lg a$. ■

6. Натуральные логарифмы

Особую роль в математическом анализе играют показательная и логарифмическая функции с основанием e (см. п. 6 § 13), т. е. функции $y=e^x$ и $y=\log_e x$. Последнюю записывают обычно в виде $y=\ln x$ ($\ln x$ называют *натуральным логарифмом* числа x). На рис. 71 изображены графики функций $y=e^x$ и $y=\ln x$.

Так как функции $y=e^x$ и $y=\ln x$ взаимно обратны, то справедливы равенства

$$e^{\ln x}=x,\ \ln e^x=x.$$

Логарифмическими называются уравнения вида

$$\log_a f(x) = \log_a g(x), \tag{3}$$

где $a>0$ и $a\neq 1$, и уравнения, сводящиеся к уравнению указанного вида.

Если по аналогии с показательными уравнениями вида $a^{f(x)} = a^{g(x)}$ перейти от уравнения (3) к уравнению $f(x) = g(x)$, то утверждать, что последнее равносильно уравнению (3), вообще говоря, нельзя. Дело в том, что уравнение $f(x) = g(x)$ может иметь своим корнем такое число x_0, при котором $f(x_0)\leqslant 0$ и $g(x_0)\leqslant 0$, а под знаком логарифма может находиться лишь положительное число. Это значит, что при переходе от уравнения (3) к уравнению $f(x) = g(x)$ нужно, найдя корни последнего уравнения, отобрать из них те, которые удовлетворяют системе неравенств

$$\begin{cases} f(x)>0, \\ g(x)>0. \end{cases}$$

Эта система задает область определения уравнения (3). Заметим, впрочем, что можно ограничиться одним из этих неравенств, так как другое вытекает из него и из уравнения $f(x) = g(x)$.

В результате приходим к следующей теореме.

Теорема 1. *Уравнение* $\log_a f(x) = \log_a g(x)$ *равносильно смешанной системе*

$$\begin{cases} f(x)>0, \\ g(x)>0, \\ f(x) = g(x). \end{cases}$$

Эта система, в свою очередь, равносильна каждой из следующих систем:

$$\begin{cases} f(x)>0, \\ f(x) = g(x) \end{cases} \text{ или } \begin{cases} g(x)>0, \\ f(x) = g(x). \end{cases}$$

Это замечание бывает полезным на практике.

При решении логарифмических уравнений используются различные свойства логарифмов. Рассмотрим, например, уравнение

$$\log_a f(x) + \log_a g(x) = \log_a h(x). \tag{4}$$

Оно приводится к виду

$$\log_a (f(x)g(x)) = \log_a h(x). \tag{5}$$

Однако уравнения (4) и (5) могут быть неравносильными. В самом деле, область определения выражения $\log_a f(x) + \log_a g(x)$ задается системой неравенств $\begin{cases} f(x)>0 \\ g(x)>0 \end{cases}$, тогда, как область определения выражения $\log_a (f(x)\,g(x))$ задается неравенством $f(x)\,g(x)>0$, которое, в свою очередь, равносильно совокупности систем неравенств

$$\begin{cases} f(x)>0, \\ g(x)>0; \end{cases} \quad \begin{cases} f(x)<0, \\ g(x)<0. \end{cases}$$

Таким образом, при переходе от уравнения (4) к уравнению (5) может произойти расширение области определения за счет решений последней системы неравенств, а значит, могут появиться посторонние корни. Поэтому, решив уравнение (5), следует из найденных его корней отобрать те, которые принадлежат области определения исходного уравнения (4), т. е. удовлетворяют системе неравенств

$$\begin{cases} f(x)>0, \\ g(x)>0, \\ h(x)>0, \end{cases}$$

либо сделать проверку найденных корней подстановкой в исходное уравнение (4).

Рассмотрим теперь уравнения вида

$$\log_{a(x)} f(x) = \log_{a(x)} g(x). \qquad (6)$$

Их решение основано на следующей теореме (приведем ее без доказательства).

Теорема 2. *Уравнение* (6) *равносильно смешанной системе*

$$\begin{cases} f(x) = g(x), \\ f(x) > 0, \\ g(x) > 0, \\ a(x) > 0, \\ a(x) \neq 1. \end{cases}$$

Иными словами, корнями уравнения (6) являются те и только те корни уравнения $f(x) = g(x)$, которые одновременно удовлетворяют следующим условиям: $f(x)>0$, $g(x)>0$, $a(x)>0$, $a(x)\neq 1$ [этими условиями задается область определения уравнения (6)]. Заметим, что одно из неравенств $f(x)>0$ или $g(x)>0$ можно опустить.

Примеры. 1. Решить уравнение $\ln(x^2-3x-5)=\ln(7-2x)$.

○Данное уравнение равносильно следующей смешанной системе:

$$\begin{cases} 7-2x>0, \\ x^2-3x-5=7-2x. \end{cases}$$

Решив уравнение системы, получим $x_1=4$, $x_2=-3$. Из этих двух значений неравенству системы удовлетворяет лишь $x=-3$. Поэтому корнем данного уравнения является $x=-3$. ●

2. Решить уравнение $\lg(x+4)+\lg(2x+3)=\lg(1-2x)$.

○Преобразуем уравнение к виду

$$\lg((x+4)(2x+3))=\lg(1-2x), \text{ или } (x+4)(2x+3)=1-2x.$$

Из последнего уравнения находим $x_1=-1$, $x_2=-5{,}5$. Область определения данного уравнения задается системой неравенств

$$\begin{cases} x+4>0, \\ 2x+3>0, \\ 1-2x>0. \end{cases}$$

Подставив поочередно найденные значения $x_1=-1$; $x_2=-5{,}5$ в систему, убеждаемся, что -1 удовлетворяет этой системе, а $-5{,}5$ не удовлетворяет. Таким образом, значение $x=-1$ является единственным корнем заданного уравнения. ●

Замечание. В рассмотренном примере система, задающая область определения, сравнительно несложная. Поэтому можно было, решив эту систему, найти область определения исходного уравнения: $-1{,}5<x<0{,}5$. Так как $x_1=-1$ принадлежит интервалу $(-1{,}5;\ 0{,}5)$, то $x_1=-1$ является корнем исходного уравнения. Так как $x_2=-5{,}5$ не принадлежит интервалу $(-1{,}5;\ 0{,}5)$, то $x_2=-5{,}5$ — посторонний корень.

3. Решить уравнение $\log_{x+4}(x^2-1)=\log_{x+4}(5-x)$.

○ Согласно теореме 3, данное уравнение равносильно системе:

$$\begin{cases} x^2-1=5-x, \\ x^2-1>0, \\ 5-x>0, \\ x+4>0, \\ x+4\neq 1. \end{cases}$$

Решив входящее в систему уравнение, получим $x_1=2$, $x_2=-3$. Из этих двух значений всем остальным условиям системы удовлетворяет лишь $x=2$. Итак, получаем ответ: $x=2$. ●

4. Решить уравнение $\lg^2 x+\lg x+1=\frac{7}{\lg(x/10)}$.

○ Так как $\lg(x/10)=\lg x-1$, то заданное уравнение можно переписать следующим образом:

$$\lg^2 x+\lg x+1=\frac{7}{\lg x-1}.$$

Полагая $\lg x=z$, получаем уравнение

$$z^2+z+1=\frac{7}{z-1},$$

откуда $z=2$. Решая теперь уравнение $\lg x=2$, находим $x=100$ — единственный корень данного уравнения. ●

5. Решить уравнение $\ln^2 x^3-\ln(x^{10}/e)=0$.

○ Имеем

$$\ln^2 x^3=(\ln x^3)^2=(3\ln x)^2=9\ln^2 x,$$

$$\ln(x^{10}/e)=\ln x^{10}-\ln e=10\ln x-1.$$

Учитывая выполненные преобразования, перепишем заданное уравнение в виде

$$9\ln^2 x-(10\ln x-1)=0.$$

Полагая $\ln x=y$, получим

$$9y^2-10y+1=0,$$

откуда $y_1=1$, $y_2=1/9$. Из уравнения $\ln x=1$ получаем $x_1=e$, а из уравнения $\ln x=1/9$ находим $x_2=e^{1/9}$. ●

6. Решить уравнение $\lg(20-x)=\lg^3 x$.

○ Это уравнение не удается решить рассмотренными в предыдущих примерах приемами. Найдем какой-либо корень уравнения методом подбора. В данном случае получаем $x_1=10$.

Однако считать, что уравнение уже решено, нельзя: оно может иметь и другие решения. Докажем, что других корней нет. Разумеется, корни уравнения следует искать в его области определения, т. е. в интервале $(0, 20)$. Замечаем, что на этом интервале функция $y=\lg(20-x)$ убывает, а функция $y=\lg^3 x$ возрастает; тогда если уравнение имеет корень, то только один (см. п. 4 § 20). Итак, $x=10$ — единственный корень уравнения. ●

8. Показательно-логарифмические уравнения

В настоящем пункте мы рассмотрим уравнения, которые можно считать и показательными, и логарифмическими.

Примеры. 1. Решить уравнение

$$x^{1-\lg x}=0{,}01. \tag{7}$$

○Область определения уравнения $x>0$. В этой области выражения, входящие в обе части уравнения (7), принимают только положительные значения; поэтому, прологарифмировав обе его части по основанию 10, получим уравнение

$$\lg x^{1-\lg x}=\lg 0{,}01,$$

равносильное уравнению (7). Далее, имеем

$$(1-\lg x)\lg x=-2.$$

Полагая $u=\lg x$, приходим к уравнению $(1-u)u=-2$, откуда $u_1=-1$, $u_2=2$. Решая совокупность уравнений $\lg x=-1$, $\lg x=2$, находим $x_1=0{,}1$; $x_2=100$ — корни уравнения (7). ●

Здесь применен *метод логарифмирования*, заключающийся в переходе от уравнения $f(x)=g(x)$ к уравнению $\log_a f(x)=\log_a g(x)$.

2. Решить уравнение $\log_x(3x^{\log_5 x}+4)=2\log_5 x$.

○Воспользовавшись определением логарифма, преобразуем уравнение к виду

$$x^{2\log_5 x}=3x^{\log_5 x}+4.$$

Полагая $u=x^{\log_5 x}$, получим уравнение $u^2-3u-4=0$, корни которого $u_1=-1$, $u_2=4$.

Теперь задача сводится к решению следующей совокупности уравнений: $x^{\log_5 x}=-1$; $x^{\log_5 x}=4$. Так как $x^{\log_5 x}>0$, а $-1<0$, то первое уравнение совокупности не имеет решения. Прологарифмировав обе части второго уравнения совокупности по основанию 5, получим

$$\log_5^2 x=\log_5 4,\ \text{т. е.}\ \log_5 x=\pm\sqrt{\log_5 4},$$

откуда $x_{1,2}=5^{\pm\sqrt{\log_5 4}}$. ●

3. Решить уравнение

$$\log_5(5^{\frac{1}{x}}+125)=\log_5 6+1+\frac{1}{2x}. \qquad (8)$$

○Сначала будем рассматривать данное уравнение как логарифмическое. Так как $1+\frac{1}{2x}=\log_5 5^{1+\frac{1}{2x}}$, то уравнение (8) запишем в виде

$$\log_5(5^{\frac{1}{x}}+125)=\log_5 6+\log_5 5^{1+\frac{1}{2x}}.$$

Далее, имеем

$$\log_5(5^{\frac{1}{x}}+125)=\log_5(6\cdot5\cdot5^{\frac{1}{2x}}),\ 5^{\frac{1}{x}}+125=30\cdot5^{\frac{1}{2x}}.$$

Получили показательное уравнение, которое можно решить методом введения новой переменной. Полагая $u=5^{1/(2x)}$, приходим к уравнению $u^2-30u+125=0$, корни которого $u_1=5$, $u_2=25$.

Теперь задача сводится к решению совокупности двух уравнений: $5^{1/(2x)}=5$; $5^{1/(2x)}=25$. Из первого уравнения получаем $1/(2x)=1$, откуда $x_1=1/2$; из второго уравнения получаем $1/(2x)=2$, откуда $x_2=1/4$. Итак, уравнение (8) имеет два корня: $x_1=1/2$ и $x_2=1/4$. ●

9. Логарифмические неравенства

Любое логарифмическое неравенство сводится в конечном счете к неравенству вида

$$\log_a f(x)>\log_a g(x), \qquad (9)$$

где, как обычно, $a>0$ и $a\neq1$. Переходя от неравенства (9) к неравенству, связывающему $f(x)$ и $g(x)$, нужно учитывать, что логарифмическая функция возрастает при $a>1$ и убывает при $0<a<1$. Значит, в случае $a>1$ от неравенства (9) следует перейти к неравенству того же смысла $f(x)>g(x)$, а в случае $0<a<1$ от неравенства (9) — к неравенству противоположного смысла $f(x)<g(x)$. Разумеется, и в том, и в другом случае следует учесть, что должны выполняться неравенства $f(x)>0$ и $g(x)>0$.

Таким образом, справедливы следующие теоремы.

Теорема 3. *Если $a>1$, то неравенство* (9) *равносильно системе неравенств*

$$\begin{cases} f(x)>0, \\ g(x)>0, \\ f(x)>g(x). \end{cases} \qquad (10)$$

Теорема 4. *Если $0<a<1$, то неравенство* (9) *равносильно системе*

$$\begin{cases} f(x)>0, \\ g(x)>0, \\ f(x)<g(x). \end{cases} \qquad (11)$$

Замечания. 1. В системе (10) неравенство $f(x)>0$ является следствием двух других неравенств и его можно опустить: аналогично можно опустить неравенство $g(x)>0$ в системе (11).

2. Обе теоремы распространяются на случай любого другого знака неравенства, отличного от знака $>$ в неравенстве (9). Подчеркнем, что в системе (10) знак последнего неравенства совпадает со знаком исходного логарифмического неравенства, а в системе (11) противоположен ему.

Примеры. 1. Решить неравенство $\log_{1/2}(2x^2-4x-14)\leqslant-1$.

○ Поскольку $-1=\log_{1/2}2$, заданное неравенство можно переписать так:

$$\log_{1/2}(2x^2-4x-14)\leqslant\log_{1/2}2.$$

Здесь основание логарифмов $a=1/2$, т. е. $0<a<1$, и, следовательно, по теореме 4 данное неравенство равносильно следующей системе неравенств:

$$\begin{cases} 2x^2-4x-14>0, \\ 2x^2-4x-14\geqslant 2. \end{cases}$$

Ясно, что полученная система равносильна неравенству $2x^2-4x-14\geqslant2$, решая которое, находим $x\leqslant-2$; $x\geqslant4$. ●

2. Решить неравенство

$$\ln(x+27)-\ln(16-2x)>\ln x. \qquad (12)$$

○ Имеем

$$\ln(x+27)>\ln x+\ln(16-2x),$$

т. е.

$$\ln(x+27)>\ln(x(16-2x)). \qquad (13)$$

Так как основание натурального логарифма удовлетворяет неравенству $e>1$, то, опуская знаки логарифмов в неравенстве (13), следует сохранить знак неравенства. Однако при этом нужно учесть, что должны выполняться неравенства $x+27>0$, $16-2x>0$ и $x>0$, задающие область определения исходного неравенства (12).

Таким образом, приходим к следующей системе неравенств:

$$\begin{cases} x+27>0, \\ 16-2x>0, \\ x>0, \\ x+27>x(16-2x). \end{cases}$$

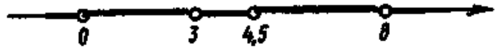


Рис. 72

Решая эту систему, получаем

$$\begin{cases} x>-27, \\ x<8, \\ x>0, \\ 2x^2-15x+27>0, \end{cases} \quad \text{т. е.} \begin{cases} 0<x<8, \\ 2(x-3)(x-4{,}5)>0. \end{cases}$$

Следовательно, $0<x<3$; $4{,}5<x<8$ (рис. 72).●

3. Решить неравенство $\log_{x-2}(2x-3)>\log_{x-2}(24-6x)$.

○Здесь надо рассмотреть два случая: 1) $x-2>1$; 2) $0<x-2<1$. В первом случае к данному неравенству применима теорема 3, а во втором — теорема 4. Таким образом, задача сводится к решению следующей совокупности двух систем неравенств:

$$\begin{cases} x-2>1, \\ 2x-3>0, \\ 24-6x>0, \\ 2x-3>24-6x; \end{cases} \quad \begin{cases} 0<x-2<1, \\ 2x-3>0, \\ 24-6x>0, \\ 2x-3<24-6x. \end{cases}$$

Из первой системы следует, что $27/8<x<4$, а из второй — что $2<x<3$.●

4. Решить неравенство $\log_2^2(x-1)^2-\log_{0,5}(x-1)>5$.

○ На основании свойств логарифмов имеем $\log_2(x-1)^2=2\log_2|x-1|=2\log_2(x-1)$ (здесь $|x-1|=x-1$, поскольку из данного неравенства следует, что $x-1>0$);

$$\log_{0,5}(x-1)=\log_{0,5^{-1}}(x-1)^{-1}=-\log_2(x-1).$$

Тогда данное неравенство можно переписать так:

$$4\log_2^2(x-1)+\log_2(x-1)>5.$$

Полагая $\log_2(x-1)=y$, получим $4y^2+y-5>0$, откуда $y<-5/4$; $y>1$.

Теперь задача сводится к решению совокупности логарифмических неравенств $\log_2(x-1)<-5/4$; $\log_2(x-1)>1$, или

$$\log_2(x-1)<\log_2 2^{-5/4};\ \log_2(x-1)>\log_2 2.$$

Из первого неравенства этой совокупности получим $0<x-1<2^{-5/4}$ и, следовательно, $1<x<1+2^{-5/4}$. Из второго неравенства совокупности имеем $x-1>2$, т. е. $x>3$. Таким образом, получаем ответ: $(1,\ 1+2^{-5/4})\cup(3,\ +\infty)$.●

§ 22. Тригонометрические функции

В школьном курсе математики изучались тригонометрические выражения и выполнялись их преобразования, в которых использовались тригонометрические формулы, например, такие, как

$$\sin^2\alpha+\cos^2\alpha=1,\ \cos 2\alpha=\cos^2\alpha-\sin^2\alpha,$$

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta,$$

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta,$$

$$\sin\alpha-\sin\beta=2\sin\frac{\alpha-\beta}{2}\cos\frac{\alpha+\beta}{2}$$

и т. д. В настоящем параграфе мы рассмотрим свойства тригонометрических функций и построим их графики.

1. Числовая окружность

Возьмем окружность единичного радиуса (рис. 73) и примем за начало отсчета на этой окружности точку A — правый конец горизонтального диаметра. В качестве положительного направления движения по окружности примем движение против часовой стрелки, а в качестве отрицательного направления — движение по часовой стрелке.

Каждой точке M окружности соответствует бесконечное множество дуг, начинающихся в точке A и кончающихся в точке M. Одной из них является кратчайшая дуга, соединяющая эти две точки, а все остальные получаются из кратчайшей дуги прибавлением или вычитанием целого числа полных оборотов (полных обходов окружности).

Тем самым каждой точке M окружности поставлено в соответствие бесконечное множество чисел — это множество состоит из величин всех дуг, начинающихся в точке A и кончающихся в точке M (при этом величины дуг берутся со знаком плюс или минус в зависимости от того, происходит ли движение от точки A к точке M против часовой стрелки или по часовой стрелке). Так как длина окружности единичного радиуса равна 2π, то величины всех дуг, оканчивающихся в данной точке M окружности, отличаются друг от друга на целое кратное 2π. Иными словами, общий вид этих величин таков: $x+2\pi k$, где k — любое целое число, т. е. $k \in \mathbf{Z}$, а x — величина кратчайшей дуги, соединяющей A и M.

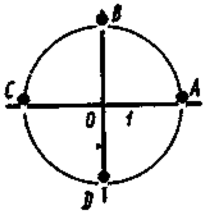
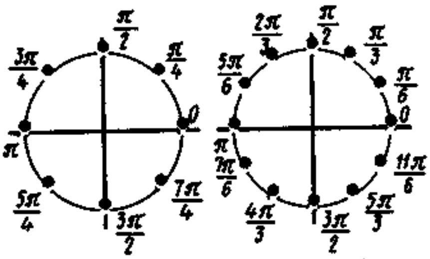


Рис. 73



а) б)

Рис. 74

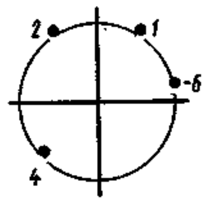


Рис. 75

Окружность единичного радиуса с выбранным началом отсчета и направлением обхода будем называть *числовой окружностью*. Каждому действительному числу x соответствует точка $M(x)$ числовой окружности такая, что величина дуги AM равна x, а каждая точка M окружности соответствует бесконечному множеству чисел вида $x+2\pi k$, $k \in \mathbf{Z}$, где x — величина одной из дуг, соединяющих точки A и M.

Вертикальный и горизонтальный диаметры разбивают числовую окружность на четыре части — это дуги AB, BC, CD и DA на рис. 73, которые называются соответственно I, II, III и IV *четвертями* окружности.

На рис. 74, *а* отмечены точки числовой окружности, соответствующие числам вида $k\pi/2$, $k\pi/4$, а на рис. 74, *б* — точки, соответствующие числам вида $k\pi/3$, $k\pi/6$ ($k \in \mathbf{Z}$).

Пример. На числовой окружности найти точки, соответствующие числам 1, 2, 4, —6.

○ Имеем $\smile{AB} = \pi/2 \approx 3{,}14/2 \approx 1{,}57$. Значит, числу 1 соответствует точка I четверти, лежащая несколько выше середины K дуги AB — точка x_1 на рис. 75. Число 2 больше числа $\pi/2$ примерно на 0,43, что составляет около четверти длины дуги BC. Так получается точка x_2 на рис. 75, соответствующая числу 2. Числу 4 соответствует точка x_3, лежащая в III четверти (рис. 75).

Остается найти точку, соответствующую числу -6. Число 6 отличается от длины окружности 2π примерно на 0,28. Исходя из точки A, опишем по окружности в отрицательном направлении путь длины 6, т. е. путь, на 0,28 меньший полного оборота. Конец этого пути обозначен на рис. 75 точкой x_4, соответствующей числу -6. ●

2. Определение тригонометрических функций

Возьмем числовую окружность и расположим ее так, чтобы центр окружности совпал с началом O прямоугольной системы координат XY (рис. 76). Пусть задано некоторое число x. Найдем соответствующую ему точку $M(x)$ числовой окружности.

Определение. Ордината точки $M(x)$ называется *синусом* числа x и обозначается $\sin x$, а абсцисса этой точки называется *косинусом* числа x и обозначается $\cos x$ (рис. 76). Отношение $\frac{\sin x}{\cos x}$ называется *тангенсом* числа x и обозначается $\operatorname{tg} x$; отношение $\frac{\cos x}{\sin x}$ называется *котангенсом* числа x и обозначается $\operatorname{ctg} x$.

Мы сопоставили каждому числу x значения $\sin x$, $\cos x$, $\operatorname{tg} x$ и $\operatorname{ctg} x$. Этим определены четыре числовые функции $y=\sin x$, $y=\cos x$, $y=\operatorname{tg} x$, $y=\operatorname{ctg} x$. Они называются *тригонометрическими функциями*.

Функции $y=\sin x$ и $y=\cos x$ определены при любых значениях x. Функция $y=\operatorname{tg} x$ определена при таких значениях x, что $\cos x \neq 0$. Функция $y=\operatorname{ctg} x$ определена при тех значениях x, для которых $\sin x \neq 0$.

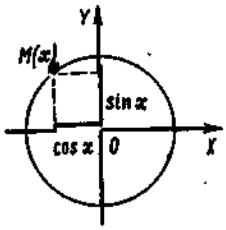


Рис. 76

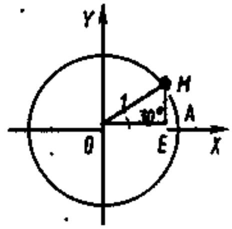


Рис. 77

Из определения следует, что $\cos 0 = 1$, $\sin 0 = 0$, $\operatorname{tg} 0 = \frac{\sin 0}{\cos 0} = \frac{0}{1} = 0$ ($\operatorname{ctg} 0$ не определен, так как $\sin 0 = 0$); $\cos(\pi/2) = 0$, $\sin(\pi/2) = 1$, $\operatorname{ctg}(\pi/2) = \frac{\cos(\pi/2)}{\sin(\pi/2)} = \frac{0}{1} = 0$ ($\operatorname{tg}(\pi/2)$ не определен, так как $\cos(\pi/2) = 0$); $\cos\pi = -1$, $\sin\pi = 0$, $\operatorname{tg}\pi = 0$ ($\operatorname{ctg}\pi$ не определен); $\cos(3\pi/2) = 0$, $\sin(3\pi/2) = -1$, $\operatorname{ctg}(3\pi/2) = 0$ ($\operatorname{tg}(3\pi/2)$ не определен).

Вычислим теперь $\sin(\pi/6)$. Пусть M — точка числовой окружности, соответствующая числу $\pi/6$ (рис. 77). Тогда $AOM = 30°$ и из треугольника EOM получаем $ME = 1/2$. Но по определению синуса $ME = \sin(\pi/6)$. Значит, $\sin(\pi/6) = 1/2$. Далее, имеем $\cos(\pi/6) = OE = \sqrt{OM^2 - ME^2} = \sqrt{1-1/4} = \sqrt{3}/2$. Тогда $\operatorname{tg}(\pi/6) = (1/2):(\sqrt{3}/2) = 1/\sqrt{3} = \sqrt{3}/3$; $\operatorname{ctg}(\pi/6) = \sqrt{3}$.

Точно так же получаем, что $\sin(\pi/3) = \sqrt{3}/2$, $\cos(\pi/3) = 1/2$, $\operatorname{tg}(\pi/3) = \sqrt{3}$, $\operatorname{ctg}(\pi/3) = \sqrt{3}/3$; $\sin(\pi/4) = \sqrt{2}/2$, $\cos(\pi/4) = \sqrt{2}/2$; $\operatorname{tg}(\pi/4) = 1$, $\operatorname{ctg}(\pi/4) = 1$.

Оси координат делят числовую окружность на четыре четверти. В каждой из четвертей координаты точек окружности сохраняют один и тот же знак. Поэтому, зная, в какой четверти числовой окружности лежит точка $M(x)$, мы можем определить знаки тригонометрических функций; они указаны в следующей таблице:

x	Четверть	$\sin x$	$\cos x$	$\operatorname{tg} x$	$\operatorname{ctg} x$
$0<x<\pi/2$	I	+	+	+	+
$\pi/2<x<\pi$	II	+	−	−	−
$\pi<x<3\pi/2$	III	−	−	+	+
$3\pi/2<x<2\pi$	IV	−	+	−	−

Рассмотрим примеры, показывающие, как с помощью числовой окружности решаются уравнения вида $\sin x=m$, $\cos x=m$, где $|m|\leqslant 1$.

Решим уравнение $\sin x=0$. Это означает, что нам нужно найти множество всех таких чисел x, которым на числовой окружности соответствуют точки с ординатой, равной нулю. На числовой окружности имеются две точки, ординаты которых равны нулю. Это точки A и C (см. рис. 75). Точка A соответствует числам 0, $\pm 2\pi$, $\pm 4\pi$, ..., а точка C — числам $\pm\pi$, $\pm 3\pi$, $\pm 5\pi$, ... Значит, точки A и C соответствуют числам вида πn, где n — любое целое число*, т. е. $n\in\mathbf{Z}$. Итак, $x=\pi n$ — корни уравнения $\sin x=0$.

Рассуждая аналогично, получаем, что корнями уравнения $\sin x=1$ являются числа вида $x=\frac{\pi}{2}+2\pi n$, а корнями уравнения $\sin x=-1$ — числа вида $x=-\frac{\pi}{2}+2\pi n$.

Решим уравнение $\cos x=0$. На числовой окружности имеются две точки, абсциссы которых равны нулю: это точки B и D (см. рис. 75). При этом точка B соответствует числам вида $\frac{\pi}{2}+2\pi k$, а точка D — числам вида $\frac{3\pi}{2}+2\pi k$ или $\frac{\pi}{2}+\pi(2k+1)$. Следовательно, точки B и D соответствуют числам вида $\frac{\pi}{2}+\pi n$. Итак, $x=\frac{\pi}{2}+\pi n$ — корни уравнения $\cos x=0$.

Аналогично получаем, что корнями уравнения $\cos x=1$ являются числа вида $x=2\pi n$, а корнями уравнения $\cos x=-1$ — числа вида $x=\pi+2\pi n$.

Теорема 1. *Если* $0<x<\pi/2$, *то* $\sin x<x<\operatorname{tg} x$.

□ Пусть M — точка числовой окружности, соответствующая числу x. Восставим из точки A перпендикуляр до пересечения в точке K с лучом OM и сравним площади трех фигур: треугольника OAM, сектора OAM и треугольника OAK (рис. 78). Имеем

$$S_{\Delta OAM}<S_{\text{сект. } OAM}<S_{\Delta OAK}, \tag{1}$$

где

$$S_{\Delta OAM}=\frac{1}{2}OA\cdot MH=\frac{1}{2}\cdot 1\cdot\sin x=\frac{1}{2}\sin x,$$

$$S_{\text{сект. } OAM}=\frac{1}{2}OA^2\cdot x=\frac{1}{2}x,\ S_{\Delta OAK}=\frac{1}{2}OA\cdot AK=\frac{1}{2}AK.$$

* В дальнейшем всюду, записывая ответ в тригонометрических уравнениях и неравенствах, мы подразумеваем, что параметр n, k, m, ... принимает любые целые значения, не оговаривая этого (для краткости) каждый раз особо.

Из подобия треугольников OHM и OAK получаем

$$\frac{MH}{OH}=\frac{AK}{OA}, \text{ т. е. } \frac{\sin x}{\cos x}=\frac{AK}{1}.$$

Значит, $AK=\operatorname{tg} x$ и потому $S_{\Delta OAK}=\frac{1}{2}AK=\frac{1}{2}\operatorname{tg} x$. Воспользовавшись полученными результатами, перепишем неравенство (1) в виде

$$\frac{1}{2}\sin x<\frac{1}{2}x<\frac{1}{2}\operatorname{tg} x, \text{ т. е. } \sin x<x<\operatorname{tg} x. \blacksquare$$

3. Свойства и графики функций $y=\sin x$, $y=\cos x$

Отметим свойства функции $y=\sin x$.

1^0. *Область определения:* $(-\infty, +\infty)$.

2^0. *Функция периодическая с основным периодом* 2π.

□ Если два числа отличаются на целое кратное 2π, то им соответствует одна и та же точка $M(x)$ числовой окружности, а потому $\sin(x\pm2\pi)=\sin x$, т. е. функция $y=\sin x$ является периодической и 2π — один из ее периодов. Легко видеть, что это основной период. В самом деле, ордината точки B на

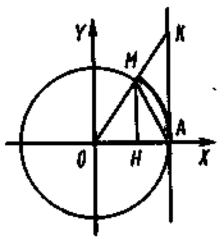


Рис. 78

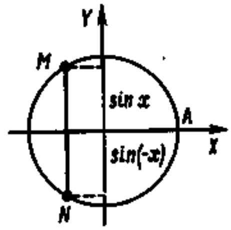


Рис. 79

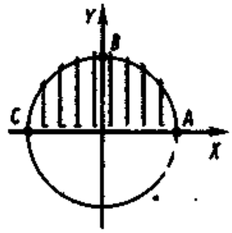


Рис. 80

рис. 75 равна 1 и, значит, $\sin(\pi/2)=1$. Если двигаться по окружности от точки B в положительном направлении, то в очередной раз ордината движущейся точки станет равной единице только тогда, когда эта точка совпадет с B, т. е. при значении $\pi/2+2\pi$. Следовательно, период функции $y=\sin x$ не может быть меньше чем 2π, а поэтому 2π есть ее основной период. ■

Так как 2π — основной период функции $y=\sin x$, то для построения графика этой функции во всей области определения достаточно построить его ветвь на отрезке $[-\pi, \pi]$.

3^0. $y=\sin x$ — *нечетная функция.*

□ Точки $M(x)$ и $N(-x)$ числовой окружности симметричны относительно оси X. Поэтому они имеют ординаты, равные по модулю, но противоположные по знаку. Таким образом, $\sin(-x)=-\sin x$ (рис. 79). ■

График нечетной функции симметричен относительно начала координат, поэтому для построения графика функции $y=\sin x$ на отрезке $[-\pi, \pi]$ достаточно сначала построить ветвь графика на отрезке $[0, \pi]$.

4^0. *Функция возрастает на отрезке* $[0, \pi/2]$ *и убывает на отрезке* $[\pi/2, \pi]$.

□ При движении по числовой окружности от A к B против часовой стрелки ордината увеличивается от 0 до 1, а при движении от B к C — уменьшается от 1 до 0 (рис. 80). Значит, на отрезке $[0, \pi/2]$ функция $y=\sin x$ возрастает, а на отрезке $[\pi/2, \pi]$ убывает. ■

5^0. *Область значений функции — отрезок* $[-1, 1]$.

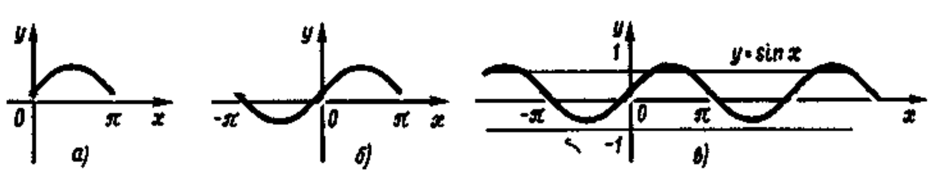


Рис. 81

Теорема 2. *Для любых x справедливо неравенство* $|\sin x| \leqslant |x|$.

□ Если $0 < x < \pi/2$, то $\sin x < x$ (согласно теореме 1). Но в интервале $(0, \pi/2)$ имеем $|\sin x| = \sin x$, а $|x| = x$. Значит, $|\sin x| < |x|$.

Если $-\pi/2 < x < 0$, то $0 < -x < \pi/2$, а потому $\sin(-x) < -x$. Используя нечетность функции $y = \sin x$, получаем $-\sin x < -x$. Но в интервале $(-\pi/2, 0)$ имеем $\sin x < 0$, $x < 0$, а следовательно, $|\sin x| = -\sin x$, $|x| = -x$. Поэтому $|\sin x| < |x|$.

Если $x = 0$, то $|\sin x| = |x| = 0$.

Пусть, наконец, $|x| \geqslant \pi/2$. Тогда $|x| > 1$, а $|\sin x| \leqslant 1$. Таким образом, $|\sin x| < |x|$.

Обобщая все сказанное выше, получаем, что неравенство $|\sin x| \leqslant |x|$ истинно для любых значений переменной x. ■

Теорема 3. *Функция* $y = \sin x$ *непрерывна на* $(-\infty, +\infty)$.

□ Возьмем произвольную точку a и докажем, что в ней функция $y = \sin x$ непрерывна. Воспользуемся определением «на языке бесконечно малых» (см. определение 2 из § 17). Для этого рассмотрим приращение функции $\Delta f = \sin x - \sin a$. Имеем

$$|\Delta f| = \left| 2\sin\frac{x-a}{2}\cos\frac{x+a}{2} \right| = 2\left| \sin\frac{x-a}{2} \right| \cdot \left| \cos\frac{x+a}{2} \right|.$$

Так как $\left|\cos\frac{x+a}{2}\right| \leqslant 1$ и $\left|\sin\frac{x-a}{2}\right| \leqslant \left|\frac{x-a}{2}\right|$ (согласно теореме 2), то $|\Delta f| \leqslant 2\left|\frac{x-a}{2}\right| \cdot 1 = |x-a|$. При $x \to a$ функция $x-a$ является бесконечно малой. Но $|\Delta f| \leqslant |x-a|$, а потому и Δf — бесконечно малая при $x \to a$ (см. теорему 2 § 16), что и означает непрерывность функции в точке a. ■

Используя результаты проведенного исследования и взяв контрольные точки $(0; 0)$, $(\pi/6; 1/2)$, $(\pi/2; 1)$, $(\pi; 0)$, сначала построим график функции $y = \sin x$ на отрезке $[0, \pi]$ (рис. 81, *а*). Далее, так как функция $y = \sin x$ — нечетная, то отобразив построенный график симметрично относительно начала координат, получим график функции на отрезке $[-\pi, \pi]$ (рис. 81, *б*). Наконец, воспользовавшись периодичностью функции $y = \sin x$, продолжим график на всю область определения (рис. 81, *в*). Отметим, что график заключен в полосе $-1 \leqslant y \leqslant 1$.

Исследование функции $y = \cos x$ аналогично исследованию функции $y = \sin x$ (соответствующие доказательства рекомендуем провести самостоятельно).

1°. *Область определения*: $(-\infty, +\infty)$.

2°. *Функция периодическая с основным периодом* 2π.

3°. *Функция четная.*

4°. *Функция убывает на отрезке* $[0, \pi]$.

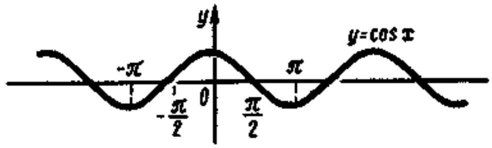


Рис. 82

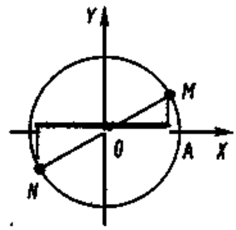


Рис. 83

5^0. *Область значений функции — отрезок* $[-1, 1]$.

6^0. *Функция непрерывна на* $(-\infty, +\infty)$.

График функции $y=\cos x$ изображен на рис. 82.

4. Свойства и графики функций $y=\operatorname{tg} x$, $y=\operatorname{ctg} x$

Отметим свойства функции $y=\operatorname{tg} x$.

1^0. *Область определения:* $x \neq \pi/2+\pi k$ (так как $\operatorname{tg} x=\sin x/\cos x$, то в области определения функции $\operatorname{tg} x$ имеем $\cos x \neq 0$, т. е. $x \neq \pi/2+\pi k$).

2^0. *Функция периодическая с основным периодом* π.

□ Заметим прежде всего, что точки $M(x)$ и $N(x+\pi)$ числовой окружности диаметрально противоположны, так как длина единичной полуокружности равна π. Следовательно, координаты этих точек отличаются только знаками: $\sin(x+\pi)=-\sin x$, $\cos(x+\pi)=-\cos x$ (рис. 83). Тогда

$$\operatorname{tg}(x+\pi)=\frac{\sin(x+\pi)}{\cos(x+\pi)}=\frac{-\sin x}{-\cos x}=\operatorname{tg} x.$$

Таким образом, функция $y=\operatorname{tg} x$ периодическая и π — один из ее периодов. Покажем, что π — основной период; в самом деле, $\operatorname{tg} x=0$ в тех точках, где $\sin x=0$, т. е. в точках $0, \pm\pi, \pm 2\pi, \ldots$. Ближайшие из этих точек отличаются друг от друга на π и, значит, период не может быть меньше, чем π. ■

Так как π — основной период функции $y=\operatorname{tg} x$, то для построения графика этой функции на всей области определения достаточно сначала построить его ветвь на интервале $(-\pi/2, \pi/2)$.

3^0. *Функция нечетная.*

□ В самом деле,

$$\operatorname{tg}(-x)=\frac{\sin(-x)}{\cos(-x)}=\frac{-\sin x}{\cos x}=-\operatorname{tg} x. \ ■$$

График нечетной функции симметричен относительно начала координат, поэтому для построения графика функции $y=\operatorname{tg} x$ на интервале $(-\pi/2, \pi/2)$ достаточно сначала построить ветвь графика на промежутке $[0, \pi/2)$.

4^0. *Функция возрастает на промежутке* $[0, \pi/2)$.

□ Пусть $0 \leqslant x_1 < x_2 < \pi/2$. Тогда $\sin x_1 < \sin x_2$ ($y=\sin x$ возрастает на рассматриваемом промежутке), а $\cos x_1 > \cos x_2$ ($y=\cos x$ убывает на этом промежутке). Следовательно,

$$\operatorname{tg} x_1=\frac{\sin x_1}{\cos x_1}<\frac{\sin x_2}{\cos x_2}=\operatorname{tg} x_2.$$

Итак, из $x_1<x_2$ следует $\operatorname{tg} x_1<\operatorname{tg} x_2$, а это и означает, что $y=\operatorname{tg} x$ — возрастающая функция на $[0, \pi/2)$. ■

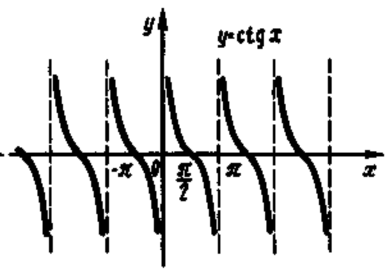


Рис. 84 Рис. 85

Заметим, что $\lim\limits_{x\to\pi/2-0}\frac{\sin x}{\cos x}=+\infty$. Это означает, что прямая $x=\pi/2$ является вертикальной асимптотой графика функции $y=\operatorname{tg}x$.

5°. *Функция непрерывна в любой точке* $x\neq\pi/2+\pi k$.

□ В самом деле, $\operatorname{tg}x=\frac{\sin x}{\cos x}$, а $\sin x$ и $\cos x$ непрерывны в любой точке. По теореме о непрерывности частного функция $\frac{\sin x}{\cos x}$ непрерывна в любой точке, где $\cos x\neq 0$, т. е. в любой точке $x\neq\pi/2+\pi k$. ■

Используя результаты проведенного исследования и взяв контрольные точки (0; 0), ($\pi/4$; 1), ($\pi/3$; $\sqrt{3}$), сначала построим график функции $y=\operatorname{tg}x$ на промежутке $[0, \pi/2)$ (рис. 84, *а*). Затем воспользовавшись нечетностью функции $y=\operatorname{tg}x$, получим график на интервале $(-\pi/2, \pi/2)$ (рис. 84, *б*). Наконец, воспользовавшись периодичностью функции $y=\operatorname{tg}x$, продолжим график на всю область определения (рис. 84, *в*).

Исследование функции $y=\operatorname{ctg}x$ аналогично исследованию функции $y=\operatorname{tg}x$. Поэтому отметим свойства функции $y=\operatorname{ctg}x$ (рекомендуем их доказательства провести самостоятельно).

1°. *Область определения:* $x\neq\pi k$.

2°. *Функция периодическая с основным периодом* π.

3°. *Функция нечетная.*

4°. *Функция убывает на промежутке* $(0, \pi)$.

5°. *Функция непрерывна в любой точке* $x\neq\pi k$.

График функции $y=\operatorname{ctg}x$ изображен на рис. 85.

§ 23. Обратные тригонометрические функции

1. Функция y = arcsin x

Рассмотрим функцию $y=\sin x$. Она монотонна и непрерывна на каждом из следующих отрезков: ..., $[-\pi/2, \pi/2]$, $[\pi/2, 3\pi/2]$, $[3\pi/2, 5\pi/2]$, ... и принимает на каждом из них все значения от -1 до 1 (см. рис. 81). Значит, на каждом из этих промежутков для функции $y=\sin x$ существует обратная функция (по теореме об обратной функции). Следует подчеркнуть, что эти обратные функции являются различными. Из них выделяется одна — функция, обратная по отношению к функции $y=\sin x$, рассматриваемой на отрезке $[-\pi/2, \pi/2]$. Эту функцию обозначают $x=\arcsin y$. Поменяв x и y местами, пишут $y=\arcsin x$.

Итак, $y=\arcsin x$ — это функция, обратная к функции $y=\sin x$, где $-\pi/2\leqslant x\leqslant\pi/2$.

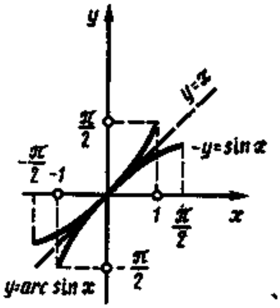


Рис. 86

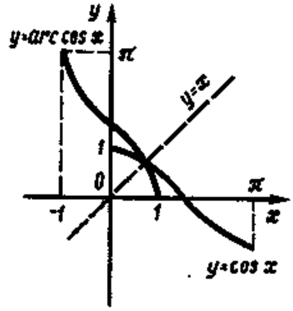


Рис. 87

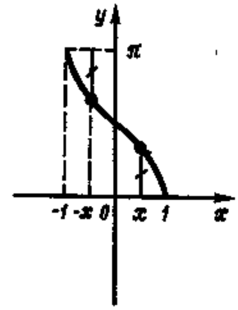


Рис. 88

График функции $y = \arcsin x$ можно получить из графика функции $y = \sin x$, $-\pi/2 \leqslant x \leqslant \pi/2$ преобразованием симметрии относительно прямой $y = x$ (рис. 86).

Отметим свойства функции $y = \arcsin x$.

1^0. *Область определения:* $[-1, 1]$.

2^0. *Функция нечетная:* $\arcsin(-x) = -\arcsin x$.

3^0. *Функция возрастающая.*

4^0. *Область значений функции:* $[-\pi/2, \pi/2]$.

5^0. *Функция непрерывна на* $[-1, 1]$.

Из сказанного выше следует, что записи $y = \arcsin x$ и $x = \sin y$, $-\pi/2 \leqslant y \leqslant \pi/2$ эквивалентны. Подставив в равенство $x = \sin y$ вместо y его выражение, т. е. $\arcsin x$, получим $x = \sin(\arcsin x)$. Следовательно, для любого $x \in [-1, 1]$ имеем

$$\sin(\arcsin x) = x; \quad -\pi/2 \leqslant \arcsin x \leqslant \pi/2.$$

2. Функция $y = \arccos x$

Функция $y = \cos x$ монотонна на отрезке $[0, \pi]$ (убывает на этом отрезке), непрерывна и принимает на нем все значения от -1 до 1 (см. рис. 82). Значит, для функции $y = \cos x$, рассматриваемой на отрезке $[0, \pi]$, существует обратная функция. Она обозначается $y = \arccos x$. График этой функции получается из графика функции $y = \cos x$, $0 \leqslant x \leqslant \pi$ преобразованием симметрии относительно прямой $y = x$ (рис. 87).

Отметим свойства функции $y = \arccos x$.

1^0. *Область определения*: $[-1, 1]$.

2^0. *Функция убывающая.*

3^0. *Область значений функции:* $[0, \pi]$.

4^0. *Функция непрерывна на* $[-1, 1]$.

Функция не является ни четной, ни нечетной. Это следует хотя бы из того, что график функции не является симметричным ни относительно оси ординат, ни относительно начала координат. Заметим, однако, что график функции $y = \arccos x$ симметричен относительно точки $(0; \pi/2)$. Отсюда получаем $\arccos(-x) - \pi/2 = \pi/2 - \arccos x$, т. е. $\arccos(-x) = \pi - \arccos x$ (рис. 88).

Из сказанного выше следует, что записи $y = \arccos x$ и $x = \cos y$, $0 \leqslant y \leqslant \pi$ эквивалентны. Подставив в равенство $x = \cos y$ вместо y выражение $\arccos x$, получим $x = \cos(\arccos x)$. Следовательно, для любого $x \in [-1, 1]$ имеем

$$\cos(\arccos x) = x; \quad 0 \leqslant \arccos x \leqslant \pi.$$

3. Функция $y = \operatorname{arctg} x$

Функция $y = \operatorname{tg} x$ монотонна на интервале $(-\pi/2, \pi/2)$ (возрастает на этом интервале), непрерывна и принимает на нем все значения (см. рис. 84). Поэтому на указанном интервале для функции $y = \operatorname{tg} x$ существует обратная функция. Она обозначается $y = \operatorname{arctg} x$. График этой функции получается из графика функции $y = \operatorname{tg} x$, $-\pi/2 < x < \pi/2$ преобразованием симметрии относительно прямой $y = x$ (рис. 89).

Отметим свойства функции $y = \operatorname{arctg} x$.

1°. *Область определения:* $(-\infty, +\infty)$.

2°. *Функция нечетная:* $\operatorname{arctg}(-x) = -\operatorname{arctg} x$.

3°. *Функция возрастающая.*

4°. *Область значений функции:* $(-\pi/2, \pi/2)$.

5°. *Функция непрерывна на* $(-\infty, +\infty)$.

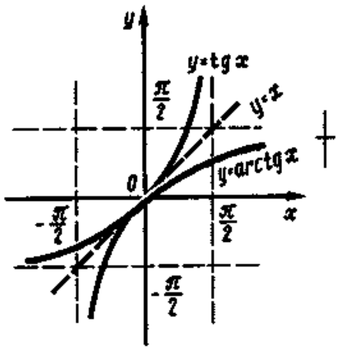


Рис. 89

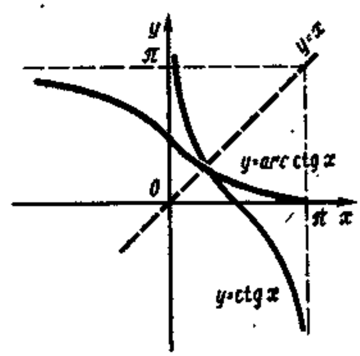


Рис. 90

Из сказанного выше следует, что записи $y = \operatorname{arctg} x$ и $x = \operatorname{tg} y$, $-\pi/2 < y < \pi/2$ эквивалентны. Для любого x, $-\infty < x < +\infty$ имеем

$$\operatorname{tg}(\operatorname{arctg} x) = x; \quad -\pi/2 < \operatorname{arctg} x < \pi/2.$$

Прямые $y = \pi/2$ и $y = -\pi/2$ являются горизонтальными асимптотами графика функции $y = \operatorname{arctg} x$ соответственно при $x \to +\infty$ и при $x \to -\infty$. Значит,

$$\lim_{x \to +\infty} \operatorname{arctg} x = \pi/2, \quad \lim_{x \to -\infty} \operatorname{arctg} x = -\pi/2.$$

4. Функция $y = \operatorname{arcctg} x$

Функция $y = \operatorname{ctg} x$ монотонна на интервале $(0, \pi)$ (убывает на этом интервале), непрерывна и принимает на нем все значения (см. рис. 85). Следовательно, на этом интервале для функции $y = \operatorname{ctg} x$ существует обратная функция. Она обозначается $y = \operatorname{arcctg} x$. График функции $y = \operatorname{arcctg} x$ получается из графика функции $y = \operatorname{ctg} x$, $0 < x < \pi$ преобразованием симметрии относительно прямой $y = x$ (рис. 90).

Отметим свойства функции $y = \operatorname{arcctg} x$.

1°. *Область определения:* $(-\infty, +\infty)$.

2°. *Функция убывающая.*

3°. *Область значений функции:* $(0, \pi)$.

4°. *Функция непрерывна на* $(-\infty, +\infty)$.

Функция не является ни четной, ни нечетной. График этой функции, как и график функции $y = \arccos x$, симметричен относительно точки $(0; \pi/2)$. Отсюда получаем, что $\operatorname{arcctg}(-x) = \pi - \operatorname{arcctg} x$.

Из сказанного выше следует, что записи $y = \operatorname{arcctg} x$ и $x = \operatorname{ctg} y$, $0 < y < \pi$ эквивалентны. Для любого x, $-\infty < x < +\infty$ имеем

$$\operatorname{ctg}(\operatorname{arcctg} x) = x, \ 0 < \operatorname{arcctg} x < \pi.$$

Прямые $y = 0$ и $y = \pi$ являются горизонтальными асимптотами графика функции $y = \operatorname{arcctg} x$ соответственно при $x \to +\infty$ и при $x \to -\infty$. Значит,

$$\lim_{x\to+\infty} \operatorname{arcctg} x = 0, \quad \lim_{x\to-\infty} \operatorname{arcctg} x = \pi.$$

Первое из этих равенств означает, что $\operatorname{arcctg} x$ — бесконечно малая функция при $x \to +\infty$.

Функции $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$, $y = \operatorname{arcctg} x$ называются *обратными тригонометрическими функциями* (или аркфункциями).

5. Преобразования выражений, содержащих обратные тригонометрические функции

Прежде чем переходить к примерам, выпишем наиболее важные соотношения для обратных тригонометрических функций, особенно часто используемые на практике (все они получены выше):

$$-\pi/2 \leqslant \arcsin x \leqslant \pi/2; \ \arcsin(-x) = -\arcsin(x);$$

$$0 \leqslant \arccos x \leqslant \pi; \ \arccos(-x) = \pi - \arccos(x);$$

$$-\pi/2 < \operatorname{arctg} x < \pi/2; \ \operatorname{arctg}(-x) = -\operatorname{arctg}(x);$$

$$0 < \operatorname{arcctg} x < \pi; \ \operatorname{arcctg}(-x) = \pi - \operatorname{arcctg} x.$$

Примеры. 1. Вычислить:

а) $\arcsin(\sqrt{3}/2)$; б) $\arccos(-\sqrt{2}/2)$; в) $\operatorname{arctg}(-\sqrt{3})$.

○ а) По определению, $y = \arcsin(\sqrt{3}/2)$ — это такое число, что $\sin y = \sqrt{3}/2$ и $-\pi/2 \leqslant y \leqslant \pi/2$. Отсюда следует, что $y = \pi/3$. Таким образом, $\arcsin(\sqrt{3}/2) = \pi/3$.

б) Мы отметили, что $\arccos(-x) = \pi - \arccos x$. Значит, $\arccos(-\sqrt{2}/2) = \pi - \arccos(\sqrt{2}/2) = \pi - \pi/4 = 3\pi/4$.

в) Так как $\operatorname{arctg}(-x) = -\operatorname{arctg} x$, то

$$\operatorname{arctg}(-\sqrt{3}) = -\operatorname{arctg}(\sqrt{3}) = -\pi/3. \ ●$$

2. Упростить выражение:

а) $\cos(\arcsin x)$, где $-1 \leqslant x \leqslant 1$; б) $\sin(\operatorname{arctg} x)$.

○ а) Положим $\arcsin x = y$; тогда $\sin y = x$, $-\pi/2 \leqslant y \leqslant \pi/2$. Нужно найти $\cos y$.

Известно, что $\cos^2 y = 1 - \sin^2 y$. Значит, $\cos^2 y = 1 - x^2$. Но $-\pi/2 \leqslant y \leqslant \pi/2$, а на отрезке $[-\pi/2, \pi/2]$ косинус принимает неотрицательные значения. Поэтому

$$\cos y = \sqrt{1-x^2}, \text{ т. е. } \cos(\arcsin x) = \sqrt{1-x^2},$$

где $-1 \leqslant x \leqslant 1$.

б) Положим $\operatorname{arctg} x = y$; тогда $\operatorname{tg} y = x$, $-\pi/2 < y < \pi/2$. Нужно найти $\sin y$.

Известно, что $1 + \operatorname{tg}^2 y = 1/\cos^2 y$, откуда $\cos^2 y = 1/(1 + \operatorname{tg}^2 y)$. Но $-\pi/2 < y < \pi/2$, а на интервале $(-\pi/2, \pi/2)$ косинус принимает лишь положительные значения. Поэтому $\cos y = 1/\sqrt{1 + \operatorname{tg}^2 y}$, т. е. $\cos(\operatorname{arctg} x) = 1/\sqrt{1+x^2}$. Так как $\sin y = \operatorname{tg} y \cos y$, то $\sin(\operatorname{arctg} x) = x\sqrt{1+x^2}$. ●

3. Доказать, что для любого $x \in [-1, 1]$ справедливо тождество

$$\arcsin x = \pi/2 - \arccos x. \tag{1}$$

○ Вычислим значение синуса от обеих частей доказываемого равенства:

$$\sin(\arcsin x) = x; \ \sin(\pi/2 - \arccos x) = \cos(\arccos x) = x.$$

Так как синусы обеих частей равны, то чтобы убедиться в справедливости равенства (1), остается показать, что $\arcsin x$ и $\pi/2 - \arccos x$ принадлежат одному и тому же промежутку монотонности функции $y = \sin x$ (без проверки этого условия можно получить неверный результат — ведь тригонометрические функции могут принимать одинаковые значения и для различных значений аргумента: например, $\pi/6 \neq 5\pi/6$, но $\sin(\pi/6) = \sin 5(\pi/6) = 1/2$).

Имеем $-\pi/2 \leqslant \arcsin x \leqslant \pi/2$; кроме того, $0 \leqslant \arccos x \leqslant \pi$, а потому $-\pi/2 \leqslant \leqslant \pi/2 - \arccos x \leqslant \pi/2$. Итак, $\arcsin x$ и $\pi/2 - \arccos x$ принадлежат одному промежутку монотонности $[-\pi/2, \pi/2]$ функции $y = \sin x$. Теперь можно сказать, что тождество (1) доказано. ●

Аналогично можно доказать, что $\operatorname{arctg} x = \pi/2 - \operatorname{arcctg} x$.

§ 24. Тригонометрические уравнения

1. Решение простейших тригонометрических уравнений

Выведем формулы для решения *простейших тригонометрических уравнений*, т. е. уравнений вида $T(x) = a$, где $T(x)$ — некоторая тригонометрическая функция.

Рассмотрим сначала уравнение $\sin x = a$, где $-1 \leqslant a \leqslant 1$. По определению, $\sin x$ — это ордината такой точки числовой окружности, которая соответствует числу x. На числовой окружности имеются две симметричные друг другу относительно оси OY точки M и N, ординаты которых равны a (рис. 91); одна из этих точек лежит на полуокружности DAB, вторая — на полуокружности BCD. Если $a = 1$ или $a = -1$, то эти точки сливаются в одну. Одно из чисел, которому на числовой окружности соответствует точка M, есть $\arcsin a$; поэтому все числа, которым соответствует точка M, имеют вид $\arcsin a + 2k\pi$. Точка N также соответствует бесконечному множеству чисел. Из равенства $\sin(\pi - \alpha) = \sin \alpha$ видно, что одно из чисел такого множества равно $\pi - \arcsin a$, а потому все числа этого множества имеют вид $\pi - \arcsin a + 2k\pi = -\arcsin a + (2k + 1)\pi$.

Итак, корни уравнения $\sin x = a$ записываются в следующем виде:

$$x = \arcsin a + \pi \cdot 2k, \tag{1}$$

$$x = -\arcsin a + \pi(2k + 1). \tag{2}$$

Заметим, что формулы (1) и (2) можно объединить в одну:

$$x = (-1)^n \arcsin a + \pi n. \tag{3}$$

Действительно, при четном n $(n = 2k)$ из формулы (3) получается формула (1), а при нечетном n $(n = 2k + 1)$ — формула (2).

Далее, рассмотрим уравнение $\cos x = a$, $-1 \leqslant a \leqslant 1$. По определению, $\cos x$ — это абсцисса такой точки числовой окружности, которая соответствует числу x. На числовой окружности имеются две точки M и N, абсциссы которых

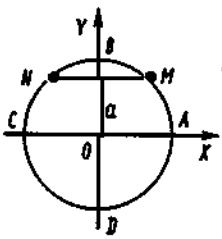


Рис. 91

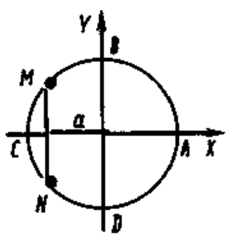


Рис. 92

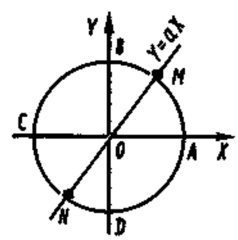


Рис. 93

равны a. Эти точки симметричны относительно оси абсцисс, и потому величины дуг AM и AN равны по модулю, но противоположны по знаку (рис. 92). Отсюда следует, что точка M соответствует числам вида $\arccos a + 2\pi n$, а точка N — числам вида $-\arccos a + 2\pi n$. Поэтому решение уравнения $\cos x = a$ имеет вид

$$x = \pm \arccos a + 2\pi n. \tag{4}$$

Рассмотрим теперь уравнение $\operatorname{tg} x = a$. Мы знаем, что $\operatorname{tg} x = \sin x / \cos x$, т. е. $\operatorname{tg} x$ равен отношению ординаты точки M числовой окружности к ее абсциссе. Поэтому найдем сначала точки окружности, для которых $Y/X = a$. Эти точки лежат на пересечении окружности с прямой $Y = aX$. Мы получаем две диаметрально противоположные точки M и N, одна из которых (M) лежит на полуокружности DAB, а вторая (N) — на полуокружности BCD (рис. 93). Точка M соответствует числам вида $\operatorname{arctg} a + 2k\pi$. Поскольку длина дуги MCN равна π, точка N соответствует числам вида $\operatorname{arctg} a + \pi + 2k\pi$. Эти две формулы можно объединить в одну: $\operatorname{arctg} a + \pi n$.

Итак, решение уравнения $\operatorname{tg} x = a$ имеет вид

$$x = \operatorname{arcctg} a + \pi n. \tag{5}$$

Точно так же для уравнения $\operatorname{ctg} x = a$ получаем

$$x = \operatorname{arctg} a + \pi n. \tag{6}$$

Примеры. 1. Решить уравнение $\sin x = 1/2$.

○ По формуле (3) получаем $x = (-1)^n \arcsin(1/2) + \pi n$. Так как $\arcsin(1/2) = \pi/6$, то окончательно имеем

$$x = (-1)^n \frac{\pi}{6} + \pi n. \; ●$$

2. Решить уравнение $\cos 3x = -\sqrt{2}/2$.

○ Используя формулу (4), получим $3x = \arccos(-\sqrt{2}/2) + 2\pi n$. Так как $\arccos(-\sqrt{2}/2) = \pi - \arccos(\sqrt{2}/2) = \pi - \pi/4 = 3\pi/4$, то

$$3x = \pm \frac{3\pi}{4} + 2\pi n, \text{ т. е. } x = \pm \frac{\pi}{4} + \frac{2\pi n}{3}. \; ●$$

3. Решить уравнение $\operatorname{tg}\left(x - \frac{\pi}{6}\right) = -\sqrt{3}$.

○Используя формулу (5), получим $x - \frac{\pi}{6} = \operatorname{arctg}(-\sqrt{3}) + \pi n$. Так как $\operatorname{arctg}(-\sqrt{3}) = -\operatorname{arctg}(\sqrt{3}) = -n/3$, то

$$x - \frac{\pi}{6} = -\frac{\pi}{3} + \pi n, \text{ т. е. } x = -\frac{\pi}{6} + \pi n. \; ●$$

Заметим, что в некоторых случаях удобнее пользоваться частными формулами, которые легко выводятся с помощью числовой окружности:

$$\begin{aligned}
\sin x = 0&: \qquad x = \pi n;\\
\sin x = 1&: \qquad x = \pi/2 + 2\pi n;\\
\sin x = -1&: \qquad x = -\pi/2 + 2\pi n;\\
\cos x = 0&: \qquad x = \pi/2 + \pi n;\\
\cos x = 1&: \qquad x = 2\pi n;\\
\cos x = -1&: \qquad x = \pi + 2\pi n.
\end{aligned}$$

2. Решение тригонометрических уравнений методом введения новой переменной

Отметим, что имеются два основных метода, приводящих к упрощению тригонометрического уравнения: метод введения новой переменной и метод разложения на множители. Оба этих метода применяются для сведения данного тригонометрического уравнения к более простому уравнению или к совокупности более простых уравнений. Рассмотрим *метод введения новой переменной*.

Пусть тригонометрическое уравнение имеет вид $f(\sin x) = 0$, где f — некоторая алгебраическая функция. Тогда подстановкой $\sin x = z$ оно сводится к алгебраическому уравнению $f(z) = 0$. Пусть $z_1, z_2, \ldots, z_n$ — такие корни этого уравнения, для которых $|z_k| \leqslant 1$. Тогда уравнение $f(\sin x) = 0$ равносильно совокупности уравнений

$$\sin x = z_1;\ \sin x = z_2; \ldots;\ \sin x = z_n.$$

Аналогично решаются уравнения вида $f(\cos x) = 0$, $f(\operatorname{tg} x) = 0$, $f(\operatorname{ctg} x) = 0$.

Часто перед выполнением подстановки приходится делать те или иные тождественные преобразования. Если в уравнение входят тригонометрические функции одного и того же аргумента, то нужно выразить все эти функции через одну из них, например через $\sin x$, а затем подстановкой $\sin x = z$ свести это уравнение к алгебраическому.

Примеры. 1. Решить уравнение $2\cos^2 x + 4\cos x = 3\sin^2 x$.

○ Так как $\sin^2 x = 1 - \cos^2 x$, то уравнение можно переписать следующим образом:

$$2\cos^2 x + 4\cos x = 3(1 - \cos^2 x), \text{ т. е. } 5\cos^2 x + 4\cos x - 3 = 0.$$

Положим $\cos x = z$; тогда получим квадратное уравнение $5z^2 + 4z - 3 = 0$, решая которое, находим $z_1 = (-2 - \sqrt{19})/5$; $z_2 = (-2 + \sqrt{19})/5$. Последний корень не годится, так как $|(2 + \sqrt{19})/5| > 1$. Значит, остается решить уравнение $\cos x = (-2 + \sqrt{19})/5$. Из него получаем

$$x = \pm \arccos \frac{-2+\sqrt{19}}{5} + 2\pi n. \ \bullet$$

2. Решить уравнение $\cos^4 x + 3\sin x - \sin^4 x - 2 = 0$.

○ Целесообразно заменить $\cos^2 x$ на $1 - \sin^2 x$, т. е. $\cos^4 x$ на $(1 - \sin^2 x)^2$. Тогда уравнение примет вид

$$(1 - \sin^2 x)^2 + 3\sin x - \sin^4 x - 2 = 0.$$

Подстановкой $\sin x = z$ последнее уравнение сводится к алгебраическому уравнению $(1 - z^2)^2 + 3z - z^4 - 2 = 0$, и, далее, $2z^2 - 3z + 1 = 0$, откуда $z_1 = 1$, $z_2 = 1/2$. Теперь задача сводится к решению совокупности уравнений $\sin x = 1$; $\sin x = 1/2$; в результате получаем

$$x = \frac{\pi}{2} + 2\pi n,\ x = (-1)^k \frac{\pi}{6} + \pi k. \ \bullet$$

3. Однородные уравнения

В некоторых случаях тригонометрическое уравнение вида $R(\cos x, \sin x)=0$, где $R(u, v)$ — рациональное выражение, можно свести к алгебраическому уравнению относительно $\operatorname{tg} x$. Примерами таких уравнений могут служить однородные уравнения:

$$a\sin x+b\cos x=0;\ a\sin^2 x+b\sin x\cos x+c\cos^2 x=0;$$

$$a\sin^3 x+b\sin^2 x\cos x+c\sin x\cos^2 x+d\cos^3 x=0.$$

Рассмотрим случай, когда $a\neq 0$. Разделим обе части первого уравнения на $\cos x$, обе части второго — на $\cos^2 x$, а обе части третьего — на $\cos^3 x$. В результате получим соответственно следующие уравнения, алгебраические относительно $\operatorname{tg} x$, а потому решаемые подстановкой $\operatorname{tg} x=z$:

$$a\operatorname{tg} x+b=0;\ a\operatorname{tg}^2 x+b\operatorname{tg} x+c=0;$$

$$a\operatorname{tg}^3 x+b\operatorname{tg}^2 x+c\operatorname{tg} x+d=0.$$

Нетрудно видеть, что при $a\neq 0$ однородным уравнениям не удовлетворяют те значения x, при которых $\cos x=0$. Поэтому деление на $\cos x$, $\cos^2 x$, $\cos^3 x$ обеих частей однородных уравнений в случае $a\neq 0$ не приводит к потере корней.

Примеры. 1. Решить уравнение $8\sin x-7\cos x=0$.

○ Разделив обе части уравнения почленно на $\cos x$, получим $8\operatorname{tg} x-7=0$. Далее, имеем $\operatorname{tg} x=7/8$, откуда $x=\operatorname{arctg}(7/8)+\pi n$. ●

2. Решить уравнение $\sin^2 x+2\sin x\cos x-3\cos^2 x=0$.

○ Разделив обе части этого однородного уравнения второй степени на $\cos^2 x$, получим $\operatorname{tg}^2 x+2\operatorname{tg} x-3=0$. Далее, положим $u=\operatorname{tg} x$; тогда приходим к квадратному уравнению $u^2+2u-3=0$, откуда $u_1=-3$, $u_2=1$.

Решив совокупность уравнений $\operatorname{tg} x=-3$; $\operatorname{tg} x=1$, находим $x=\operatorname{arctg}(-3)+\pi k$; $x=\pi/4+\pi n$. ●

3. Решить уравнение $5\sin^2 x+\sqrt{3}\sin x\cos x+6\cos^2 x=5$.

○ Имеем

$$5\sin^2 x+\sqrt{3}\sin x\cos x+6\cos^2 x=5(\sin^2 x+\cos^2 x),$$

откуда

$$\sqrt{3}\sin x\cos x+\cos^2 x=0. \tag{7}$$

В полученном уравнении отсутствует член вида $a\sin^2 x$, т. е. $a=0$. Здесь делить обе части уравнения на $\cos^2 x$ нельзя, так как те значения x, при которых $\cos^2 x=0$, удовлетворяют уравнению (7), а потому деление на $\cos^2 x$ приведет к потере корней. Поступим иначе: разложив левую часть уравнения (7) на множители, получим

$$\cos x(\sqrt{3}\sin x+\cos x)=0.$$

Теперь задача сводится к решению совокупности уравнений

$$\cos x=0;\quad \sqrt{3}\sin x+\cos x=0. \tag{8}$$

Из первого уравнения совокупности (8) находим $x=\pi/2+\pi k$. Разделив обе части однородного уравнения первой степени $\sqrt{3}\sin x+\cos x=0$ на $\cos x$, имеем $\sqrt{3}\operatorname{tg} x+1=0$, откуда $\operatorname{tg} x=-1/\sqrt{3}$; $x=\operatorname{arctg}(-1/\sqrt{3})+\pi n$, т. е. $x=-\pi/6+\pi n$.

Итак, получаем две серии решений: $x=\pi/2+\pi n$; $x=-\pi/6+\pi n$. ●

4. Универсальная подстановка

Подстановки, которые мы рассматривали до сих пор, годились лишь для специальных случаев уравнений вида $R(\cos x, \sin x)=0$. Рассмотрим теперь подстановку, позволяющую свести к рациональному уравнению любое уравнение вида $R(\cos x, \sin x)=0$. Это подстановка $\operatorname{tg}(x/2)=u$. Если $x\neq\pi+2\pi n$, то справедливы следующие тождества:

$$\cos x=\frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)};\ \sin x=\frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)}. \tag{9}$$

В самом деле, имеем

$$\frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)}=\frac{1-\dfrac{\sin^2(x/2)}{\cos^2(x/2)}}{1/\cos^2(x/2)}=\cos^2(x/2)-\sin^2(x/2)=\cos x,$$

$$\frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)}=\frac{2\dfrac{\sin(x/2)}{\cos(x/2)}}{1/\cos^2(x/2)}=2\sin(x/2)\cos(x/2)=\sin x.$$

Поэтому подстановка $\operatorname{tg}(x/2)=u$ преобразует уравнение $R(\cos x, \sin x)=0$ в уравнение

$$R\left(\frac{1-u^2}{1+u^2}, \frac{2u}{1+u^2}\right)=0.$$

Левая часть этого уравнения является рациональной функцией от u. Значит, указанная подстановка привела уравнение к рациональному виду. Подстановка $\operatorname{tg}(x/2)=u$ называется *универсальной*.

Так как использование универсальной подстановки возможно лишь при $x\neq\pi+2\pi n$, то нужно проверять, не являются ли числа вида $x=\pi+2\pi n$ решениями заданного уравнения.

Примеры. 1. Решить уравнение $3\sin x+4\cos x=5$.

○ Выражая $\sin x$ и $\cos x$ через $\operatorname{tg}(x/2)$ по формулам (9) и полагая $\operatorname{tg}(x/2)=u$, придем к рациональному уравнению

$$3\cdot\frac{2u}{1+u^2}+4\cdot\frac{1-u^2}{1+u^2}=5.$$

Решив это уравнение, получим $u=1/3$. Из уравнения $\operatorname{tg}(x/2)=1/3$ находим

$$\frac{x}{2}=\operatorname{arctg}\frac{1}{3}+\pi n,\ \text{т. е. } x=2\operatorname{arctg}\frac{1}{3}+2\pi n.$$

Проверкой убеждаемся, что значения $x=\pi+2\pi n$ не удовлетворяют заданному уравнению. Итак, получаем ответ: $x=2\operatorname{arctg}(1/3)+2\pi n$. ●

2. Решить уравнение $3\sin 2x+\cos 2x+1=0$.

○ Воспользуемся универсальной подстановкой. Выражая $\sin 2x$ и $\cos 2x$ через $\operatorname{tg} x$ и полагая $\operatorname{tg} x=u$, получим рациональное уравнение

$$\frac{6u}{1+u^2}+\frac{1-u^2}{1+u^2}+1=0,$$

откуда $u=-1/3$. Из уравнения $\operatorname{tg} x=-1/3$ находим $x=\operatorname{arctg}(-1/3)+\pi k$.

Однако нужно еще проверить, не удовлетворяют ли заданному уравнению те значения x, при которых $2x=\pi+2\pi n$, т. е. значения $x=\pi/2+\pi n$. Имеем

$$3\sin(\pi+2\pi n)+\cos(\pi+2\pi n)+1=3\sin\pi+\cos\pi+1=3\cdot0-1+1=0.$$

Проверка показывает, что значения $\pi/2+\pi n$ являются решениями уравнения. Итак, заданное уравнение имеет следующие решения: $x=\operatorname{arctg}(-1/3)+\pi n$; $x=\pi/2+\pi n$. ●

5. Решение тригонометрических уравнений методом разложения на множители

Сущность *метода разложения на множители* заключается в следующем. Пусть дано тригонометрическое уравнение $f(x)=0$, где $f(x)=g(x)h(x)$. Тогда все решения уравнения $f(x)=0$ являются и решениями совокупности уравнений $g(x)=0$; $h(x)=0$. Обратное утверждение неверно, так как среди решений одного из этих уравнений могут найтись такие, которые не принадлежат области определения другого уравнения, а значит, и области определения уравнения $f(x)=0$. Так, преобразовав уравнение $\cos x(\operatorname{tg} x-1)=0$ в совокупность уравнений $\cos x=0$; $\operatorname{tg} x=1$, получим $x=\pi/2+\pi k$, $x=\pi/4+\pi k$. Однако значения $x=\pi/2+\pi k$ являются посторонними корнями для уравнения $\cos x(\operatorname{tg} x-1)=0$, поскольку при указанных значениях не определен $\operatorname{tg} x$.

Переход от уравнения $g(x)\,h(x)=0$ к совокупности уравнений $g(x)=0$; $h(x)=0$ является равносильным преобразованием, если область определения каждого из уравнений совокупности совпадает с областью определения исходного уравнения (или шире ее). Это имеет место, например, в том случае, когда левая часть уравнения $g(x)\,h(x)=0$ определена при всех x.

Пример. Решить уравнение $\sin 5x+\sin x+2\sin^2 x=1$.

○ Применим к первым двум слагаемым формулу суммы синусов и воспользуемся тем, что $2\sin^2 x=1-\cos 2x$. Тогда уравнение примет вид

$$2\sin 3x\cos 2x+1-\cos 2x=1, \text{ или } \cos 2x\,(2\sin 3x-1)=0.$$

Теперь задача сводится к решению совокупности уравнений $\cos 2x=0$; $2\sin 3x-1=0$, равносильной заданному уравнению. Решая уравнения этой совокупности, получаем следующие две серии решений заданного уравнения: $x=\frac{\pi}{4}+\frac{\pi n}{2}$; $x=(-1)^n\frac{\pi}{18}+\frac{\pi n}{3}$.●

6. Метод введения вспомогательного аргумента

Иногда при решении тригонометрических уравнений оказывается полезным заменить выражение $a\cos x+b\sin x$ на $A\sin(x+\varphi)$, где $A=\sqrt{a^2+b^2}$, $\sin\varphi=a/\sqrt{a^2+b^2}$, $\cos\varphi=b/\sqrt{a^2+b^2}$. В этом случае φ называют *вспомогательным аргументом*.

Примеры. 1. Решить уравнение $8\cos x+15\sin x=17$.

○ Разделив обе части уравнения на $\sqrt{8^2+15^2}=17$, получим

$$\frac{8}{17}\cos x+\frac{15}{17}\sin x=1. \qquad (10)$$

Так как $(8/17)^2+(15/17)^2=1$, то существует такое φ, что $8/17=\sin\varphi$ и $15/17=\cos\varphi$. Перепишем уравнение (10) следующим образом:

$$\sin\varphi\cos x+\sin x\cos\varphi=1.$$

Но $\sin\varphi\cos x+\sin x\cos\varphi=\sin(x+\varphi)$. Значит, $\sin(x+\varphi)=1$, откуда $x=\pi/2+2\pi n-\varphi$. Поскольку $\varphi=\arcsin(8/17)$, окончательно получаем следующие решения заданного уравнения: $x=\pi/2-\arcsin(8/17)+2\pi n$.●

2. Решить уравнение $5\sin x-12\cos x+13\sin 3x=0$.

○ Имеем

$$\sqrt{5^2+12^2}\left(\frac{5}{\sqrt{5^2+12^2}}\sin x-\frac{12}{\sqrt{5^2+12^2}}\cos x\right)=-13\sin 3x;$$

$$\frac{5}{13}\sin x-\frac{12}{13}\cos x=-\sin 3x.$$

Полагая $5/13 = \cos\varphi$, $12/13 = \sin\varphi$, получим

$$\sin x\cos\varphi - \cos x\sin\varphi = -\sin 3x$$

и, далее,

$$\sin(x-\varphi) + \sin 3x = 0;\ 2\sin\left(2x - \frac{\varphi}{2}\right)\cos\left(x + \frac{\varphi}{2}\right) = 0.$$

Решая совокупность уравнений $\sin\left(2x - \frac{\varphi}{2}\right) = 0$; $\cos\left(x + \frac{\varphi}{2}\right) = 0$, находим $x = \frac{\varphi}{4} + \frac{\pi n}{2}$; $x = \frac{\pi}{2} - \frac{\varphi}{2} + \pi n$. Учитывая, что $\varphi = \arccos(5/13)$, в итоге получаем следующие две серии решений заданного уравнения: $x = \frac{1}{4}\arccos\frac{5}{13} + \frac{\pi n}{2}$, $x = \frac{\pi}{2} - \frac{1}{2}\arccos\frac{5}{13} + \pi n$.●

§ 25. Построение графиков функций с помощью геометрических преобразований известных графиков

1. Преобразования графиков

Рассмотрим в общем виде задачи на построение графиков функций с помощью преобразований известных графиков.

Задача 1. Построить график функции $y = mf(x)$, где $m > 0$, $m \neq 1$, если задан график функции $y = f(x)$.

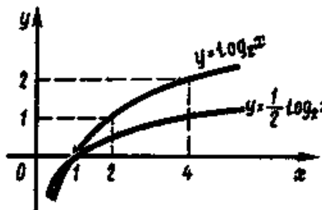


Рис. 94

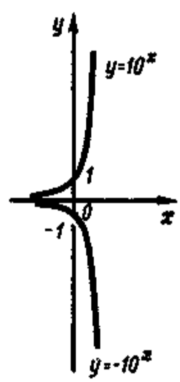


Рис. 95

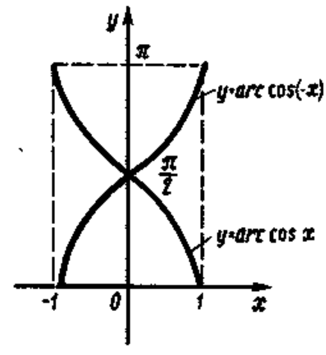


Рис. 96

○ Ординаты точек графика функции $y = f(x)$ получаются умножением на m соответствующих ординат точек графика функции $y = f(x)$. Такое преобразование графика функции $y = f(x)$ называется его растяжением от оси x с коэффициентом m, если $m > 1$, и сжатием к оси x, если $0 < m < 1$.●

На рис. 94 изображены графики функций $y = \log_2 x$ и $y = \frac{1}{2}\log_2 x$.

Задача 2. Построить график функции $y = -f(x)$, если задан график функции $y = f(x)$.

○ При одном и том же значении x ординаты точек графиков функций $y = f(x)$ и $y = -f(x)$ отличаются только знаком. Значит, график функции $y = -f(x)$ можно получить из графика функции $y = f(x)$ преобразованием симметрии последнего относительно оси x.●

На рис. 95 изображены графики функций $y = 10^x$ и $y = -10^x$.

Задача 3. Построить график функции $y = mf(x)$, где $m < 0$, $m \neq -1$, если задан график функции $y = f(x)$.

○ Так как $mf(x) = -|m| f(x)$, то график функции $y = mf(x)$ можно получить с помощью растяжения или сжатия графика функции $y = f(x)$ к оси x с коэффициентом $|m|$ и последующим преобразованием симметрии относительно оси x (см. задачи 1 и 2).●

Задача 4. Построить график функции $y = f(x)$, если задан график функции $y = f(-x)$.

○ Ордината графика функции $y = f(x)$ в точке x равна ординате графика функции $y = f(-x)$ в точке $-x$. Это означает, что график функции $y = f(-x)$ можно получить из графика функции $y = f(x)$ преобразованием симметрии относительно оси y.●

На рис. 96 изображены графики функций $y = \arccos x$ и $y = \arccos(-x)$.

Задача 5. Построить график функции $y = f(kx)$, если задан график функции $y = f(x)$.

○ Рассмотрим случай, когда $k > 0$. Ясно, что ордината графика функции $y = f(kx)$ в точке x равна ординате графика функции $y = f(x)$ в точке kx. Так, $f(x) = f(k)$ при $x = k$, а $f(kx) = f(k)$ при $x = 1$; $f(x) = f(2k)$ при $x = 2k$, а $f(kx) = f(2k)$ при $x = 2$ и т. д. Это означает, что график функции $y = f(kx)$ получается из графика функции $y = f(x)$ сжатием с коэффициентом k к оси y, если $k > 1$, и растяжением, если $0 < k < 1$.

Пусть теперь $k < 0$. Тогда $f(kx) = f(-|k|)$. Поэтому график функции $y = f(kx)$ можно получить сжатием (или растяжением) графика функции $y = f(x)$ с коэффициентом $|k|$ к оси y с последующим преобразованием симметрии полученного графика функции $y = f(|k|x)$ относительно оси y.●

На рис. 97 изображены графики функций $y = \sin x$ и $y = \sin 3x$, а на рис. 98 — графики функций $y = \cos x$ и $y = \cos\frac{x}{3}$.

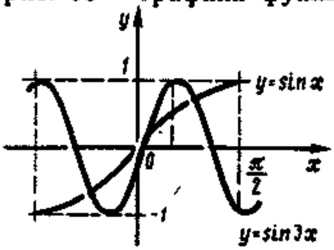


Рис. 97

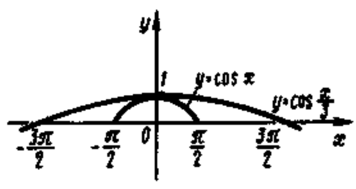


Рис. 98

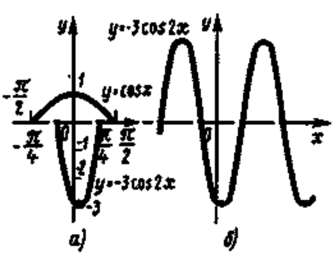


Рис. 99

Пример 1. Построить график функции $y = -3\cos 2x$.

○ Построим график функции $y = \cos x$ (при этом можно ограничиться одной полуволной графика). Произведя его сжатие к оси y с коэффициентом 2, получим график функции $y = \cos 2x$. Теперь произведем растяжение полученного графика от оси x с коэффициентом 3, а затем преобразование симметрии относительно оси x. В результате мы получим график функции $y = -3\cos 2x$. На рис. 99, *а* показана одна полуволна графика, а на рис. 99, *б* — весь график.●

Задача 6. Построить график функции $y = f(x - m) + n$, если задан график функции $y = f(x)$.

○ Положим $x' = x - m$, $y' = y - n$. Тогда формулу $y = f(x - m) + n$, или, что то же самое, $y - n = f(x - m)$, можно переписать в виде $y' = f(x')$. Таким образом, график функции $y = f(x - m) + n$, построенный в координатной плоскости xy, совпадает с графиком функции $y' = f(x')$, построенным в координатной плоскости $x'y'$.

Формулы $x' = x - m$, $y' = y - n$, или, что то же самое, $x = x' + m$, $y = y' + n$, задают параллельный перенос, при котором любая точка $(x'; y')$ переходит в точку $(x' + m; y' + n)$ и, в частности, начало координат переходит в точку $(m; n)$.

Таким образом, чтобы построить график функции $y = f(x - m) + n$, нужно:

1°. Выполнить параллельный перенос плоскости, выбрав в качестве начала новой системы координат $x'y'$ точку O' $(m; n)$.

2°. В плоскости $x'y'$ построить график функции $y' = f(x')$. ●

Пример 2. Построить график функции $y = \sqrt{x-2} + 4$.

○ 1°. Выполним параллельный перенос плоскости, поместив начало новой системы $x'y'$ координат в точку O' (2; 4).

2°. В плоскости $x'y'$ построим график функции $y' = \sqrt{x'}$. Это и есть искомый график (рис. 100). ●

2. График гармонического колебания

Один из наиболее часто встречающихся на практике колебательных процессов описывается формулой

$$y = A \sin(\omega x + \alpha), \tag{1}$$

которая называется *формулой гармонических* (или синусоидальных) *колебаний*. Величина A называется *амплитудой* колебания; она характеризует размах колебания. Величина ω называется *частотой* колебания. Чем больше ω, тем большее число колебаний совершается за единицу времени. Наконец, α называется *начальной фазой* колебания.

Например, если груз, висящий на пружине, вывести из положения равновесия, то он начнет совершать вертикальные колебания. Оказывается, что

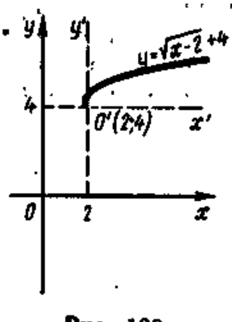


Рис. 100

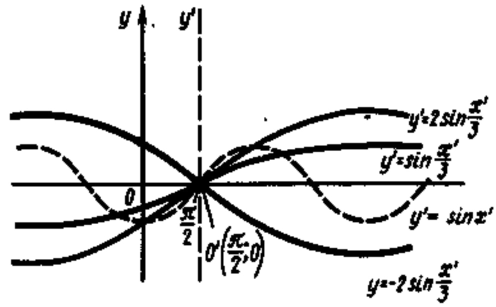


Рис. 101

закон движения выражается формулой (1), где y — отклонение груза от положения равновесия, а x — время. Тот же закон (1) встречается в теории переменного электрического тока.

Пример. Построить график функции $y = 2\sin\left(\frac{x}{3} - \frac{\pi}{6}\right)$.

○ Запишем функцию в виде $y = 2\sin\left(\frac{1}{3}\left(x - \frac{\pi}{2}\right)\right)$. Построение графика выполним в несколько этапов:

1°. Произведем параллельный перенос системы координат, выбрав в качестве начала новой системы точку $O'(\pi/2; 0)$. В системе $x'y'$ нам нужно построить график функции $y' = 2\sin(x'/3)$.

2°. Строим график функции $y' = \sin x'$.

3°. Выполнив сжатие графика к оси y' с коэффициентом 1/3 (т. е. растяжение с коэффициентом 3), получим график функции $y' = \sin(x'/3)$.

4°. Произведем растяжение последнего графика от оси x' с коэффициентом 2. Это и есть график функции $y=2\sin\left(\frac{x}{3}-\frac{\pi}{6}\right)$ (рис. 101).●

В общем случае построение графика функции $y=A\sin(\omega x+\alpha)$ проводится аналогично тому, как это сделано в примере. Прежде всего функция преобразуется к виду $y=A\sin\omega\left(x+\frac{\alpha}{\omega}\right)$. Далее нужно:

1°. Осуществить параллельный перенос системы координат, поместив начало новой системы $x'y'$ в точку $O'(-\alpha/\omega; 0)$,

2°. В системе $x'y'$ построить график функции $y'=\sin x'$ (при этом можно ограничиться одной полуволной),

3°. Произведя сжатие построенного графика к оси y' с коэффициентом ω, построить график функции $y'=\sin\omega x'$,

4°. Произведя растяжение последнего графика от оси x' с коэффициентом A, получить искомый график.

§ 26. Непрерывность элементарных функций

1. Понятие композиции функций. Непрерывность композиции

В математическом анализе приходится рассматривать не только показательную, логарифмическую, степенную, тригонометрические и т. д. функции, но и «функции от функций», например, $y=\sin x^2$, $y=\ln \operatorname{tg} x$ и т. д. — их называют сложными функциями. Более строго понятие сложной функции (или композиции функций) определяется следующим образом.

Определение. **Пусть функция $u=g(x)$ определена на множестве X и U — область ее значений. Далее, пусть функция $y=f(u)$ определена на множестве U. Поставим в соответствие каждому x из X число $f(g(x))$. Тем самым на множестве X будет задана функция $y=f(g(x))$. Она называется *композицией* функций f и g (или *сложной функцией*).**

Теорема 1. *Пусть функция $u=g(x)$ непрерывна в точке x_0 и $g(x_0)=u_0$; пусть функция $y=f(u)$ непрерывна в точке u_0. Тогда сложная функция $y=f(g(x))$ непрерывна в точке x_0.*

□ В доказательстве мы несколько раз воспользуемся определением непрерывности «на языке $\varepsilon-\delta$» (см. определение 3 из § 17).

Возьмем произвольное $\varepsilon>0$. Поскольку функция $y=f(u)$ непрерывна в точке u_0, существует $\lambda>0$ (по причинам, которые станут ясны позже, мы отказываемся здесь от привычного обозначения $\delta>0$) такое, что из неравенства $|u-u_0|<\lambda$ следует неравенство $|f(u)-f(u_0)|<\varepsilon$.

Так как функция $u=g(x)$ непрерывна в точке x_0, то для $\lambda>0$, найденного выше, существует $\delta>0$ такое, что из неравенства $|x-x_0|<\delta$ следует неравенство $|g(x)-g(x_0)|<\lambda$.

Пусть $|x-x_0|<\delta$. Тогда согласно доказанному на втором шаге $|u-u_0|<\lambda$, а из последнего неравенства в силу доказанного на первом шаге следует $|f(u)-f(u_0)|<\varepsilon$. Таким образом, для любого $\varepsilon>0$ существует $\delta>0$ такое, что из $|x-x_0|<\delta$ следует $|f(g(x))-f(g(x_0))|<\varepsilon$, а это и означает непрерывность функции $y=f(g(x))$ в точке x_0. ■

Следствие (**о переходе к пределу под знаком непрерывной функции**). *Если $\lim\limits_{x\to a} g(x)=b$, а функция $y=f(u)$ непрерывна в точке b, то $\lim\limits_{x\to a} f(g(x))=f(b)$.*

Иными словами, $\lim\limits_{x\to a} f(g(x)) = f(\lim\limits_{x\to a} g(x))$.

Примеры. 1. Вычислить $\lim\limits_{x\to\infty} \frac{\sqrt{x^4+x+1}+\sqrt{9x^4-1}}{2x^2+5x+2}$.

○ Применим прием, который мы неоднократно применяли выше для вычисления предела отношения двух многочленов при $x\to\infty$: разделим почленно и числитель, и знаменатель на x в наивысшей из имеющихся степеней, т. е. на x^2. Получим

$$\lim_{x\to\infty} \frac{\sqrt{x^4+x+1}+\sqrt{9x^4-1}}{2x^2+5x+2} = \lim_{x\to\infty} \frac{\frac{\sqrt{x^4+x+1}}{x^2}+\frac{\sqrt{9x^4-1}}{x^2}}{\frac{2x^2}{x^2}+\frac{5x}{x^2}+\frac{2}{x^2}} =$$

$$= \lim_{x\to\infty} \frac{\sqrt{\frac{x^4+x+1}{x^4}}+\sqrt{\frac{9x^4-1}{x^4}}}{2+\frac{5}{x}+\frac{2}{x^2}} = \frac{\lim\limits_{x\to\infty}\sqrt{1+\frac{1}{x^3}+\frac{1}{x^4}}+\lim\limits_{x\to\infty}\sqrt{9-\frac{1}{x^4}}}{\lim\limits_{x\to\infty}\left(2+\frac{5}{x}+\frac{2}{x^2}\right)}.$$

Согласно теореме о переходе к пределу под знаком непрерывной функции, предел корня равен корню из предела; значит,

$$\lim_{x\to\infty}\sqrt{1+\frac{1}{x^3}+\frac{1}{x^4}} = \sqrt{\lim_{x\to\infty}\left(1+\frac{1}{x^3}+\frac{1}{x^4}\right)} = \sqrt{1} = 1,$$

$$\lim_{x\to\infty}\sqrt{9-\frac{1}{x^4}} = \sqrt{\lim_{x\to\infty}\left(9-\frac{1}{x^4}\right)} = \sqrt{9} = 3.$$

В результате получаем

$$\lim_{x\to\infty} \frac{\sqrt{x^4+x+1}+\sqrt{9x^4-1}}{2x^2+5x+2} = \frac{1+3}{2} = 2. \ \bullet$$

2. Доказать непрерывность показательно-степенной функции $y = f(x)^{g(x)}$ в любой точке x, в которой $f(x) > 0$ и функции $f(x)$ и $g(x)$ непрерывны.

○ Имеем $f(x)^{g(x)} = e^{\ln f(x)^{g(x)}} = e^{g(x)\cdot\ln f(x)}$. Так как функции $g(x)$ и $f(x)$ непрерывны, причем $f(x) > 0$, то непрерывны сложная функция $\ln f(x)$ и произведение $h(x) = g(x)\ln f(x)$. Наконец, $e^{h(x)}$ непрерывна по теореме 1. ●

Отсюда, в частности, следует непрерывность в любой точке $x > 0$ степенной функции $y = x^\alpha$ с любым действительным показателем (а не только с рациональным, чем мы ограничивались выше в § 19).

2. Гиперболические функции

В математическом анализе рассматривают функции $y = \frac{e^x+e^{-x}}{2}$ и $y = \frac{e^x-e^{-x}}{2}$; первую из них называют *гиперболическим косинусом* и обозначают $\operatorname{ch} x$, вторую — *гиперболическим синусом* и обозначают $\operatorname{sh} x$. Графики функций $y = \operatorname{ch} x$, $y = \operatorname{sh} x$ изображены на рис. 102, *а* и *б*; кривую, изображенную на рис. 102, *а*, называют *цепной линией* — по дуге этой кривой провисает цепь или канат под действием силы тяжести.

Использование в названии введенных функций терминов «косинус» и «синус» не случайно — эти функции по своим свойствам во многом аналогичны обычным (круговым) функциям $y = \cos x$, $y = \sin x$ (на самом деле здесь имеется более глубокая аналогия, но на этом мы не будем останавливаться). Приведем ряд этих свойств.

1°. *Функция* $\operatorname{ch} x$ *— четная функция,* $\operatorname{sh} x$ *— нечетная.*

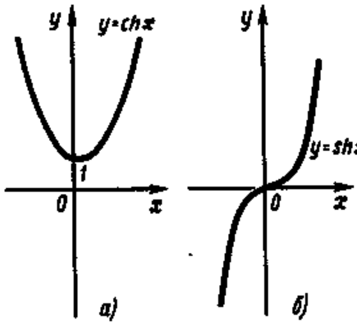


Рис. 102

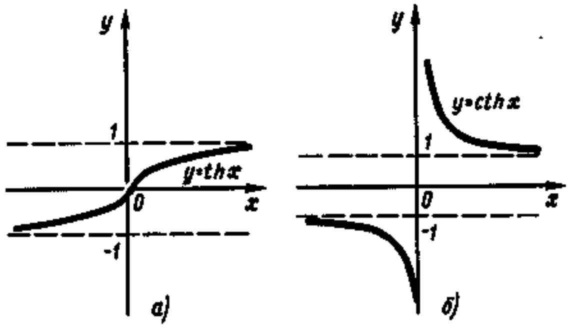


Рис. 103

□ Имеем $\operatorname{ch}(-x) = \frac{e^{-x}+e^{x}}{2} = \operatorname{ch} x$;
$\operatorname{sh}(-x) = \frac{e^{-x}-e^{x}}{2} = -\frac{e^{x}-e^{-x}}{2} = -\operatorname{sh} x$ ■

2^0. *Справедливо равенство* $\operatorname{sh}2x = 2\operatorname{sh}x \cdot \operatorname{ch}x$.

□ Имеем

$$2\operatorname{sh}x \cdot \operatorname{ch}x = 2 \cdot \frac{e^{x}-e^{-x}}{2} \cdot \frac{e^{x}+e^{-x}}{2} = \frac{e^{2x}-e^{-2x}}{2} = \operatorname{sh}2x. \blacksquare$$

Аналогично можно доказать и такие равенства (сделайте это самостоятельно):

3^0. $\operatorname{ch}2x = \operatorname{ch}^2 x + \operatorname{sh}^2 x$.

4^0. $\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$.

Наряду с гиперболическими синусом и косинусом рассматривают также гиперболические тангенс и котангенс: отношение $\frac{\operatorname{sh}x}{\operatorname{ch}x}$ называют *гиперболическим тангенсом* и обозначают $\operatorname{th}x$, отношение $\frac{\operatorname{ch}x}{\operatorname{sh}x}$ называют *гиперболическим котангенсом* и обозначают $\operatorname{cth}x$. Эти функции по своим свойствам напоминают обычные $\operatorname{tg}x$ и $\operatorname{ctg}x$. Так, обе они нечетные; справедливы равенства

$$1 - \operatorname{th}^2 x = \frac{1}{\operatorname{ch}^2 x}, \quad \operatorname{cth}^2 x - 1 = \frac{1}{\operatorname{sh}^2 x}.$$

Графики функций $y = \operatorname{th}x$ и $y = \operatorname{cth}x$ изображены на рис. 103, *а* и *б*.

3. Элементарные функции и их непрерывность

К элементарным функциям обычно относят функции, изученные нами в предыдущих параграфах: это постоянная, степенная, показательная, логарифмическая, тригонометрические, обратные тригонометрические, гиперболические функции, а также функции, полученные из перечисленных с помощью конечного числа арифметических операций и композиций (сложные функции). Исследование элементарных функций на непрерывность, проведенное в предыдущих параграфах, и теорема о непрерывности сложной функции позволяют сформулировать в качестве итогового вывода одно из наиболее важных положений математического анализа.

Теорема 2. *Любая элементарная функция непрерывна в любой точке своей области определения.*

§ 27. Техника вычисления пределов функций

1. Простейшие случаи раскрытия неопределенностей

Если функции $P(x)$ и $Q(x)$ непрерывны в точке $x=a$, причем $Q(a)\neq a$, то $\lim\limits_{x\to a}\frac{P(x)}{Q(x)}=\frac{P(a)}{Q(a)}$. Если $Q(a)=0$, $P(a)\neq 0$, то $\lim\limits_{x\to a}\frac{P(x)}{Q(x)}=\infty$. Если же и $P(a)=0$, и $Q(a)=0$, то $\lim\limits_{x\to a}\frac{P(x)}{Q(x)}$, как мы увидим ниже, может быть равным любому числу, а также ∞. Выражения вида $\lim\limits_{x\to a}\frac{P(x)}{Q(x)}$, где $\lim\limits_{x\to a}P(x)=\lim\limits_{x\to a}Q(x)=0$, называют *неопределенностями вида* 0/0.

Для вычисления таких пределов (или, как говорят, для раскрытия неопределенностей) применяют следующий метод: заменяют дробь $\frac{P(x)}{Q(x)}$ дробью $\frac{P_1(x)}{Q_1(x)}$, тождественно равной $\frac{P(x)}{Q(x)}$ при $x\neq a$, но такой, что хотя бы одно из чисел $P_1(a)$, $Q_1(a)$ было отлично от нуля. Так как значение функции в точке $x=a$ не играет роли при вычислении пределов (см. п. 2 § 16), то $\lim\limits_{x\to a}\frac{P(x)}{Q(x)}=\lim\limits_{x\to a}\frac{P_1(x)}{Q_1(x)}$.

Пример. Вычислить предел $\lim\limits_{x\to 4}\frac{\sqrt{x+5}-3}{x^2-16}$.

○ Здесь имеем неопределенность вида 0/0. Для ее раскрытия умножим числитель и знаменатель на выражение, сопряженное числителю, что позволит избавиться от иррациональности в числителе, а затем сократим дробь. Получим

$$\lim_{x\to 4}\frac{\sqrt{x+5}-3}{x^2-16}=\lim_{x\to 4}\frac{(\sqrt{x+5}-3)(\sqrt{x+5}+3)}{(x^2-16)(\sqrt{x+5}+3)}=\lim_{x\to 4}\frac{(\sqrt{x+5})^2-9}{(x+4)(x-4)(\sqrt{x+5}+3)}=$$

$$=\lim_{x\to 4}\frac{x-4}{(x+4)(x-4)(\sqrt{x+5}+3)}=\lim_{x\to 4}\frac{1}{(x+4)(\sqrt{x+5}+3)}=\frac{1}{(4+4)(\sqrt{4+5}+3)}=$$

$$=\frac{1}{48}. \bullet$$

2. Первый замечательный предел

Теорема 1. *Справедливо равенство*

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

(первый замечательный предел).

□ Так как $\frac{\sin x}{x}$ — четная функция, то достаточно рассмотреть случай $x>0$. Выше (см. п. 2 § 22) мы доказали, что при $0<x<\pi/2$ справедливо неравенство $\sin x<x<\operatorname{tg} x$. Разделив все части этого двойного неравенства на $\sin x$ (учитывая, что $\sin x>0$ при $0<x<\pi/2$), получим

$$1<\frac{x}{\sin x}<\frac{1}{\cos x},\ \text{откуда}\ \cos x<\frac{\sin x}{x}<1.$$

Устремим x к нулю. Так как $y=\cos x$ — всюду непрерывная функция, то $\lim\limits_{x\to 0}\cos x=\cos 0=1$. С другой стороны, предел постоянной функции равен значению этой постоянной, т. е. $\lim\limits_{x\to 0}1=1$. Тогда по теореме о пределе промежуточной функции (см. теорему 9 § 16) получим $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$. ■

Заметим, что справедливо более общее равенство:

$$\lim_{x\to 0}\frac{\sin kx}{kx}=1\,,\ \text{где}\ k\neq 0\,. \qquad (1)$$

Примеры. 1. Вычислить предел $\lim\limits_{x\to 0}\frac{\sin 3x}{4x}$.

○ Перепишем этот предел так:

$$\lim_{x\to 0}\frac{\sin 3x}{4x}=\lim_{x\to 0}\frac{\sin 3x}{3x}\cdot\frac{3x}{4x}=\frac{3}{4}\lim_{x\to 0}\frac{\sin 3x}{3x}.$$

Поскольку $\lim\limits_{x\to 0}\frac{\sin 3x}{3x}=1$, находим $\lim\limits_{x\to 0}\frac{\sin 3x}{4x}=\frac{3}{4}$. ●

2. Вычислить предел $\lim\limits_{x\to 0}\frac{1-\cos x}{x^2}$.

○ Так как $1-\cos x=2\sin^2\frac{x}{2}$, а $\lim\limits_{x\to 0}\frac{\sin\frac{x}{2}}{\frac{x}{2}}=1$, то

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

3. Вычислить предел $\lim\limits_{x\to\pi}\frac{\sin 3x}{2(\pi-x)}$.

○ Положим $\pi-x=t$. Тогда $x=\pi-t$ и $\lim\limits_{x\to\pi}t=\lim\limits_{x\to\pi}(\pi-x)=0$. Поэтому

$$\lim_{x\to\pi}\frac{\sin 3x}{2(\pi-x)}=\lim_{t\to 0}\frac{\sin 3(\pi-t)}{2t}=\lim_{t\to 0}\frac{\sin 3t}{2t}=\lim_{t\to 0}\frac{\sin 3t}{3t}\cdot\frac{3t}{2t}=\frac{3}{2}.\ ●$$

3. Порядок бесконечно малых. Эквивалентные бесконечно малые

Если α и β — бесконечно малые функции при $x\to a$, то $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}$ (неопределенность вида $\frac{0}{0}$) может быть равен либо нулю, либо бесконечности, либо какому-нибудь числу, отличному от нуля; наконец, предел может не существовать.

Если $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}=0$, то при стремлении x к a функция α стремится быстрее к нулю, чем β. Говорят, что α — *бесконечно малая более высокого порядка, чем* β.

Если $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}=\infty$, то α — называют *бесконечно малой более низкого порядка, чем* β.

Если $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}=c\neq 0$, то α и β называют *бесконечно малыми одного и того же порядка.*

Особенно важен частный случай, когда $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}=1$. В этом случае α и β называют *эквивалентными бесконечно малыми* (при $x\to a$) и пишут $\alpha\sim\beta$.

Если $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}$ не существует, то α и β называют *несравнимыми бесконечно малыми.*

Функция $x-a$ при $x\to a$ называется *бесконечно малой первого порядка.*

Если $\lim\limits_{x\to a}\frac{\alpha(x)}{(x-a)^k}=c\neq 0$, то бесконечно малую α называют *бесконечно малой порядка k*.

Примеры. 1. Если $n>m$, то $\lim\limits_{x\to 0}\frac{x^n}{x^m}=\lim\limits_{x\to 0}x^{n-m}=0$, т. е. при $x\to 0$ функция x^n есть бесконечно малая более высокого порядка, чем x^m (x^n — бесконечно малая n-го порядка, x^m — m-го порядка).

2. $\lim\limits_{x\to 0}\frac{1-\cos x}{x^2}=\frac{1}{2}$ (см. пример 2 п. 2). Поэтому $1-\cos x$ и x^2 — бесконечно малые одного и того же порядка при $x\to 0$. Отсюда следует, что $1-\cos x$ — бесконечно малая второго порядка по сравнению с x.

Из равенств

$$\lim_{x\to 0}\frac{\sin x}{x}=1,\ \lim_{x\to 0}\frac{\operatorname{tg} x}{x}=1,\ \lim_{x\to 0}\frac{1-\cos x}{\frac{1}{2}x^2}=1$$

следует, что *при $x\to 0$ функции $\sin x$, $\operatorname{tg} x$ являются бесконечно малыми, эквивалентными x, а функция $1-\cos x$ эквивалентна функции $\frac{1}{2}x^2$*:

$$\sin x\sim x, \tag{2}$$

$$\operatorname{tg} x\sim x, \tag{3}$$

$$1-\cos x\sim\frac{1}{2}x^2. \tag{4}$$

Справедливо более общее утверждение: если $f(x)$ — бесконечно малая функция при $x\to a$, то

$$\sin f(x)\sim f(x),\ \operatorname{tg} f(x)\sim f(x),\ 1-\cos f(x)\sim\frac{1}{2}f^2(x).$$

При вычислении пределов весьма полезной оказывается следующая теорема.

Теорема 2. *Если при $x\to a$ функции α, β, α_1, β_1 являются бесконечно малыми и $\alpha\sim\alpha_1$, $\beta\sim\beta_1$, то $\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)}=\lim\limits_{x\to a}\frac{\alpha_1(x)}{\beta_1(x)}$, либо оба предела не существуют.*

□ Согласно условию, имеем $\lim\limits_{x\to a}\frac{\alpha(x)}{\alpha_1(x)}=1$, $\lim\limits_{x\to a}\frac{\beta(x)}{\beta_1(x)}=1$. Поэтому

$$\lim_{x\to a}\frac{\alpha(x)}{\beta(x)}=\lim_{x\to a}\frac{\alpha(x)}{\alpha_1(x)}\cdot\frac{\alpha_1(x)}{\beta_1(x)}\cdot\frac{\beta_1(x)}{\beta(x)}=$$

$$=\lim_{x\to a}\frac{\alpha(x)}{\alpha_1(x)}\cdot\lim_{x\to a}\frac{\alpha_1(x)}{\beta_1(x)}\cdot\lim_{x\to a}\frac{\beta_1(x)}{\beta(x)}=1\cdot\lim_{x\to a}\frac{\alpha_1(x)}{\beta_1(x)}\cdot 1=\lim_{x\to a}\frac{\alpha_1(x)}{\beta_1(x)}.\ ■$$

Таким образом, при вычислении предела отношения бесконечно малых можно заменять и числитель, и знаменатель эквивалентными бесконечно малыми функциями.

Пример 3. Вычислить предел $\lim\limits_{x\to 0}\operatorname{ctg} 5x\cdot\sin 3x$.

○ Имеем $\operatorname{ctg} 5x=\frac{1}{\operatorname{tg} 5x}$. Так как $\sin 3x\sim 3x$, $\operatorname{tg} 5x\sim 5x$ при $x\to 0$, то

$$\lim_{x\to 0}\operatorname{ctg} 5x\cdot\sin 3x=\lim_{x\to 0}\frac{\sin 3x}{\operatorname{tg} 5x}=\lim_{x\to 0}\frac{3x}{5x}=\frac{3}{5}.\ ●$$

Замечание. Так как $\lim\limits_{x\to 0}\operatorname{ctg} 5x=\infty$, $\lim\limits_{x\to 0}\sin 3x=0$, то заданный предел относится к случаю, который называется *неопределенностью вида* $\infty\cdot 0$. Такую неопределенность всегда можно свести к неопределенности вида $0/0$, что мы и сделали в рассмотренном примере.

4. Второй замечательный предел

Выше (см. п. 3 § 15) мы отмечали, что справедливо равенство

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

Если положить в нем $t=1/x$, то оно примет вид

$$\lim_{t\to 0}(1+t)^{1/t}=e. \qquad (5)$$

Этот предел называют *вторым замечательным пределом.*

Заметим, что $\lim\limits_{t\to 0}(1+t)=1$, а $\lim\limits_{t\to 0}\frac{1}{t}=\infty$. Поэтому предел, написанный в левой части равенства (5), относят обычно к *неопределенностям вида* 1^{∞}. Раскрытие подобных неопределенностей, как правило, связано с использованием второго замечательного предела.

Примеры. 1. Вычислить $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^{5x}$.

○ Находим $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^{5x}=\lim\limits_{x\to\infty}\left(\left(1+\frac{1}{x}\right)^{x}\right)^{5}=e^{5}$. ●

2. Вычислить $\lim\limits_{x\to\infty}\left(\frac{2x+1}{2x-1}\right)^{4x-3}$.

○ Имеем $\frac{2x+1}{2x-1}=\frac{(2x-1)+2}{2x-1}=1+\frac{2}{2x-1}$. Преобразуем выражение под знаком предела следующим образом:

$$\left(\frac{2x+1}{2x-1}\right)^{4x-3}=\left(\left(1+\frac{2}{2x-1}\right)^{\frac{2x-1}{2}}\right)^{\frac{2}{2x-1}\cdot(4x-3)}$$

и положим $f(x)=\left(1+\frac{2}{2x-1}\right)^{\frac{2x-1}{2}}$, $g(x)=\frac{2(4x-3)}{2x-1}$.

Вычислим $\lim\limits_{x\to\infty}f(x)$ и $\lim\limits_{x\to\infty}g(x)$. Для нахождения первого предела введем новую переменную $t=\frac{2}{2x-1}$. Если $x\to\infty$, то $t\to 0$, и мы получаем

$$\lim_{x\to\infty}f(x)=\lim_{x\to\infty}\left(1+\frac{2}{2x-1}\right)^{\frac{2x-1}{2}}=\lim_{t\to 0}(1+t)^{1/t}=e.$$

Найдем второй предел:

$$\lim_{x\to\infty}g(x)=\lim_{x\to\infty}\frac{2(4x-3)}{2x-1}=4.$$

Так как $\lim\limits_{x\to\infty}f(x)^{g(x)}=\lim\limits_{x\to\infty}f(x)^{\lim\limits_{x\to\infty}g(x)}$, то

$$\lim_{x\to\infty}f(x)^{g(x)}=\lim_{x\to\infty}\left(\frac{2x-1}{2x-1}\right)^{4x-3}=e^{4}. \ ●$$

Теорема 3. *Справедливо соотношение*

$$\lim_{t\to 0}\frac{\ln(1+t)}{t}=1. \qquad (6)$$

□ Имеем

$$1=\ln e=\ln\left(\lim_{t\to 0}(1+t)^{1/t}\right)=\lim_{t\to 0}\ln(1+t)^{1/t}=\lim_{t\to 0}\frac{1}{t}\cdot\ln(1+t)=\lim_{t\to 0}\frac{\ln(1+t)}{t}.$$

При переходе от $\ln\left(\lim_{t\to 0}(1+t)^{1/t}\right)$ к $\lim_{t\to 0}\ln(1+t)^{1/t}$ мы пользовались непрерывностью логарифмической функции. ■

Равенство (6) означает, что *при* $t\to 0$ *функции* $\ln(1+t)$ *и* t *являются эквивалентными бесконечно малыми:* $\ln(1+t)\sim t$, если $t\to 0$.

Введем новую переменную z, полагая $\ln(1+t)=z$. Тогда $(1+t)=e^z$, $t=e^z-1$. Поэтому

$$\lim_{t\to 0}t=\lim_{z\to 0}(e^z-1)=0,$$

и из формулы (6) следует, что

$$\lim_{z\to 0}\frac{z}{e^z-1}=1. \tag{7}$$

Мы доказали, что *при* $z\to 0$ *функции* e^z-1 *и* z *— эквивалентные бесконечно малые:* $e^z-1\sim z$, если $z\to 0$.

В частности, так как $a^x=e^{x\ln a}$, то $e^{x\ln a}-1\sim x\ln a$. Иными словами, получаем, что $a^x-1\sim x\ln a$, если $x\to 0$.

Наконец, докажем, что

$$\lim_{x\to 0}\frac{(1+x)^\alpha-1}{x}=\alpha. \tag{8}$$

Для этого заметим, что $1+x=e^{\ln(1+x)}$, и потому $(1+x)^\alpha=e^{\alpha\ln(1+x)}$. Значит,

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

Равенство (8) означает, что $(1+x)^\alpha-1\sim\alpha x$, если $x\to 0$.

Примеры. 3. Вычислить предел $\lim_{x\to 0}\frac{\ln(1+\operatorname{tg}2x)}{e^{\sin 3x}-1}$.

○ Так как $\ln(1+x)\sim x$, а $\operatorname{tg}2x$ — бесконечно малая при $x\to 0$, то $\ln(1+\operatorname{tg}2x)\sim\operatorname{tg}2x$ при $x\to 0$. Точно так же из $e^x-1\sim x$ получаем $e^{\sin 3x}-1\sim\sin 3x$. Поэтому, заменяя числитель и знаменатель дроби эквивалентными бесконечно малыми, имеем

$$\lim_{x\to 0}\frac{\ln(1+\operatorname{tg}2x)}{e^{\sin 3x}-1}=\lim_{x\to 0}\frac{\operatorname{tg}2x}{\sin 3x}.$$

Но $\operatorname{tg}2x\sim 2x$, $\sin 3x\sim 3x$ при $x\to 0$ и потому

$$\lim_{x\to 0}\frac{\operatorname{tg}2x}{\sin 3x}=\lim_{x\to 0}\frac{2x}{3x}=\frac{2}{3}.$$

Итак, искомый предел равен 2/3. ●

4. Вычислить предел $\lim_{x\to 0}\frac{\sqrt[3]{1+2x}-1}{\sqrt[4]{1+\sin 3x}-1}$.

○ Числитель запишем в виде $(1+2x)^{1/3}-1$, а знаменатель — в виде $(1+\sin 3x)^{1/4}-1$. Используя соотношение $(1+x)^\alpha-1\sim\alpha x$, получаем

$$(1+2x)^{\frac{1}{3}}-1\sim\frac{1}{3}\cdot 2x,\ (1+\sin 3x)^{\frac{1}{4}}-1\sim\frac{1}{4}\sin 3x.$$

Значит,

$$\lim_{x\to 0}\frac{\sqrt[3]{1+2x}-1}{\sqrt[4]{1+\sin 3x}-1}=\lim_{x\to 0}\frac{\frac{2}{3}x}{\frac{1}{4}\sin 3x}=\lim_{x\to 0}\frac{8x}{3\cdot 3x}=\frac{8}{9}.\ ●$$

5

Производная и ее приложения

§ 28. Производная

1. Задача о проведении касательной к графику функции

Прежде всего выясним, что именно мы будем понимать под касательной к произвольной плоской кривой. Касательную нельзя определить как прямую, имеющую с кривой только одну общую точку. В самом деле, ось параболы $y = ax^2$ имеет с ней лишь одну общую точку, но не касается параболы (см. рис. 18). В то же время прямая $y = 1$ имеет бесконечное множество общих точек с синусоидой $y = \sin x$ (см. рис. 81), но касается синусоиды в каждой из этих точек. Нельзя принять за определение касательной и требование, чтобы кривая располагалась по одну сторону от прямой, — ось абсцисс касается кривой $y = ax^3$ в точке $(0; 0)$, хотя в этой точке кривая пересекает ось абсцисс (см. рис. 20).

Чтобы дать определение касательной, придется использовать понятие предела. Пусть Γ — дуга некоторой кривой и M_0 — точка этой кривой. Проведем через точку M_0 секущую M_0N. Если точка N приближается по кривой к точке M_0, то секущая M_0N поворачивается вокруг точки M_0. Может случиться так, что по мере приближения точки N к M_0 секущая стремится к некоторому предельному положению M_0T. Тогда M_0T называют *касательной* к кривой Γ в точке M_0 (рис. 104).

Найдем угловой коэффициент касательной для случая, когда кривая Γ является графиком некоторой функции $y = f(x)$. Пусть M_0 — точка графика с абсциссой a и ординатой $f(a)$. Предположим, что касательная к кривой Γ в точке M_0 существует. Возьмем на этой кривой еще одну точку $N(x; y)$ (рис. 105) и проведем через точки M_0 и N прямую; пусть φ — угол наклона этой секущей к положительному направлению оси x. Тогда

$$\operatorname{tg} \varphi = \frac{f(x) - f(a)}{x - a}.$$

Угол наклона касательной к оси абсцисс обозначим через θ. Таким образом,

$$k_{\text{кас}} = \operatorname{tg} \theta = \lim_{x \to a} \operatorname{tg} \varphi = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

Выше мы условились называть разность $f(x) - f(a)$ приращением функции в точке a и обозначать Δf или Δy. Аналогично разность $x - a$ называют *при-*

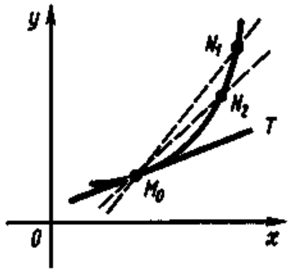


Рис. 104

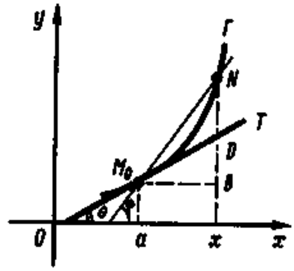


Рис. 105

ращением аргумента и обозначают Δx; условие $x \to a$ эквивалентно условию $\Delta x \to 0$. Значит,

$$k_{\text{кас}} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Итак, для того чтобы можно было провести невертикальную касательную к графику функции $y = f(x)$ в точке с абсциссой a, необходимо и достаточно, чтобы в этой точке существовал предел $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, причем этот предел равен угловому коэффициенту касательной.

2. Задача о мгновенной скорости прямолинейного движения

Пусть по прямой, на которой выбраны начало отсчета, единица измерения и направление, движется точка по закону $s = s(t)$ ($s(t)$ — координата точки на прямой в момент времени t). Под *средней скоростью* движения за некоторый промежуток времени в физике понимают отношение перемещения к промежутку времени, т. е. средняя скорость за промежуток времени от t_1 до t_2 выражается равенством $v_{\text{ср}} = \frac{s(t_2) - s(t_1)}{t_2 - t_1}$. Полагая $t_1 = t$, $t_2 = t + \Delta t$, получим

$$v_{\text{ср}} = \frac{s(t + \Delta t) - s(t)}{\Delta t} = \frac{\Delta s}{\Delta t}.$$

Мгновенной скоростью в момент времени t (точнее, числовым значением скорости в момент времени t) называют предел средней скорости движения за промежуток времени $[t, t + \Delta t]$ при условии $\Delta t \to 0$. Таким образом,

$$v_{\text{мгн}} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}.$$

3. Определение производной, ее геометрический и физический смысл

Две различные задачи, рассмотренные в предыдущих пунктах, привели в процессе решения к одному и тому же результату (или, как чаще говорят в математике, к одной и той же математической модели) — пределу отношения приращения функции к вызвавшему его приращению аргумента при условии, что приращение аргумента стремится к нулю. Оказывается, что многие задачи приводят к необходимости вычисления такого же предела, поэтому есть смысл специально заняться его изучением.

Определение. Если для функции $y=f(x)$ в фиксированной точке x существует предел отношения приращения функции к приращению аргумента Δx при условии $\Delta x \to 0$, то этот предел называется *значением производной* функции $y=f(x)$ в точке x и обозначается $f'(x)$ или y'. Итак,

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

Отметим, что $f'(x)$ — это новая функция, определенная во всех таких точках x, в которых существует указанный предел; эту функцию называют *производной* функции $y=f(x)$.

Используя это определение, можно рекомендовать следующий план отыскания производной для функции $y=f(x)$.

АЛГОРИТМ
нахождения производной функции $y=f(x)$

1^0. Зафиксировав значение x, найти $f(x)$.
2^0. Придав аргументу x приращение Δx так, чтобы не выйти из области определения функции $f(x)$, найти $f(x+\Delta x)$.
3^0. Вычислить приращение функции $\Delta f=f(x+\Delta x)-f(x)$.
4^0. Составить отношение $\frac{\Delta f}{\Delta x}$.
5^0. Найти предел отношения $\frac{\Delta f}{\Delta x}$ при $\Delta x\to 0$.

Примеры. 1. Найти производную функции $y=kx+b$.

○ Воспользуемся алгоритмом нахождения производной.

1^0. Зафиксируем x; $f(x)=kx+b$.

2^0. Дадим аргументу приращение Δx; тогда $f(x+\Delta x)=k(x+\Delta x)+b=kx+k\Delta x+b$.

3^0. $\Delta y=f(x+\Delta x)-f(x)=(kx+k\Delta x+b)-(kx+b)=k\Delta x$.

4^0. $\frac{\Delta y}{\Delta x}=\frac{k\Delta x}{\Delta x}=k$.

5^0. $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}k=k$.

Итак, $(kx+b)'=k$.

В частности, для постоянной функции $y=b$ (здесь $k=0$) имеем $(b)'=0$; для функции $y=x$ (здесь $k=1$) имеем $(x)'=1$.●

2. Найти производную функции $y=x^2$.

○ 1^0. $f(x)=x^2$.

2^0. $f(x+\Delta x)=(x+\Delta x)^2=x^2+2x\Delta x+(\Delta x)^2$.

3^0. $\Delta y=f(x+\Delta x)-f(x)=(x^2+2x\Delta x+(\Delta x)^2)-x^2=2x\Delta x+(\Delta x)^2=\Delta x(2x+\Delta x)$.

4^0. $\frac{\Delta y}{\Delta x}=\frac{\Delta x(2x+\Delta x)}{\Delta x}=2x+\Delta x$.

5^0. $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}(2x+\Delta x)=2x$.

Итак, $(x^2)'=2x$. ●

3. Найти производную функции $y=1/x$ в точке $x\neq 0$.

○ Не ограничивая общности рассуждения, будем считать, что $x>0$ и $x+\Delta x>0$. Имеем:

1^0. $f(x)=\frac{1}{x}$.

2^0. $f(x+\Delta x)=\frac{1}{x+\Delta x}$.

3^0. $\Delta y = f(x+\Delta x) - f(x) = \frac{1}{x+\Delta x} - \frac{1}{x} = -\frac{\Delta x}{x(x+\Delta x)}$.

4^0. $\frac{\Delta y}{\Delta x} = -\frac{\Delta x}{x(x+\Delta x)\Delta x} = -\frac{1}{x^2+x\Delta x}$.

5^0. $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\frac{1}{\lim\limits_{\Delta x \to 0}(x^2+x\Delta x)} = -\frac{1}{x^2}$.

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

Выше мы установили, что если в точке $M(a;\ f/a)$ к графику функции $y=f(x)$ можно провести касательную, то угловой коэффициент касательной равен $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, а этот предел есть значение производной функции $y=f(x)$ в точке a. Таким образом, *угловой коэффициент касательной к графику функции $y=f(x)$ в точке $x=a$ равен значению производной в этой точке:* $k_{кас}=f'(a)$. В этом состоит геометрический смысл производной. Физический смысл производной фактически был установлен нами в п. 2: *при прямолинейном движении, описываемом функцией $s=s(t)$, числовое значение скорости в момент времени t равно значению производной: $v=s'(t)$.*

Понятие производной позволяет определить не только мгновенную скорость прямолинейного движения, но и мгновенную скорость протекания других физических процессов. Рассмотрим, например, вращение тела вокруг оси. Оно задается функцией $y=\varphi(t)$, описывающей угол поворота тела за промежуток времени $[0,\ t]$. Если это вращение происходит равномерно, то его *угловой скоростью* называют отношение угла поворота к величине промежутка времени: $\omega = \frac{\varphi(t)}{t}$. Если же вращение происходит неравномерно, то *средней угловой скоростью* этого вращения за промежуток времени $[t,\ t+\Delta t]$ называют отношение $\frac{\varphi(t+\Delta t)-\varphi(t)}{\Delta t}$ т. е. $\omega_{ср} = \frac{\Delta\varphi}{\Delta t}$.

Мгновенной угловой скоростью в момент времени t называют предел средней угловой скорости за промежуток времени $[t,\ t+\Delta t]$ при условии $\Delta t \to 0$:

$$\omega_{мгн} = \lim_{\Delta t \to 0} \omega_{ср} = \lim_{\Delta t \to 0} \frac{\Delta\varphi}{\Delta t}.$$

Таким образом, $\omega = \varphi'(t)$.

Точно так же можно установить, что мгновенная скорость радиоактивного распада выражается формуой $v=m'(t)$, где $m(t)$ — масса вещества в момент времени t.

Обобщая все сказанное, введем понятие *средней скорости изменения функции* на промежутке $[x,\ x+\Delta x]$ — под этим понимают отношение $\frac{f(x+\Delta x)-f(x)}{\Delta x}$, т. е. $\frac{\Delta y}{\Delta x}$. Тогда $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = y'$ — скорость изменения y относительно x в данной точке («мгновенная скорость» изменения функции).

4. Уравнение касательной

Пусть функция $y=f(x)$ имеет производную в точке a и $M(a;\ f(a))$ — точка графика этой функции. Составим уравнение касательной к графику функции в заданной точке.

Будем искать это уравнение в виде $y=kx+b$. Поскольку искомая прямая проходит через точку $M(a;\ f(a))$, выполняется равенство $f(a)=ka+b$,

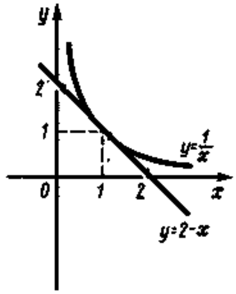


Рис. 106

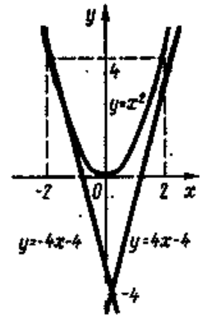


Рис. 107

откуда $b = f(a) - ka$. Поэтому уравнение касательной есть $y = kx + (f(a) - ka)$ или $y = f(a) + k(x-a)$. Далее, так как угловой коэффициент k касательной равен $f'(a)$, то уравнение касательной к графику функции $y = f(x)$ в точке $M(a; f(a))$ имеет вид

$$y = f(a) + f'(a)(x-a). \qquad (1)$$

Примеры. 1. Составить уравнение касательной к гиперболе $y = 1/x$ в точке с абсциссой $x = 1$.

○ Здесь $a = 1$ (абсцисса точки касания), $f(a) = 1$, $f'(x) = -1/x^2$, $f'(a) = f'(1) = -1$. Подставив найденные значения $a = 1$, $f(a) = 1$, $f'(a) = -1$ в уравнение (1), получим $y = 1 - (x-1)$, т. е. $y = 2 - x$ (рис. 106).●

2. Из точки $(0; -4)$ провести касательную к параболе $y = x^2$.

○ Пусть a — абсцисса точки касания. Тогда $f(a) = a^2$, $f'(x) = 2x$, $f'(a) = 2a$. Воспользовавшись уравнением (1), получим

$$y = a^2 + 2a(x-a),$$

т. е.

$$y = 2ax - a^2. \qquad (2)$$

По условию касательная должна проходить через точку $(0; -4)$; подставив вместо x и y в уравнение (2) значения 0 и -4, имеем $-4 = 2a \cdot 0 - a^2$, откуда $a_1 = 2$; $a_2 = -2$.

Если $a = 2$, то уравнение (2) примет вид $y = 4x - 4$; если $a = -2$, то оно примет вид $y = -4x - 4$.

Итак, получили две касательные, удовлетворяющие заданному условию: $y = 4x - 4$ и $y = -4x - 4$ (рис. 107).●

§ 29. Дифференциал

1. Дифференцируемость функции в точке

Определение 1. Функция $y = f(x)$ называется *дифференцируемой* **в точке** x, если ее приращение Δy представимо в виде $\Delta y = A\Delta x + \alpha\Delta x$, где A — число, а α — бесконечно малая функция при $\Delta x \to 0$.

Для линейной функции $y = kx + b$ нетрудно убедиться в том, что выполняется равенство $\Delta y = k\Delta x$, т. е. приращение функции пропорционально приращению аргумента. Равенство же $\Delta y = A\Delta x + \alpha\Delta x$ означает, что приращение функции «почти пропорционально» приращению аргумента: $\Delta y \approx A\Delta x$; это равенство тем точнее, чем меньше $|\Delta x|$. Геометрически это означает, что дифференцируемая в точке x функция «линейна в малом», т. е. ее график в

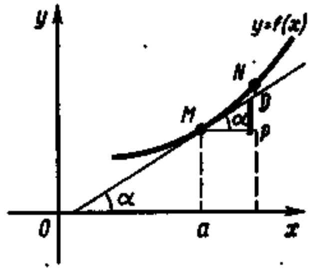


Рис. 108

некоторой достаточно малой окрестности точки x почти сливается с некоторой невертикальной прямой, т. е. с графиком некоторой линейной функции. Таким образом, с геометрической точки зрения дифференцируемость функции в точке x означает возможность «спрямления» графика в некоторой достаточно малой окрестности точки x. Отсюда, например, вытекает недифференцируемость в точке $x = 0$ такой функции, как $y = |x|$ (см. рис. 23) — этот график в некоторой окрестности точки $x = 0$ «спрямить» нельзя.

Теорема 1. *Для того чтобы функция $y = f(x)$ была дифференцируемой в точке x, необходимо и достаточно, чтобы в этой точке существовала производная.*

□ Необходимость. Пусть существует $f'(x)$. По определению, $f'(x) = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, но тогда $\frac{\Delta y}{\Delta x} = f'(x) + \alpha$, где α — бесконечно малая при $\Delta x \to 0$ (см. п. 2 § 16). Значит, $\Delta y = f'(x)\Delta x + \alpha\Delta x$ или $\Delta y = A\Delta x + \alpha\Delta x$, где $A = f'(x)$.

Достаточность. Пусть $\Delta y = A\Delta x + \alpha\Delta x$, где α — бесконечно малая при $\Delta x \to 0$. Тогда $\frac{\Delta y}{\Delta x} = A + \alpha$, а потому $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = A$. Значит, производная в точке существует, причем $f'(x) = A$. ■

Данная теорема означает, что для функции одной переменной дифференцируемость равнозначна существованию производной, поэтому, в частности, операцию отыскания производной называют *дифференцированием* функции, а раздел математического анализа, в котором изучается производная, — *дифференциальным исчислением*.

2. Дифференциал функции в точке

Определение 2. Пусть функция $y = f(x)$ дифференцируема в точке x, т. е. $\Delta y = f'(x)\Delta x + \alpha\Delta x$, где α — бесконечно малая при $\Delta x \to 0$. Тогда главная линейная часть приращения функции, т. е. $f'(x)\Delta x$, называется *дифференциалом* функции $f(x)$ в точке x и обозначается $d(f(x))$ или dy: $dy = f'(x)\Delta x$.

Например, для функции $y = x^2$ имеем $dy = 2x\Delta x$, для функции $y = x^3$ имеем $dy = 3x^2\Delta x$.

Пусть функция $y = f(x)$ дифференцируема в точке a; тогда $dy = f'(a)\Delta x$. Проведем касательную к графику функции в точке a (рис. 108); имеем $\operatorname{tg}\alpha = f'(a)$. Из треугольника MDP находим $DP = MP \operatorname{tg}\alpha = \Delta x f'(a) = dy$. Значит, *$dy$ есть приращение ординаты касательной при переходе от точки a к точке $a + \Delta x$* (в то время как Δy, т. е. NP на рис. 108, есть приращение ординаты графика функции). В этом состоит геометрический смысл дифференциала.

Если $f(x)=x$, то $f'(x)=1$ и потому $d(f(x))=1\cdot\Delta x$, т. е. $dx=\Delta x$. Учитывая это, приращение аргумента обычно обозначают через dx, и основная формула для дифференциала принимает вид

$$dy=f'(x)dx \quad \text{или} \quad dy=y'dx,$$

откуда $y'=\frac{dy}{dx}$; выражение $\frac{dy}{dx}$ часто используется в качестве обозначения производной (читается «дэ игрек по дэ икс»).

3. Использование дифференциала в приближенных вычислениях

Основная роль дифференциала в математическом анализе и в его приложениях заключается в том, что для дифференцируемой в точке a функции $y=f(x)$ справедливо приближенное равенство $\Delta y\approx dy$ (к вопросу об оценке точности этого равенства мы вернемся позднее). Полагая в нем $\Delta y=f(x)-f(a)$, $\Delta x=x-a$, получаем

$$f(x)-f(a)\approx f'(a)(x-a),$$

или

$$f(x)\approx f(a)+f'(a)(x-a). \tag{3}$$

Выше мы отмечали, что уравнение касательной к графику функции $y=f(x)$ в точке $x=a$ имеет вид

$$y=f(a)+f'(a)(x-a). \tag{4}$$

Сопоставив правые части формул (3) и (4), мы еще раз убеждаемся в том, что дифференцируемость с геометрической точки зрения означает возможность приближенной замены графика (в достаточно малой окрестности рассматриваемой точки) прямой (касательной). Если же график функции имеет в некоторой точке излом, заострение, а тем более разрыв, то функция не является дифференцируемой в такой точке.

Итак, существуют непрерывные, но не дифференцируемые функции. С другой стороны, почти очевидна следующая теорема.

Теорема 2. *Если функция дифференцируема в точке, то она непрерывна в этой точке.*

□ Пусть функция дифференцируема в точке a; тогда $\Delta y=f'(a)\Delta x+\alpha\Delta x$. Если $x\to a$, т. е. $\Delta x\to 0$, то из написанного равенства следует, что $\Delta y\to 0$. Итак, Δy — бесконечно малая функция при $x\to a$, а это и означает, что функция $y=f(x)$ непрерывна в точке a (см. п. 1 § 17). ■

При использовании формулы (3) в приближенных вычислениях удобно несколько изменить обозначения: вместо a писать x_0, а вместо x писать a. Тогда получим

$$f(a)\approx f(x_0)+f'(x_0)(a-x_0). \tag{5}$$

На этой формуле основан следующий алгоритм.

АЛГОРИТМ

использования дифференциала в приближенных вычислениях

Чтобы найти приближенное значение некоторой величины A, нужно:

1°. Представить A в виде значения некоторой функции в точке $x = a$: $A = f(a)$.
2°. Подобрать x_0 так, чтобы точка x_0 была достаточно близка к точке a и чтобы значение $f(x_0)$ было вычислить легко.
3°. Вычислить $f(x_0)$.
4°. Для функции $f(x)$ найти $f'(x)$ и $f'(x_0)$.
5°. Подставить найденные значения a, x_0, $f(x_0)$ и $f'(x_0)$ в формулу (5).

Пример. Найти приближенное значение величины $2{,}008^3$.

○ 1°. Здесь $A = 2{,}008^3$; положим $f(x) = x^3$, тогда $A = f(a)$, где $a = 0{,}008$.
2°. Возьмем $x_0 = 2$.
3°. Имеем $f(x_0) = 2^3 = 8$.
4°. $f'(x) = 3x^2$; $f'(x_0) = f'(2) = 12$.
5°. По формуле (5) находим $f(a) \approx 8 + 12\cdot(2{,}008 - 2) = 8 + 12\cdot 0{,}008 = 8{,}096$, т. е. $2{,}008^3 \approx 8{,}096$ (заметим, что, вычислив $2{,}008^3$ с помощью микрокалькулятора, получим 8,0963845). ●

§ 30. Правила дифференцирования

1. Дифференцирование суммы, произведения, частного

В предыдущих параграфах мы использовали понятие производной при решении различных физических задач, для составления уравнения касательной, в приближенных вычислениях; с другими приложениями производной мы познакомимся позднее. Это означает, что необходимо уметь дифференцировать различные функции, которые могут встретиться в реальных задачах. В этом пункте речь идет о том, как, зная производные функций $u(x)$ и $v(x)$, находить производные суммы, произведения, частного. При доказательстве каждой теоремы будем использовать алгоритм нахождения производной (см. § 28) и соответствующую теорему об арифметических операциях над пределами, не оговаривая этого каждый раз особо (при этом мы рекомендуем приведенные ниже довольно краткие записи доказательств теорем дополнить необходимыми комментариями). Кроме того, учитывая, что $\Delta u = u(x+\Delta x) - u(x)$, $\Delta v = v(x+\Delta x) - v(x)$, будем пользоваться тем, что $u(x+\Delta x) = u(x) + \Delta u$, $v(x+\Delta x) = v(x) + \Delta v$.

Теорема 1. *Если функции $u(x)$ и $v(x)$ дифференцируемы в точке x, то их сумма $f(x) = u(x) + v(x)$ дифференцируема в точке x, причем $f'(x) = u'(x) + v'(x)$.*

□ 1°. $f(x) = u(x) + v(x)$.
2°. $f(x+\Delta x) = u(x+\Delta x) + v(x+\Delta x) = u(x) + \Delta u + v(x) + \Delta v$.
3°. $\Delta y = f(x+\Delta x) - f(x) = \Delta u + \Delta v$.
4°. $\dfrac{\Delta y}{\Delta x} = \dfrac{\Delta u + \Delta v}{\Delta x} = \dfrac{\Delta u}{\Delta x} + \dfrac{\Delta v}{\Delta x}$.
5°. $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x} = \lim\limits_{\Delta x \to 0} \dfrac{\Delta u}{\Delta x} + \lim\limits_{\Delta x \to 0} \dfrac{\Delta v}{\Delta x} = u'(x) + v'(x)$.

Итак, $f'(x) = u'(x) + v'(x)$. ■

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

Следствие. *Дифференциал суммы двух дифференцируемых в точке x функций равен сумме их дифференциалов:*

$$d\,(u(x) + v(x)) = du + dv.$$

□ Имеем $d\,(u(x) + v(x)) = (u(x) + v(x))'\,dx = (u'(x) + v'(x))\,dx = u'(x)\,dx + v'(x)dx = du + dv$. ■

Теорема 2. *Если функции $u(x)$ и $v(x)$ дифференцируемы в точке x, то и их произведение $f(x) = u(x)\,v(x)$ дифференцируемо в точке x, причем $f'(x) = u'(x)\,v(x) + u(x)\,v'(x)$.*

□ 1^0. $f(x) = u(x)\,v(x)$.

2^0. $f(x + \Delta x) = u(x + \Delta x)v(x + \Delta x) = (u(x) + \Delta u)(v(x) + \Delta v) = u(x)v(x) + \Delta u \cdot v(x) + \Delta v \cdot u(x) + \Delta u \Delta v$.

3^0. $\Delta y = f(x + \Delta x) - f(x) = \Delta u \cdot v(x) + \Delta v \cdot u(x) + \Delta u \Delta v$.

4^0. $\frac{\Delta y}{\Delta x} = \frac{\Delta u \cdot v(x) + \Delta v \cdot u(x) + \Delta u \Delta v}{\Delta x} = \frac{\Delta u}{\Delta x} v(x) + \frac{\Delta v}{\Delta x} u(x) + \frac{\Delta u \Delta v}{\Delta x}$.

5^0. $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \left(\lim\limits_{\Delta x \to 0} \frac{\Delta u}{\Delta x}\right) \cdot v(x) + u(x)\left(\lim\limits_{x \to 0} \frac{\Delta v}{\Delta x}\right) +$

$+ \lim\limits_{\Delta x \to 0} \frac{\Delta u}{\Delta x} \cdot \lim\limits_{x \to 0} \Delta v = u'(x)\,v(x) + u(x)\,v'(x) + u'(x) \cdot \lim\limits_{\Delta x \to 0} \Delta v$.

Согласно теореме 2 § 29 имеем $\lim\limits_{\Delta x \to 0} \Delta v = 0$; значит, в результате получаем

$$f'(x) = u'(x)\,v(x) + u(x)\,v'(x). \blacksquare$$

Следствие 1. *Справедлива формула* $(Cu(x))' = Cu'(x)$.

□ В самом деле, по теореме 2 имеем $(Cu(x))' = (C)'u(x) + C(u(x))'$. Но $(C)' = 0$; значит, $(Cu(x))' = Cu'(x)$. ■

Следствие 2. *Справедлива формула* $d\,(u(x)\,v(x)) = du \cdot v(x) + u(x)dv$.

□ $d\,(u(x)\,v(x)) = (u(x)\,v(x))'\,dx = (u'(x)\,v(x) + u(x)\,v'(x))\,dx = (u'(x)\,dx) \cdot v(x) + u(x)\,(v'(x)\,dx) = du \cdot v(x) + u(x)\,dv$. ■

Следствие 3. *Справедлива формула* $d\,(Cu(x)) = Cdu$.

Примеры. 1. $(5x^2)' = 5(x^2)' = 5 \cdot 2x = 10x$.

2. $(x^3)' = (x^2 \cdot x) = (x^2)'x + x^2(x)' = 2x \cdot x + x^2 \cdot 1 = 3x^2$.

3. $\left(\frac{x^3}{3} - 2x^2 + 4x - 3\right)' = \left(\frac{1}{3}x^3\right)' + (-2x^2)' +$

$+ (4x - 3)' = \frac{1}{3} \cdot 3x^2 - 2 \cdot 2x + 4 = x^2 - 4x + 4$.

Теорема 3. *Если функции $u(x)$ и $v(x)$ дифференцируемы в точке x и, кроме того, $v(x) \neq 0$, то их частное $f(x) = \frac{u(x)}{v(x)}$ дифференцируемо в точке x, причем*

$$f'(x) = \frac{u'(x)\,v(x) - u(x)\,v'(x)}{v^2(x)}.$$

□ 1^0. $f(x) = \frac{u(x)}{v(x)}$.

2^0. $f(x + \Delta x) = \frac{u(x + \Delta x)}{v(x + \Delta x)} = \frac{u(x) + \Delta u}{v(x) + \Delta v}$.

3^0. $\Delta y = f(x+\Delta x) - f(x) = \frac{u(x)+\Delta u}{v(x)+\Delta v} - \frac{u(x)}{v(x)} = \frac{\Delta u\cdot v(x) - \Delta v\cdot u(x)}{(v(x)+\Delta v)\,v(x)}$.

4^0. $\frac{\Delta y}{\Delta x} = \frac{\Delta u\cdot v(x) - \Delta v u(x)}{(v(x)+\Delta v)v(x)\Delta x} = \left(\frac{\Delta u}{\Delta x}v(x) - u(x)\frac{\Delta v}{\Delta x}\right)\frac{1}{v^2(x)+v(x)\Delta v}$.

5^0. Переходя к пределу при $\Delta x\to 0$, воспользуемся, как и в доказательстве теоремы 2, тем, что $\lim\limits_{\Delta x\to 0}\Delta v = 0$. Тогда получим

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \left(\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\cdot v(x) - u(x)\lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x}\right)\frac{1}{v^2(x)} = \frac{u'(x)v(x) - u(x)v'(x)}{v^2(x)}. \blacksquare$$

Следствие. *Справедлива формула*

$$d\left(\frac{u(x)}{v(x)}\right) = \frac{du\cdot v(x) - u(x)\cdot dv}{v^2(x)}.$$

Например,

$$\left(\frac{2x+3}{3x-4}\right)' = \frac{(2x+3)'(3x-4)-(2x+3)(3x-4)'}{(3x-4)^2} = \frac{2\cdot(3x-4)-(2x+3)\cdot 3}{(3x-4)^2} = -\frac{17}{(3x-4)^2}.$$

Таким образом, в этом пункте мы вывели следующие правила дифференцирования:

I. *Производная суммы двух и вообще любого конечного числа функций равна сумме производных.*

II. *Постоянный множитель можно вынести за знак производной.*

III. *Производная произведения двух функций $u(x)$ и $v(x)$ равна $u'v+uv'$.*

IV. *Производная частного двух функций $u(x)$ и $v(x)$ равна $\frac{u'v-uv'}{v^2}$ (если $v(x)\neq 0$).*

2. Дифференцирование сложной функции

Теорема 4. *Пусть дана сложная функция $y=f(u)$, $u=g(x)$ и пусть функция $u=g(x)$ дифференцируема в точке x, а функция $y=f(u)$ дифференцируема в соответствующей точке $u=g(x)$. Тогда сложная функция $y=f(g(x))$ дифференцируема в точке x, причем*

$$y' = f'(u)g'(x). \quad (1)$$

□ Так как функция $u=g(x)$ дифференцируема в точке x, то приращение функции при переходе от точки x к точке $x+\Delta x$ можно записать в виде

$$\Delta u = g'(x)\Delta x + \alpha\Delta x, \quad (2)$$

где $\alpha\to 0$ при $\Delta x\to 0$.

Аналогично, так как функция $y=f(u)$ дифференцируема в точке u, то

$$\Delta y = f'(u)\Delta u + \beta\Delta u. \quad (3)$$

где $\beta\to 0$ при $\Delta u\to 0$.

Подставив в равенство (3) вместо Δu его значение, определяемое равенством (2), получим

$$\Delta y = f'(u)(g'(x)\Delta x + \alpha\Delta x) + \beta(g'(x)\Delta x + \alpha\Delta x) =$$
$$= f'(u)g'(x)\Delta x + (f'(u)\alpha + g'(x)\beta + \alpha\beta)\Delta x.$$

Если $\Delta x \to 0$, то из равенства (2) следует, что $\alpha \to 0$ и $\Delta u \to 0$; если же $\Delta u \to 0$, то из равенства (3) следует, что $\beta \to 0$. В итоге получаем, что $f'(u)\alpha + g'(x)\beta + \alpha\beta$ есть бесконечно малая при $\Delta x \to 0$; обозначим ее через γ.

Итак, $\Delta y = f'(u)g'(x)\Delta x + \gamma\Delta x$. Отсюда имеем $\frac{\Delta y}{\Delta x} = f'(u)g'(x) + \gamma$, а $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(u)g'(x)$. Это означает, что $y' = f'(u)g'(x)$. ■

Пример. Найти производную функции $y = \left(x^2 + \frac{1}{x}\right)^3$.

○ Здесь $y = u^3$, $u = x^2 + \frac{1}{x}$. Значит,

$$y' = (u^3)'\left(x^2 + \frac{1}{x}\right)' = 3u^2\left(2x - \frac{1}{x^2}\right) = 3\left(x^2 + \frac{1}{x}\right)^2\left(2x - \frac{1}{x^2}\right). \ \bullet$$

На практике равенство (1) записывают в виде

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} \text{ или } y'_x = y'_u u'_x$$

и формулируют в виде следующего правила:

V. *Производная композиции функций по независимой переменной равна произведению производной по промежуточному аргументу на производную промежуточного аргумента по независимой переменной.*

Этому правилу можно дать наглядное истолкование: y'_u есть скорость изменения y относительно u, а u'_x — скорость изменения u относительно x; иными словами, в данной точке y изменяется в y'_u раз быстрее, чем u, а u — в u'_x раз быстрее, чем x. Ясно, что тогда y изменяется в $y'_u u'_x$ раз быстрее, чем x, т. е. $y'_x = y'_u u'_x$.

Правило V распространяется на случай композиции трех и большего числа дифференцируемых функций. Например, если $y = f(t)$, $t = g(u)$, $u = h(x)$, то $t'_x = t'_u u'_x$; тогда

$$y'_x = y'_t t'_x = y'_t t'_u u'_x \text{ или } \frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{du}\frac{du}{dx}.$$

3. Инвариантность формы записи дифференциала

По определению имеем $dy = y'_x\Delta x$ или, поскольку мы условились вместо приращения Δx независимой переменной писать dx, получим $dy = y'_x dx$; здесь $y = f(x)$ — дифференцируемая функция в точке x.

Пусть теперь x не является независимой переменной, а, в свою очередь, представляет собой дифференцируемую функцию независимой переменной t: $x = \varphi(t)$. Тогда y есть сложная функция независимой переменной t и, следовательно, $dy = y'_t\Delta t$ или, в силу указанного соглашения, $dy = y'_t dt$. Так как $y'_t = y'_x x'_t$, то $dy = y'_x x'_t dt$. Далее, $x'_t dt = dx$ и, значит, $dy = y'_x dx$.

Таким образом, форма дифференциала сохранилась; *дифференциал функции имеет один и тот же вид как в том случае, когда x — независимая переменная, так и в том, когда x — зависимая промежуточная переменная*: дифференциал равен произведению производной на дифференциал того аргумента, по которому взята производная. Это свойство называют *инвариантностью* (т. е. неизменностью) формы дифференциала. Подчеркнем, что речь идет только о сохранении формы: если x независимая переменная, то $dx = \Delta x$; если же x — зависимая переменная, то, как правило, $dx \neq \Delta x$.

§ 31. Формулы дифференцирования

1. Дифференцирование степенной функции

Выведем формулу дифференцирования функции $y=x^\alpha$, где α — любое действительное число, $x>0$ (это ограничение можно снять в случае целого α). Здесь и далее будем использовать алгоритм нахождения производной (см. § 28). Имеем

1^0. $f(x)=x^\alpha$.

2^0. $f(x+\Delta x)=(x+\Delta x)^\alpha$.

3^0. $\Delta y=(x+\Delta x)^\alpha-x^\alpha=x^\alpha\left(\left(1+\frac{\Delta x}{x}\right)^\alpha-1\right)$.

4^0. $\frac{\Delta y}{\Delta x}=x^\alpha\cdot\frac{\left(1+\frac{\Delta x}{x}\right)^\alpha-1}{\Delta x}$.

5^0. Переходя к пределу при $\Delta x\to 0$, заметим, что $\frac{\Delta x}{x}$ — бесконечно малая при $\Delta x\to 0$, а значит, $\left(1+\frac{\Delta x}{x}\right)^\alpha-1$ и $\alpha\cdot\frac{\Delta x}{x}$ — эквивалентные бесконечно малые (см. § 27). Следовательно,

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=x^\alpha\cdot\lim_{\Delta x\to 0}\frac{\left(1+\frac{\Delta x}{x}\right)^\alpha-1}{\Delta x}=x^\alpha\cdot\lim_{\Delta x\to 0}\frac{\alpha\cdot\frac{\Delta x}{x}}{\Delta x}=x^\alpha\cdot\frac{\alpha}{x}=\alpha x^{\alpha-1}.$$

Итак,

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

Например, $(x^{100})'=100x^{99}$, $\left(\frac{1}{x^{11}}\right)'=(x^{-11})'=-11x^{-12}$; $(\sqrt[5]{x^3})'=(x^{3/5})'=$ $=\frac{3}{5}x^{\frac{3}{5}-1}=\frac{3}{5}x^{-\frac{2}{5}}$.

Заметим, что выведенные ранее формулы $(x)'=1$, $(x^2)'=2x$, $(x^3)'=3x^2$ являются частными случаями формулы (1).

2. Дифференцирование показательной функции

Применим алгоритм нахождения производной к функции $y=a^x$, где $a>0$, $a\neq 1$.

1^0. $f(x)=a^x$.

2^0. $f(x+\Delta x)=a^{x+\Delta x}$.

3^0. $\Delta y=a^{x+\Delta x}-a^x=a^x(a^{\Delta x}-1)$.

4^0. $\frac{\Delta y}{\Delta x}=a^x\cdot\frac{a^{\Delta x}-1}{\Delta x}$.

5^0. $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=a^x\cdot\lim\limits_{\Delta x\to 0}\frac{a^{\Delta x}-1}{\Delta x}=a^x\cdot\ln a$ (см. § 27).

Итак,

$$(a^x)'=a^x\cdot\ln a. \tag{2}$$

В частности,

$$(e^x)'=e^x. \tag{3}$$

Примеры. 1. Найти производную функции $y=5^x+x^5$.

○ Имеем

$$y'=(5^x+x^5)'=(5^x)'+(x^5)'=5^x\ln 5+5x^4. \bullet$$

2. Найти производную функции $y=e^{x^2-5x+6}$.

○ Здесь $y=e^u$, $u=x^2-5x+6$. Применив правило дифференцирования сложной функции, получим

$$y'=(e^u)'(x^2-5x+6)'=e^u(2x-5)=e^{x^2-5x+6}(2x-5). \bullet$$

3. Составить уравнение касательной к графику функции $y=e^x$ в точке с абсциссой $x=0$.

○ Здесь $a=0$ — абсцисса точки касания; $f(a)=f(0)=e^0=1$; $f'(x)=(e^x)'=e^x$; $f'(a)=e^0=1$. Подставив в уравнение касательной $y=f(a)+f'(a)\cdot(x-a)$ вместо a, $f(a)$, $f'(a)$ их значения, получим

$$y=1+1\cdot(x-0), \text{ т. е. } y=x+1. \bullet$$

Следовательно, касательная к графику функции $y=e^x$ в точке $(0;\ 1)$ составляет с осью x угол 45° (см. рис. 71); этим функция e^x выделяется среди показательных функций с другим основанием.

3. Дифференцирование логарифмической функции

Применим алгоритм нахождения производной к функции $y=\ln x$ в точке $x>0$.

1⁰. $f(x)=\ln x$.

2⁰. $f(x+\Delta x)=\ln(x+\Delta x)$.

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

4⁰. $\frac{\Delta y}{\Delta x}=\frac{\ln\left(1+\frac{\Delta x}{x}\right)}{\Delta x}$.

5⁰. Переходя к пределу при $\Delta x\to 0$, заметим, что $\frac{\Delta x}{x}$ — бесконечно малая при $\Delta x\to 0$, а значит, $\ln\left(1+\frac{\Delta x}{x}\right)$ и $\frac{\Delta x}{x}$ — эквивалентные бесконечно малые (см. § 27). Следовательно,

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{\ln\left(1+\frac{\Delta x}{x}\right)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\frac{\Delta x}{x}}{\Delta x}=\frac{1}{x}.$$

Итак,

$$(\ln x)'=\frac{1}{x}. \qquad (4)$$

Согласно формуле перехода к новому основанию логарифма, для функции $y=\log_a x$ имеем $\log_a x=\frac{\ln x}{\ln a}$. Значит,

$$(\log_a x)'=\left(\frac{\ln x}{\ln a}\right)'=\frac{1}{\ln a}\cdot(\ln x)'=\frac{1}{\ln a}\cdot\frac{1}{x}.$$

Итак,

$$(\log_a x)'=\frac{1}{x\ln a}. \qquad (5)$$

Пример. Найти приближенное значение величины $\ln 1{,}1$.

○ Применим алгоритм использования дифференциала в приближенных вычислениях (см. § 29).

1^0. $A=\ln 1{,}1$; $f(x)=\ln x$, $a=1{,}1$; $A=f(a)$.

2^0. $x_0=1$.

3^0. $f(x_0)=\ln 1=0$.

4^0. $f'(x)=(\ln x)'=\frac{1}{x}$; $f'(x_0)=1$.

5^0. Так как $f(a)\approx f(x_0)+f'(x_0)(a-x_0)$, то $\ln 1{,}1\approx 0+1\cdot(1{,}1-1)=0{,}1$. ●

4. Дифференцирование тригонометрических функций

Применим алгоритм нахождения производной к функции $y=\sin x$.

1^0. $f(x)=\sin x$.

2^0. $f(x+\Delta x)=\sin(x+\Delta x)$.

3^0. $\Delta y=\sin(x+\Delta x)-\sin x$. Используя формулу разности синусов

$$\sin\alpha-\sin\beta=2\sin\frac{\alpha-\beta}{2}\cos\frac{\alpha+\beta}{2},$$

получим

$$\Delta y=2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right).$$

4^0. $\frac{\Delta y}{\Delta x}=\frac{2\sin\frac{\Delta x}{2}}{\Delta x}\cos\left(x+\frac{\Delta x}{2}\right)=\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\cdot\cos\left(x+\frac{\Delta x}{2}\right)$.

5^0. Переходя к пределу при $\Delta x\to 0$, заметим, что $\lim\limits_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}=1$ (первый замечательный предел), а $\lim\limits_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right)=\cos x$ в силу непрерывности функции $\cos x$ в любой точке x. Значит,

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\cdot\lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right)=1\cdot\cos x=\cos x.$$

Итак,

$$(\sin x)'=\cos x. \qquad (6)$$

Аналогично доказывается, что

$$(\cos x)'=-\sin x. \qquad (7)$$

Для функции $y=\operatorname{tg}x$, воспользовавшись правилом дифференцирования частного, получим

$$y'=(\operatorname{tg}x)'=\left(\frac{\sin x}{\cos x}\right)'=\frac{(\sin x)'\cos x-\sin x(\cos x)'}{\cos^2 x}=$$

$$=\frac{\cos x\cdot\cos x-\sin x(-\sin x)}{\cos^2 x}=\frac{\cos^2 x+\sin^2 x}{\cos^2 x}=\frac{1}{\cos^2 x}.$$

Итак,

$$(\operatorname{tg}x)'=\frac{1}{\cos^2 x}. \qquad (8)$$

Аналогично доказывается, что

$$(\mathrm{ctg}\,x)' = -\frac{1}{\sin^2 x}. \qquad (9)$$

Пример. Найти производную функций $y = \sqrt{\sin(\ln x)}$.

○ Здесь $y = \sqrt{t}$, $t = \sin u$, $u = \ln x$. Применив правило дифференцирования сложной функции, получим

$$y' = (\sqrt{t})'(\sin u)'(\ln x)' = \frac{1}{2}t^{-1/2}\cos u\frac{1}{x} =$$

$$= \frac{1}{2\sqrt{\sin(\ln x)}}\cos(\ln x)\frac{1}{x}. \bullet$$

5. Дифференцирование обратных тригонометрических функций

Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и возрастает (убывает) на этом отрезке, то для нее существует обратная функция φ, которая непрерывна и возрастает (убывает) на отрезке $[f(a), f(b)]$ (соответственно на отрезке $[f(b), f(a)]$; см. § 18). При этом $y = f(x)$ тогда и только тогда, когда $x = \varphi(y)$. Докажем, что если, кроме того, функция $f(x)$ дифференцируема в какой-нибудь внутренней точке x из $[a, b]$, причем в этой точке $f'(x) \neq 0$, то функция φ дифференцируема в точке $y = f(x)$.

Теорема 1. *Пусть функция $f(x)$ непрерывна и монотонна на отрезке $[a, b]$ и дифференцируема во внутренней точке x этого отрезка, причем ее производная в этой точке отлична от нуля. Тогда функция $\varphi(y)$, обратная к $f(x)$, дифференцируема в точке $y = f(x)$, причем*

$$\varphi'(y) = \frac{1}{f'(x)}. \qquad (10)$$

Эту формулу записывают также в виде $x'_y = 1/y'_x$.

□Дадим числу $y = f(x)$ приращение $\Delta y \neq 0$. Обратная к f функция φ получит приращение $\Delta\varphi$, равное $\varphi(y + \Delta y) - \varphi(y)$. В силу монотонности обратной функции это приращение отлично от нуля: $\Delta\varphi \neq 0$. Поскольку $y = f(x)$ и $x = \varphi(y)$, вместо Δy и $\Delta\varphi$ можно писать Δf и Δx: $\Delta y = \Delta f$, $\Delta x = \Delta\varphi$. Поэтому

$$\frac{\Delta\varphi}{\Delta y} = \frac{\Delta x}{\Delta f} = \frac{1}{\frac{\Delta f}{\Delta x}}.$$

Перейдем в полученном равенстве к пределу при $\Delta y \to 0$. Так как функция φ непрерывна (по теореме о непрерывности обратной функции), то $\lim\limits_{\Delta y \to 0}\Delta\varphi = 0$, т. е. $\lim\limits_{\Delta y \to 0}\Delta x = 0$, а потому имеем

$$\varphi'(y) = \lim_{\Delta y \to 0}\frac{\Delta\varphi}{\Delta y} = \lim_{\Delta x \to 0}\frac{1}{\frac{\Delta f}{\Delta x}} = \frac{1}{\lim\limits_{\Delta x \to 0}\frac{\Delta f}{\Delta x}} = \frac{1}{f'(x)}.$$

Этим доказано, что $\varphi'(y) = \frac{1}{f'(x)}$. ■

Формула (10) имеет простой геометрический смысл. Графики функций $y = f(x)$ и $y = \varphi(x)$ симметричны относительно прямой $y = x$. При симметрии относительно этой прямой касательная к графику функции f в точке M $(x_0; y_0)$ переходит в касательную к графику функции φ в точке N $(y_0; x_0)$ (рис. 109). Так как углы наклона этих касательных к оси абсцисс в сумме составляют

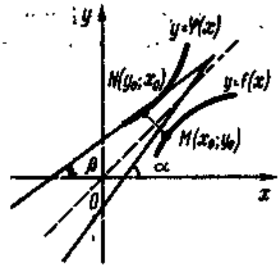


Рис. 109

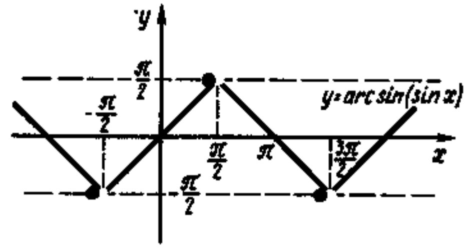


Рис. 110

$\pi/2$, то $\operatorname{tg}\alpha = \operatorname{tg}\left(\frac{\pi}{2} - \beta\right) = \operatorname{ctg}\beta = \frac{1}{\operatorname{tg}\beta}$. Но $\operatorname{tg}\alpha = f'(x)$, $\operatorname{tg}\beta = \varphi'(y)$, и, следовательно, $f'(x) = 1/\varphi'(y)$.

Применим теорему о дифференцировании обратной функции к нахождению производных обратных тригонометрических функций. Начнем с дифференцирования функции $\arcsin x$. Если $y = \arcsin x$, где $-1 \leqslant x \leqslant 1$, то $x = \sin y$, где $-\pi/2 \leqslant y \leqslant \pi/2$.

Функция $\sin y$ непрерывна и монотонна на отрезке $[-\pi/2, \pi/2]$, причем ее производная по y равна $\cos y$, а потому отлична от нуля в любой внутренней точке этого отрезка. Значит, выполнены все условия теоремы 1. Согласно этой теореме, получаем

$$(\arcsin x)' = \frac{1}{\cos y} = \frac{1}{\cos(\arcsin x)}.$$

Но

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

Итак, мы доказали, что

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}. \qquad (11)$$

Аналогично выводится формула

$$(\arccos x)' = -\frac{1}{\sqrt{1 - x^2}}. \qquad (12)$$

Обе формулы справедливы на интервале $(-1, 1)$.

Найдем производную функции $\operatorname{arctg} x$. Если $y = \operatorname{arctg} x$, $-\infty < x < +\infty$, то $x = \operatorname{tg} y$, $-\pi/2 < y < \pi/2$.

Здесь также выполнены все условия теоремы 1, причем производная от $\operatorname{tg} y$ по y равна $1/\cos^2 y$. Значит,

$$(\operatorname{arctg} x)' = \frac{1}{1/\cos^2 y} = \frac{1}{1 + \operatorname{tg}^2 y} = \frac{1}{1 + \operatorname{tg}^2(\operatorname{arctg} x)} = \frac{1}{1 + x^2}.$$

Таким образом,

$$(\operatorname{arctg} x)' = \frac{1}{1 + x^2}. \qquad (13)$$

Аналогично выводится формула

$$(\operatorname{arcctg} x)' = -\frac{1}{1 + x^2}. \qquad (14)$$

Примеры. 1. Найти дифференциал функции $\operatorname{arctg}^3\sqrt{x}$ в точке $x = 1$ при $dx = 0{,}08$.

○ Имеем $t = \sqrt{x}$, $u = \operatorname{arctg} t$, $y = u^3$. Воспользовавшись правилом дифференцирования сложной функции, получим

$$y'_x = y'_u u'_t t'_x = 3u^2 \frac{1}{1+t^2} \frac{1}{2\sqrt{x}} =$$

$$= 3\operatorname{arctg}^2\sqrt{x} \frac{1}{1+(\sqrt{x})^2} \frac{1}{2\sqrt{x}} = \frac{3\operatorname{arctg}^2\sqrt{x}}{2\sqrt{x}(1+x)}.$$

Значит,

$$y'\Big|_{x=1} = \frac{3\operatorname{arctg}^2 1}{2\cdot 1(1+1)} = \frac{3}{4}\left(\frac{\pi}{4}\right)^2 = \frac{3\pi^2}{64}.$$

Тогда

$$dy = y'dx = \frac{3\pi^2}{64}\, 0{,}08 = \frac{3\pi^2}{800}. \; ●$$

2. Найти производную функции $\arcsin(\sin x)$.

○ Имеем

$$f'(x) = \frac{1}{\sqrt{1-\sin^2 x}} \cos x = \frac{\cos x}{\sqrt{\cos^2 x}} = \frac{\cos x}{|\cos x|}.$$

Замечаем, что если $\cos x > 0$, то $f'(x) = 1$; если $\cos x < 0$, то $f'(x) = -1$; если $\cos x = 0$, то $f'(x)$ не существует. Полученный результат хорошо иллюстрирует график функции $y = \arcsin(\sin x)$, изображенный на рис. 110. В точках $x = \pi/2 + \pi n$ график имеет излом. ●

6. Дифференцирование гиперболических функций

Найдем производную функции $\operatorname{sh} x$ (см. п. 2 § 26). Имеем

$$(\operatorname{sh} x)' = \left(\frac{e^x - e^{-x}}{2}\right)' = \frac{e^x + e^{-x}}{2} = \operatorname{ch} x.$$

Аналогично находим

$$(\operatorname{ch} x)' = \left(\frac{e^x + e^{-x}}{2}\right) = \frac{e^x - e^{-x}}{2} = \operatorname{sh} x.$$

Итак, $(\operatorname{sh} x)' = \operatorname{ch} x$, $(\operatorname{ch} x)' = \operatorname{sh} x$. Это еще одно подтверждение отмеченной выше аналогии между гиперболическими и тригонометрическими функциями.

Воспользовавшись правилом дифференцирования частного, получим еще две формулы дифференцирования:

$$(\operatorname{th} x)' = \left(\frac{\operatorname{sh} x}{\operatorname{ch} x}\right)' = \frac{(\operatorname{sh} x)'\operatorname{ch} x - \operatorname{sh} x(\operatorname{ch} x)'}{\operatorname{ch}^2 x} = \frac{\operatorname{ch}^2 x - \operatorname{sh}^2 x}{\operatorname{ch}^2 x} = \frac{1}{\operatorname{ch}^2 x};$$

$$(\operatorname{cth} x)' = -\frac{1}{\operatorname{sh}^2 x}.$$

7. Дифференцирование функций, заданных параметрически

Плоские кривые часто задаются уравнениями вида $x = x(t)$, $y = y(t)$, где переменная t, называемая *параметром*, пробегает некоторый промежуток значений T. Чтобы построить кривую, заданную параметрически, нужно задать ряд значений параметра t по формулам $x = x(t)$, $y = y(t)$, вычислить соответствующие значения x и y, отметить полученные точки $(x; y)$ на координатной плоскости и соединить их плавной линией. Из курса геометрии известны параметрические уравнения окружности $x^2+y^2=a^2$ и эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, которые соответственно имеют вид

$$\begin{cases} x=a\cos t, \\ y=a\sin t, \text{ где } t\in[0,\ 2\pi]; \end{cases} \quad \begin{cases} x=a\cos t, \\ y=b\sin t, \text{ где } t\in[0,\ 2\pi]. \end{cases}$$

В качестве нового примера рассмотрим кривую, заданную уравнениями

$$\begin{cases} x=a\cos^3 t, \\ y=a\sin^3 t, \text{ где } t\in[0,\ 2\pi]. \end{cases}$$

Составим таблицу значений:

t	0	$\pi/2$	π	$3\pi/2$	2π	$\pi/4$	$3\pi/4$	$5\pi/4$	$7\pi/4$
x	a	0	$-a$	0	a	$a\sqrt{2}/4$	$-a\sqrt{2}/4$	$-a\sqrt{2}/4$	$a\sqrt{2}/4$
y	0	a	0	$-a$	0	$a\sqrt{2}/4$	$a\sqrt{2}/4$	$-a\sqrt{2}/4$	$-a\sqrt{2}/4$

Отметим точки $(a;\ 0)$, $(0;\ a)$, $(-a;\ 0)$, $(0;\ -a)$, $(a\sqrt{2}/4;\ a\sqrt{2}/4)$, $(-a\sqrt{2}/4;\ a\sqrt{2}/4)$, $(-a\sqrt{2}/4;\ -a\sqrt{2}/4)$, $(a\sqrt{2}/4;\ -a\sqrt{2}/4)$ на координатной плоскости и соединим их плавной линией (рис. 111). Построенная кривая называется *астроидой*.

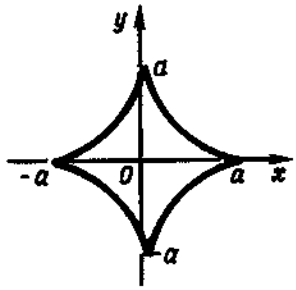


Рис. 111

Если функция $x=x(t)$ обратима, т. е. существует обратная функция $t=t(x)$, то уравнение $y=y(t)$ можно переписать в виде $y=y(t(x))$ и тогда ставить вопрос об отыскании производной функции $y=y(t(x))$ или, как говорят, функции, заданной параметрически уравнениями $x=x(t)$, $y=y(t)$.

Теорема 2. *Пусть функции $x=x(t)$, $y=y(t)$ непрерывны на $[\alpha,\ \beta]$, дифференцируемы в $(\alpha,\ \beta)$, причем $x'(t)$ сохраняет постоянный знак на этом интервале. Пусть, далее $[a,\ b]$ — область значений функции $x=x(t)$. Тогда уравнения $x=x(t)$, $y=y(t)$ определяют непрерывную на $[a,\ b]$ и дифференцируемую в $(a,\ b)$ функцию $y=y(x)$, причем $y'_x=y'(t)/x'(t)$.*

□ По условию, $x'(t)$ сохраняет постоянный знак; пусть для определенности $x'(t)>0$. Тогда функция $x(t)$ монотонна и непрерывна на $[\alpha,\ \beta]$ (см. теорему 3 из § 34); значит, она обратима и, согласно теореме 1, производная обратной функции $t(x)$ вычисляется по формуле $t'(x)=1/x'(t)$.

Имеем $y = y(t)$, $t = t(x)$. Воспользовавшись правилом дифференцирования сложной функции, получим

$$y'_x = y'_t t'_x = y'_t \frac{1}{x'_t} = \frac{y'(t)}{x'(t)}.$$

Итак, $y'_x = \frac{y'(t)}{x'(t)}$. ■

Например, для астроиды имеем

$$y' = \frac{(a\sin^3 t)'}{(a\cos^3 t)'} = \frac{3a\sin^2 t\cos t}{3a\cos^2 t(-\sin t)} = -\operatorname{tg} t.$$

Значит, в частности, $y' = 0$ при $t = 0$, $t = \pi$; y' не существует при $t = \pi/2$, $t = 3\pi/2$. Это наглядно иллюстрирует рис. 111: при значениях $t = 0$, $t = \pi$ астроида касается оси x (т. е. $y' = 0$), а при значениях $t = \pi/2$, $t = 3\pi/2$ она касается оси y (т. е. y' не существует).

§ 32. Производные и дифференциалы высших порядков

1. Понятие производной n-го порядка

Пусть функция $f(x)$ дифференцируема на множестве X. Производная $f'(x)$ этой функции является функцией от x на X. Следовательно, можно говорить о производной полученной функции, т. е. о производной от первой производной. Если она существует, то ее называют *производной второго порядка* функции $f(x)$ или, короче, *второй производной* и обозначают $f''(x)$. Значит, по определению $f''(x)=(f'(x))'$.

Аналогично, если существует производная от второй производной, то ее называют *третьей производной* и обозначают $f'''(x)$. Следовательно, по определению $f'''(x)=(f''(x))'$.

Вообще, *производной n-го порядка* называют производную от производной $(n-1)$-го порядка. Производную n-го порядка обозначают $f^{(n)}(x)$. Итак, по определению $f^{(n)}(x)=(f^{(n-1)}(x))'$.

Для производной n-го порядка используется также обозначение $y^{(n)}$.

Выше мы установили, что первая производная координаты x по времени t есть мгновенная скорость: $v=x'$.

Рассматривая производную скорости по времени t, получим скорость изменения скорости, т. е. ускорение: $a=v'=(x')'=x''$.

Итак, *вторая производная координаты по времени равна ускорению*. В этом состоит механический смысл второй производной.

Пример 1. Найти $f'''(1)$, если $f(x) = x^5$.

○ Имеем $f'(x) = (5x^4)'$, $f''(x) = (5x^4)' = 20x^3$, $f'''(x) = (20x^3)' = 60x^2$. Следовательно, $f'''(1) = 60\cdot 1^2 = 60$. ●

2. Найти $f^{(n)}$, если $f(x) = x^\alpha$, $x > 0$.

○ Имеем $f'(x)=\alpha x^{\alpha-1}$, $f''(x)=\alpha(\alpha-1)x^{\alpha-2}$, $f'''(x)=\alpha(\alpha-1)(\alpha-2)x^{\alpha-3}$. Можно сделать индуктивное предположение, что $(x^\alpha)^{(n)}=\alpha(\alpha-1)(\alpha-2)...(\alpha-n+1)x^{\alpha-n}$. Эту формулу нетрудно доказать методом математической индукции. ●

Отметим, что если α — натуральное число ($\alpha=m$), то $(x^m)^{(m)}=m!$. При $n>m$ получим $(x^m)^{(n)}=0$.

3. Тело движется прямолинейно по закону $x = 10t^2+3t-1$ (x — в метрах, t — в секундах). Доказать, что это движение происходит под действием постоянной силы.

○ Находим $x'=(10t^2+3t-1)'=20t+3$, $x''=(20t+3)'=20$. Значит, ускорение постоянно и равно 20 м/с². Так как по закону Ньютона действующая сила пропорциональна ускорению, то она постоянна. ●

4. Гармоническое колебание описывается законом $x=A\cos(\omega t+\alpha)$, где t — время, x — координата колеблющейся точки. Доказать что закон гармонических колебаний удовлетворяет дифференциальному уравнению $x''+\omega^2x=0$.

○ Находим $x'=-\omega\sin(\omega t+\alpha)$, $x''=-A\omega^2\cos(\omega t+\alpha)$. Тогда $x''+\omega^2x=-A\omega^2\cos(\omega t+\alpha)+\omega^2A\cos(\omega t+\alpha)=0$, что и требовалось доказать. ●

2. Формула Тейлора для многочлена

Пусть дан многочлен

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n.$$

При $x=0$ получим $a_0=f(0)$.

Покажем, что остальные коэффициенты этого многочлена выражаются через производные от f в точке $x=0$. Имеем

$$f'(x)=a_1+2a_2x+3a_3x^2+4a_4x^3+\ldots+na_nx^{n-1}. \tag{1}$$

Полагая в этом равенстве $x=0$, находим $a_1=f'(0)$.

Продифференцируем обе части равенства (1):

$$f''(x)=2a_2+3\cdot 2a_3x+4\cdot 3a_4x^2+\ldots+n(n-1)a_nx^{n-2}.$$

Полагая в этом равенстве $x=0$, получим $a_2=\frac{f''(0)}{2}=\frac{f''(0)}{2!}$.

Точно так же находим

$$a_3=\frac{f'''(0)}{3!},\ a_4=\frac{f^{\text{IV}}(0)}{4!},\ldots,a_n=\frac{f^{(n)}(0)}{n!}.$$

Итак, если $f(x)$ — многочлен n-й степени, то

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\frac{f'''(0)}{3!}+\ldots+\frac{f^{(n)}(0)}{n!}.$$

Аналогично выводится более общее равенство

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n,$$

которое называется *формулой Тейлора** для многочлена $f(x)$.

3. Свойства производной n-го порядка

1⁰. *Производная n-го порядка от суммы конечного числа n раз дифференцируемых функций равна сумме производных n-го порядка, взятых от каждой слагаемой функции.*

□ В самом деле, последовательно дифференцируя функцию $f=g+h$, получим:

$$f'=g'+h';\ f''=(f')'=(g'+h')'=(g')'+(h')'=g''+h''.$$
$$f'''=(f'')'=(g''+h'')'=(g'')'+(h'')'=g'''+h''' \text{ и т. д.} ■$$

2⁰. *Постоянный множитель можно вынести за знак производной n-го порядка, т. е.* $(Cg)^{(n)}=Cg^{(n)}$.

Доказательство рекомендуем провести самостоятельно.

* Б. Тейлор (1685—1731) — английский математик.

3^0. *Если f и g — две функции, имеющие производные до n-го порядка включительно, то*

$$(gh)^{(n)}=gh^{(n)}+C_n^1g'h^{(n-1)}+C_n^2g''h^{(n-2)}+\ldots+C_n^kg^{(k)}h^{(n-k)}+\ldots+C_n^{n-1}g^{(n-1)}h'+g^{(n)}h, \tag{2}$$

где $C_n^k=\frac{n!}{k!(n-k)!}$.

Эта формула называется *формулой Лейбница**. Она доказывается методом математической индукции.

Для запоминания формулы Лейбница полезно обратить внимание на ее аналогию с *формулой бинома Ньютона***:

$$(a+b)^n=a^n+C_n^1a^{n-1}b+C_n^2a^{n-2}b^2+\ldots+C_n^ka^{n-k}b^k+\ldots+C_n^{n-1}ab^{n-1}+b^n.$$

Пример. Найти $f^{(30)}$ для функции x^3e^x.

○ Согласно формуле Лейбница, получим

$$(x^3e^x)^{(30)}=x^3(e^x)^{(30)}+C_{30}^1(x^3)'(e^x)^{(29)}+C_{30}^2(x^3)''(e^x)^{(28)}+$$
$$+C_{30}^3(x^3)'''(e^x)^{(27)}+C_{30}^4(x^3)^{(4)}(e^x)^{(26)}+\ldots+(x^3)^{(30)}e^x.$$

Заметим, что $(x^3)'=3x^2$. $(x^3)''=6x$, $(x^3)'''=6$, $(x^3)^{(4)}=0$, а все последующие производные функции x^3 равны нулю. Таким образом,

$$f^{(30)}(x)=x^3(e^x)^{(30)}+C_{30}^1\cdot3x^2(e^x)^{(29)}+C_{30}^2\cdot6x(e^x)^{(28)}+C_{30}^3\cdot6(e^x)^{(27)}.$$

Далее, имеем $(e^x)'=(e^x)''=\ldots=(e^x)^{(30)}=e^x$. Значит,

$$f^{(30)}(x)=e^x(x^3+3C_{30}^1x^2+6C_{30}^2x+6C_{30}^3).$$

Остается вычислить коэффициенты. Находим

$$C_{30}^1=30,\ C_{30}^2=\frac{30\cdot29}{2!}=435,\ C_{30}^3=\frac{30\cdot29\cdot28}{3!}=4060.$$

Окончательно получим

$$f^{(30)}(x)=(x^3e^x)^{(30)}=e^x(x^3+90x^2+2610x+24360).$$ ●

4. Дифференциалы высших порядков

Как известно, выражение для дифференциала df функции f имеет вид

$$df=f'(x)dx.$$

При постоянном dx дифференциал dy является функцией независимой переменной x и, следовательно, от этой функции можно взять дифференциал, считая значение dx дифференциала независимой переменной неизменным: $d(df)=d(f'(x)dx)$. Этот дифференциал называют *дифференциалом второго порядка* и обозначают d^2f. Следовательно,

$$d^2f=d(df)=d(f'(x)dx).$$

Так как dx не зависит от x, то dx можно вынести за знак дифференциала:

$$d(f'(x)dx)=dx\cdot d(f'(x))=dx\cdot f''(x)dx=f''(x)dx^2.$$

где dx^2 — упрощенное обозначение для $(dx)^2$.

Таким образом, *дифференциал второго порядка равен произведению производной второго порядка на квадрат дифференциала независимой переменной:* $d^2f=f''(x)dx^2$.

Вообще *дифференциалом n-го порядка* называют дифференциал от дифференциала $(n-1)$-го порядка: $d^nf=d(d^{n-1}f)$ при постоянном значении dx. Мето-

---

* Г. Лейбниц (1646—1716) — немецкий философ и математик.

** И. Ньютон (1643—1727) — английский физик и математик.

дом математической индукции нетрудно доказать, что при любом натуральном значении n справедлива формула

$$d^n f = f^{(n)}(x)dx^n. \tag{3}$$

Замечание. Формула (3) верна, если x есть независимая переменная. В случае, когда x — промежуточный аргумент, например $x=\varphi(t)$, нельзя писать $d^n f = f^{(n)}dx^n$. В самом деле, $dx=\varphi'(t)dt$ и поэтому dx не является постоянной. Согласно правилу вычисления дифференциала произведения, имеем

$$d^2 f = d(df) = d(f'(x)dx) = d(f'(x))dx + f'(x)d(dx) =$$
$$= (f''(x)dx)dx + f'(x)d^2x = f''(x)dx^2 + f'(x)d^2x.$$

Лишь при условии, что $d^2x=0$, получаем $d^2f=f''(x)dx^2$. Это имеет место, если x — линейная функция от t: $x=kt+b$.

Таким образом, форма записи $f^{(n)}(x)dx^n$ для дифференциалов высших порядков неинвариантна.

Пример. Найти $d^5 f$ для функции e^{2x} при $x=0$, $dx=0{,}1$.

○ Имеем $f'=2e^{2x}$, $f''=4e^{2x}$, $f'''=8e^{2x}$, $f^{(4)}=16e^{2x}$, $f^{(5)}=32e^{2x}$. Таким образом,

$$d^5 f = f^{(5)}(x)dx^5 = 32e^{2x}dx^5.$$

При $x=0$, $dx=0{,}1$ получаем $d^5f=32e^0(0{,}1)^5=0{,}00032$.●

§ 33. Основные теоремы дифференциального исчисления

1. Связь между характером изменения функции и ее производной

До сих пор мы занимались исследованием функций, не пользуясь понятием производной. Производная помогает детальнее изучить свойства функций: установить интервалы возрастания и убывания функции; точки, в которых функция достигает наибольшего и наименьшего значений; интервалы выпуклости вверх и вниз графика функции и т. д. Исследованием функций с помощью производной мы займемся в следующих параграфах, в настоящем же параграфе будут получены необходимые для исследования функций теоремы.

Теорема 1. *Если $f(x)$ — дифференцируемая и неубывающая (невозрастающая) функция на промежутке X, то ее производная неотрицательна (неположительна) на X.*

□ Пусть $f(x)$ — неубывающая функция на X. Если $\Delta x > 0$, то $x < x + \Delta x$ и потому $f(x) \leqslant f(x+\Delta x)$, т. е. $\Delta f = f(x+\Delta x) - f(x) \geqslant 0$. Тогда $\frac{\Delta f}{\Delta x} \geqslant 0$. Аналогично, $\frac{\Delta f}{\Delta x} \geqslant 0$ при $\Delta x < 0$, а значит, $f'(x) = \lim\limits_{\Delta x \to 0}\frac{\Delta f}{\Delta x} \geqslant 0$. Случай, когда $f(x)$ — невозрастающая функция рассматривается аналогично.■

Теорема 1 имеет простой геометрический смысл: она означает, что касательная к графику неубывающей (невозрастающей) функции либо параллельна оси абсцисс, либо образует с положительным направлением этой оси острый (тупой) угол (рис. 112, *а* и *б*).

Для исследования функций более важную роль играет обратное утверждение, позволяющее определить характер изменения функции по знаку ее производной:

Если функция $f(x)$ непрерывна на промежутке X и имеет в каждой внутренней точке этого промежутка положительную (отрицательную) производную, то эта функция возрастает (убывает) на X.

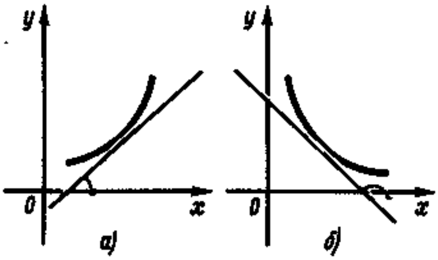


Рис. 112

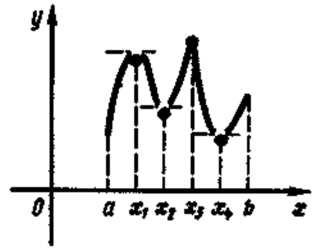


Рис. 113

Доказательство этого утверждения мы приведем позднее, а сейчас дадим лишь его физическое истолкование. Пусть точка $M(x)$ движется по прямой согласно закону $x = f(t)$. Ясно, что если $f'(t) > 0$ на отрезке $[a, b]$, т. е. скорость точки положительна в течение промежутка времени $[a, b]$, точка движется слева направо и, значит, функция f возрастает. Случай отрицательной производной рассматривается аналогично.

На рис. 113 изображен график непрерывной функции f, заданной на отрезке $[a, b]$. Значение функции в точке x_1 не является наибольшим на отрезке $[a, b]$, но если ограничиться сравнением $f(x_1)$ со значениями функции в точках, «близких» к x_1, то окажется, что $f(x_1)$ больше всех этих значений. Иными словами, существует такая окрестность U точки x_1, что для любого $x \in U$ выполняется неравенство $f(x) \leqslant f(x_1)$. Говорят, что функция f имеет в точке x_1 максимум.

В точке x_2 функция f принимает значение, меньшее значений функции в точках, «близких» к x_2. В таких случаях говорят, что функция в точке x_2 имеет минимум. Аналогично, функция f имеет минимум в точке x_4 и максимум в точке x_3.

Теперь дадим строгое определение этих понятий.

Определение. Пусть функция $f(x)$, заданная на множестве X, определена в некоторой окрестности точки $x_0 \in X$ и непрерывна в этой точке. Если для всех точек x этой окрестности выполняется неравенство $f(x) \leqslant f(x_0)$ $(f(x) \geqslant f(x_0))$, то x_0 называется *точкой максимума (минимума) функции* $f(x)$.

Точки максимума и минимума функции называют *точками экстремума* (от лат. extremum — «крайний»).

Определение точек максимума и минимума можно дать иначе, используя понятие приращения функции. В точке максимума при достаточно близких к нулю значениях приращения аргумента приращение функции отрицательно (как в случае, когда $\Delta x > 0$, так и в случае, когда $\Delta x < 0$), а в точке минимума при достаточно близких к нулю значениях Δx это приращение положительно (см. рис. 113).

Снова обратимся к рис. 113. Замечаем, что в точках экстремума x_1, x_2, x_4 касательная к графику функции параллельна оси абсцисс, т. е. в каждой из этих точек $f'(x) = 0$. В точке же x_3 касательную провести нельзя и, значит, в этой точке производная не существует. Можно сделать предположение, что если функция $y = f(x)$ имеет экстремум в точке x_0, то в этой точке производная данной функции либо не существует, либо равна нулю (в дальнейшем это утверждение будет доказано). Геометрически это означает, что *в точке экстремума касательная к графику функции либо горизонтальна, либо не существует.*

Точки, в которых производная равна нулю или не существует, называют *критическими* (иногда точки, где производная равна нулю, называют *стационарными*).

Следует иметь в виду, что не всякая критическая точка является точкой экстремума. Например, функция $y = x^3$ монотонна, но имеет стационарную точку $x = 0$ (см. рис. 20).

2. Необходимое условие экстремума

Лемма 1. *Пусть в некоторой точке x_0 производная функции $f(x)$ положительна, т. е. $f'(x) > 0$. Тогда найдется такая окрестность этой точки, в которой знак приращения функции совпадает со знаком приращения ее аргумента, т. е. в указанной окрестности $\Delta f = f(x_0 + \Delta x) - f(x_0) > 0$, если $\Delta x > 0$, и $\Delta f < 0$, если $\Delta x < 0$.*

□ Так как в точке x_0 функция дифференцируема, то ее приращение в этой точке можно представить в виде $\Delta f = (f'(x_0) + \alpha)\Delta x$, где $\lim\limits_{\Delta x \to 0} \alpha = 0$. При этом по условию $f'(x_0) > 0$.

Поскольку $\lim\limits_{\Delta x \to 0} \alpha = 0$, найдется такая окрестность точки x_0, в каждой точке которой выполняется неравенство $|\alpha| < f'(x_0)$. Тогда $f'(x_0) + \alpha > 0$ и, следовательно, $\frac{\Delta f}{\Delta x} = f'(x_0) + \alpha > 0$, а потому знаки Δx и Δf совпадают. ■

Аналогично доказывается следующее утверждение.

Лемма 2. *Пусть в некоторой точке x_0 производная функции $f(x)$ отрицательна, т. е. $f'(x_0) < 0$. Тогда найдется такая окрестность этой точки, в которой знак приращения функции противоположен знаку приращения аргумента.*

Из лемм 1 и 2 вытекает следующая теорема (о ней уже шла речь в п. 1), выражающая необходимое условие того, что x_0 — точка экстремума функции.

> **Теорема 2.** *Пусть функция $f(x)$ непрерывна в некоторой окрестности точки x_0 и имеет экстремум в этой точке. Тогда производная функции в точке x_0 либо равна нулю, либо не существует.*

□ Возможны четыре случая: 1) $f'(x_0) > 0$; 2) $f'(x_0) < 0$; 3) $f'(x_0) = 0$, 4) производная в точке x_0 не существует.

Если $f'(x_0) > 0$, то в силу леммы 1 найдется такая окрестность точки x_0, в которой знаки Δx и Δf совпадают. Поэтому Δf может принимать в указанной окрестности как положительные, так и отрицательные значения. Но если x_0 — точка максимума (минимума), то существует ее окрестность, в которой Δf неположительно (неотрицательно). Полученное противоречие показывает, что точка, в которой $f'(x_0) > 0$, не может быть точкой экстремума.

Аналогично доказывается, что отпадает и второй случай; следовательно, остаются лишь третий и четвертый случаи, т. е. в точке экстремума производная либо равна нулю, либо не существует. ■

3. Теорема Ролля

> **Теорема 3 (теорема Ролля*).** *Пусть функция $f(x)$ удовлетворяет следующим условиям: 1) она непрерывна на отрезке $[a, b]$; 2) она дифференцируема в интервале (a, b); 3) на концах отрезка $[a, b]$ она принимает равные значения: $f(a) = f(b)$. Тогда в интервале (a, b) существует по крайней мере одна точка c, в которой производная равна нулю.*

□ По условию данная функция непрерывна на отрезке $[a, b]$; следовательно, на этом отрезке она принимает свое наименьшее значение m и свое на-

* М. Ролль (1652—1719) — французский математик.

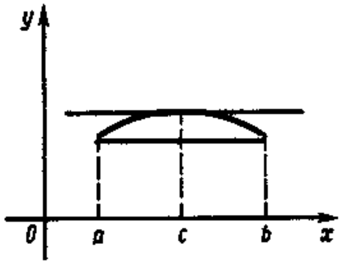


Рис. 114

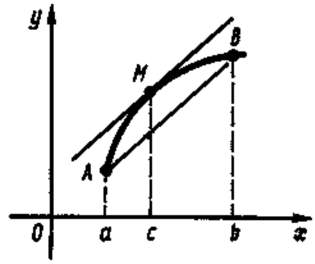


Рис. 115

нбольшее значение M. Возможны два случая: 1) значения m и M данная функция принимает на концах отрезка; 2) хотя бы одно из значений m или M данная функция принимает во внутренней точке отрезка.

Рассмотрим первый случай. Пусть $f(a)=m$, $f(b)=M$. Так как по условию $f(a)=f(b)$, то получаем $m=f(a)=f(b)=M$. Это означает, что наибольшее значение функции f на отрезке $[a, b]$ совпадает с наименьшим, и поэтому функция на этом отрезке постоянна. Но тогда всюду внутри отрезка производная равна нулю, и в качестве c можно выбрать любую точку интервала (a, b).

Рассмотрим второй случай, когда хотя бы одно из значений m или M достигается во внутренней точке отрезка; пусть для определенности $f(c)=M$, где $a<c<b$. Это означает, что c является точкой максимума функции, а тогда согласно теореме 2 в этой точке производная либо не существует, либо равна нулю. По условию, данная функция дифференцируема в любой внутренней точке отрезка; значит, она дифференцируема и в точке c. В таком случае возможно только, что $f'(c)=0$. ■

Геометрическая иллюстрация теоремы Ролля изображена на рис. 114.

4. Теорема Лагранжа

Производная позволяет характеризовать поведение функции лишь вблизи данной точки. Чтобы с ее помощью делать выводы о глобальном поведении функции, применяют теоремы, называемые теоремами о среднем значении. Важнейшей и чаще всего применяемой из них является теорема, о которой идет речь в настоящем пункте.

Проведем хорду AB графика функции $y=f(x)$ (рис. 115). Из рисунка видно, что на дуге AB найдется точка M, в которой касательная к данной кривой параллельна хорде AB. Следовательно, угловые коэффициенты этой касательной и прямой AB равны. Но $k_{\text{кас}}=f'(c)$, $k_{\text{хор}}=\frac{f(b)-f(a)}{b-a}$. Значит, внутри отрезка найдется такая точка c, в которой выполняется равенство $f'(c)=\frac{f(b)-f(a)}{b-a}$. Это утверждение и составляет суть теоремы Лагранжа. Дадим формулировку и строгое доказательство этой теоремы.

Теорема 4 (теорема Лагранжа*). *Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема во всех внутренних точках этого отрезка. Тогда в интервале (a, b) найдется точка c $(a<c<b)$, в которой выполняется равенство*

$$f'(c)=\frac{f(b)-f(a)}{b-a}. \qquad (1)$$

* Ж. Лагранж (1736—1813) — французский математик.

□ Пусть $y = kx + l$ — уравнение прямой, проходящей через точки $(a; f(a))$ и $(b; f(b))$. Как известно, угловой коэффициент k этой прямой равен $\frac{f(b)-f(a)}{b-a}$.

Рассмотрим функцию $F(x) = f(x) - (kx + l)$. Она является разностью двух непрерывных функций и, следовательно, непрерывна на отрезке $[a, b]$. Аналогично доказывается, что она дифференцируема в каждой внутренней точке отрезка. Наконец, при $x = a$ и при $x = b$ значения $f(x)$ и $kx + l$ совпадают, и, следовательно, в этих точках функция $F(x)$ обращается в нуль: $F(a) = 0$, $F(b) = 0$.

Итак, вспомогательная функция $F(x)$ удовлетворяет всем условиям теоремы Ролля; значит, в интервале (a, b) найдется по крайней мере одна точка c, в которой производная равна нулю. Так как $F'(x) = f'(x) - k$ и $k = \frac{f(b)-f(a)}{b-a}$, то

$$F'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}.$$

При $x = c$ получаем $0 = f'(c) - \frac{f(b)-f(a)}{b-a}$, откуда

$$f'(c) = \frac{f(b)-f(a)}{b-a}. \blacksquare$$

Геометрическая иллюстрация теоремы Лагранжа изображена на рис. 115.

Заметим, что если $f(a) = f(b)$, то из равенства (1) следует $f'(c) = 0$. Это означает, что теорема Ролля есть частный случай теоремы Лагранжа.

В формуле (1) фигурирует неизвестное число c. Тем не менее, как мы увидим позднее, это не является препятствием для многочисленных применений этой формулы в математическом анализе.

Пример. Доказать, что $\ln x_2 - \ln x_1 < x_2 - x_1$, где $1 < x_1 < x_2$.

○ Функция $\ln x$ на отрезке $[x_1, x_2]$ удовлетворяет всем условиям теоремы Лагранжа. Так как $(\ln x)' = 1/x$, то, применяя к функции $\ln x$ на отрезке $[x_1, x_2]$ формулу (1), получим

$$\ln x_2 - \ln x_1 = \frac{1}{c}(x_2 - x_1).$$

Поскольку $c > 1$, имеем $1/c < 1$. Значит,

$\frac{1}{c}(x_2 - x_1) < x_2 - x_1$, откуда $\ln x_2 - \ln x_1 < x_2 - x_1$. ●

5. Теорема Коши

В этом пункте рассмотрим теорему, которая также относится к числу теорем о среднем и является обобщением теоремы Лагранжа.

Теорема 5 (теорема Коши*). *Пусть на отрезке $[a, b]$ заданы две функции $f(x)$ и $g(x)$, причем: 1) они непрерывны на отрезке $[a, b]$; 2) они дифференцируемы в интервале (a, b); 3) $g'(x)$ ни в одной точке интервала (a, b) не обращается в нуль. Тогда в (a, b) найдется точка c, такая, что*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}. \qquad (2)$$

□ Составим вспомогательную функцию $F(x) = f(x) - \lambda g(x)$, где λ — число. Так как функции f и g по условию непрерывны на $[a, b]$ и дифференцируемы

* О. Коши (1789—1857) — французский математик.

в (a, b), то и функция F непрерывна на $[a, b]$ и дифференцируема в (a, b). Значит, функция F удовлетворяет первым двум условиям теоремы Ролля. Подберем постоянный множитель λ так, чтобы функция F удовлетворяла и третьему условию теоремы Ролля: $F(a)=F(b)$. Для этого должно выполняться равенство $f(a)-\lambda g(a)=f(b)-\lambda g(b)$, т. е.

$$\lambda(g(b)-g(a))=f(b)-f(a). \tag{3}$$

Заметим, что $g(b)-g(a)\neq 0$. В самом деле, если $g(b)-g(a)=0$, то, применяя к функции g на отрезке $[a, b]$ теорему Ролля, мы заключили бы, что существует точка $c_1 (a<c_1<b)$ такая, что $g'(c_1)=0$, а это противоречит условию доказываемой теоремы: $g'(x)\neq 0$ всюду в (a, b). Поэтому обе части равенства (3) можно разделить на коэффициент при λ, в результате чего получим

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

При найденном значении λ вспомогательная функция F удовлетворяет всем условиям теоремы Ролля. Тогда по теореме Ролля в (a, b) найдется по крайней мере одна точка c такая, что $F'(c)=0$. Так как $F'(x)=f'(x)-\lambda g'(x)$, то, учитывая, что $F'(c)=0$, получим $f'(c)-\lambda g'(c)=0$, откуда $\lambda=\frac{f'(c)}{g'(c)}$, т. е.

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

Заметим, что теорема Лагранжа является частным случаем теоремы Коши. Если в формуле (2) положить $g(x)=x$, то получим формулу Лагранжа $\frac{f(b)-f(a)}{b-a}=f'(c)$.

6. Формула Тейлора

В § 32 мы вывели формулу Тейлора для многочлена n-й степени:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{4}$$

Если $f(x)$ является не многочленом, а какой-либо другой функцией, имеющей в точке $x=a$ производные n-го порядка, то в формуле (4) нельзя написать знак равенства: $f(x)$ отличается от многочлена $S_n(x)$, входящего в правую часть этой формулы. Обозначив разность $f(x)-S_n(x)$ через $r_n(x)$, получим, что для любой функции $f(x)$, имеющей в точке $x=a$ производные любого порядка, справедлива формула

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+r_n(x). \tag{5}$$

Она называется *формулой Тейлора с остаточным членом* $r_n(x)$.

Разумеется, формула (5) окажется полезной, если удастся найти удобное выражение для остаточного члена $r_n(x)$. Этим мы и займемся ниже, рассмотрев сначала случай, когда $a=0$; в этом случае формула (5) принимает вид

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+r_n(x). \tag{6}$$

Положим

$$S_n(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n \tag{7}$$

и назовем этот многочлен *многочленом Тейлора*. Рассмотрим две леммы о свойствах многочлена Тейлора и остаточного члена $r_n(x)$.

Лемма 1. *Если $S_n(x)$ — многочлен Тейлора для функции $f(x)$, то $S_n(0)=f(0)$, $S_n'(0)=f'(0)$, $S_n''(0)=f''(0)$, ..., $S_n^{(n)}(0)=f^{(n)}(0)$.*

□ Применяя к $S_n(x)$ формулу Тейлора для многочлена (см. п. 2 § 32), получим

$$S_n(x)=S_n(0)+\frac{S_n'(0)}{1!}x+\frac{S_n''(0)}{2!}x^2+\ldots+\frac{S_n^{(n)}(0)}{n!}x^n \tag{8}$$

Сопоставляя правые части формул (7) и (8), приходим к выводу, что $S_n(0)=f(0)$, $S_n'(0)=f'(0)$, $S_n''(0)=f''(0)$, ..., $S_n^{(n)}(0)=f^{(n)}(0)$ (мы приравняли коэффициенты при одинаковых степенях переменных). ■

Следствие. *Для остаточного члена $r_n(x)$ справедливы равенства $r_n(0)=0$, $r_n'(0)=0$, $r_n''(0)=0$, ..., $r_n^{(n)}(0)=0$.*

Лемма 2. *Для остаточного члена $r_n(x)$ справедливо равенство $r_n^{(n+1)}(x)=f^{(n+1)}(x)$.*

□ В самом деле, $f(x)=S_n(x)+r_n(x)$. Но $(S_n(x))^{(n+1)}=0$ и, значит, $f^{(n+1)}(x)=r_n^{(n+1)}(x)$. ■

Итак, неизвестная пока функция $r_n(x)$ обладает следующими свойствами: $r_n(0)=r_n'(0)=\ldots=r_n^{(n)}(0)=0$; $r_n^{(n+1)}(x)=f^{(n+1)}(x)$. Используя эти свойства, найдем выражение для $r_n(x)$.

Рассмотрим на отрезке $[0, x]$ две функции: $h(x)=r_n(x)$ и $g(x)=x^{(n+1)}$. Не ограничивая общности рассуждений, будем считать, что $x>0$. Применяя к этим функциям теорему Коши, т. е. полученную в предыдущем пункте формулу (2), имеем

$$\frac{h(x)-h(0)}{g(x)-g(0)}=\frac{h'(c_1)}{g'(c_1)},\ \text{где}\ 0<c_1<x.$$

Учитывая, что $h(0)=r_n(0)=0$ и $g(0)=0$, получим

$$\frac{r_n(x)}{x^{n+1}}=\frac{r_n'(c_1)}{(n+1)c_1^n},$$

т. е.

$$r_n(x)=\frac{r_n'(c_1)}{(n+1)c_1^n}x^{(n+1)}. \tag{9}$$

Теперь применим теорему Коши к функциям $h_1(x)=r_n'(x)$ и $g_1(x)=x^n$ на отрезке $[0, c_1]$. Имеем

$$\frac{h_1(c_1)-h_1(0)}{g_1(c_1)-g_1(0)}=\frac{h_1'(c_2)}{g_1'(c_2)},\ \text{где}\quad 0<c_2<c_1.$$

Учитывая, что $h_1(0)=r_n'(0)=0$ и $g_1(0)=0$, получим

$$\frac{r_n'(c_1)}{c_1^n}=\frac{r_n''(c_2)}{nc_2^{n-1}}. \tag{10}$$

Сопоставляя формулы (9) и (10), находим

$$r_n(x)=\frac{r_n''(c_2)}{(n+1)nc_2^{n-1}}x^{n+1}.$$

Применяя теорему Коши к функциям $h_2(x) = r_n''(x)$ и $g_2(x) = x^{n-1}$ на отрезке $[0, c_2]$, затем к функциям $h_3(x) = r_n'''(x)$ и $g_3(x) = x^{n-2}$ на отрезке $[0, c_3]$ и т. д., в результате придем к формуле

$$r_n(x) = \frac{r_n^{(n)}(c_n)}{(n+1)!c_n} x^{n+1}. \tag{11}$$

Далее, применим теорему Лагранжа к функции $r_n^{(n)}(x)$ на отрезке $[0, c_n]$. Получим

$$\frac{r_n^{(n)}(c_n) - r_n^{(n)}(0)}{c_n - 0} = r^{(n+1)}(c), \text{ где } 0 < c < c_n.$$

Учитывая, что $r_n^{(n)}(0) = 0$, а $r^{(n+1)}(c) = f^{(n+1)}(c)$, имеем

$$\frac{r_n^{(n)}(c_n)}{c_n} = f^{(n+1)}(c). \tag{12}$$

Сопоставляя формулы (11) и (12), находим

$$r_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1}.$$

Воспользовавшись полученным выражением для остаточного члена $r_n(x)$, перепишем формулу (6) следующим образом:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}. \tag{13}$$

Здесь $0 < c < x$.

Положим в этой формуле $x = t - a$. Тогда $f(x) = f(t-a) = \varphi(t)$, $f(0) = \varphi(a)$, $f'(0) = \varphi'(a)$, $f''(0) = \varphi''(a)$, ..., $f^{(n)}(0) = \varphi^{(n)}(a)$, $f^{(n+1)}(c) = \varphi^{(n+1)}(\tilde{c})$, где $\tilde{c} = a + c$, т. е. $a < \tilde{c} < x$. Тогда формула (13) примет вид

$$\varphi(t) = \varphi(a) + \frac{\varphi'(a)}{1!}(t-a) + \frac{\varphi''(a)}{2!}(t-a)^2 + \ldots$$

$$\ldots + \frac{\varphi^{(n)}(a)}{n!}(t-a)^n + \frac{\varphi^{(n+1)}(\tilde{c})}{(n+1)!}(t-a)^{n+1}.$$

Возвращаясь к привычным обозначениям, перепишем эту формулу в виде

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \ldots$$

$$\ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\tilde{c})}{(n+1)!}(x-a)^{n+1}. \tag{14}$$

Здесь $a < c < x$ (или $x < c < a$).

Эта формула называется *формулой Тейлора с остаточным членом в форме Лагранжа:* $r_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$.

7. Оценка точности приближенного равенства $y \approx dy$

В § 29, преобразовав равенство $\Delta y \approx dy$ к виду

$$f(x) \approx f(a) + f'(a)(x-a), \tag{15}$$

мы не рассматривали вопрос о точности этого приближенного равенства. Сделаем это теперь, используя полученную в п. 6 формулу Тейлора.

Полагая в формуле (14) $n = 1$, получим

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(c)}{2!}(x-a)^2.$$

Отсюда сразу следует, что абсолютная погрешность приближенного равенства (15) не превышает $\frac{|f''(c)|}{2}(x-a)^2$.

Оценим, например, точность приближенного равенства $2{,}008^3 \approx 8{,}096$ (см. пример из п. 3 § 29). Оно было составлено для функции $f(x)=x^3$ на отрезке $[2;\ 2{,}008]$. Имеем $f'(x)=3x^2$, $f''(x)=6x$. На рассматриваемом отрезке $|f''(x)|\leqslant 12{,}5$. Значит,

$$|r(x)|\leqslant\frac{1}{2}12{,}5\cdot 0{,}008^2<\frac{1}{2}12{,}5\cdot 0{,}01^2 \doteq <7\cdot 0{,}0001=0{,}0007<0{,}001.$$

Таким образом, $2{,}008^3\approx 8{,}096$ по крайней мере с точностью 0,001.

§ 34. Применение производной к исследованию функций

1. Условия постоянства функции

Теорема 1. *Если функция $f(x)$ непрерывна на отрезке $[a,\ b]$ и во всех внутренних точках отрезка ее производная равна нулю, то функция $f(x)$ постоянна на этом отрезке.*

□ Пусть x — точка из промежутка $(a,\ b)$. Тогда в силу теоремы Лагранжа $\frac{f(x)-f(a)}{x-a}=f'(c)$, где $a<c<x$. Но по условию $f'(c)=0$ и, следовательно $\frac{f(x)-f(a)}{x-a}=0$, откуда $f(x)=f(a)$. Это и означает, что рассматриваемая функция постоянна на данном отрезке. ■

Доказанное утверждение имеет простой физический смысл: *если скорость точки все время равна нулю, то точка находится в покое и ее координата не меняется (постоянна).*

Из теоремы 1 вытекает следующая теорема.

Теорема 2. *Если функции $f(x)$ и $\varphi(x)$ непрерывны на отрезке $[a,\ b]$ и имеют равные производные во всех внутренних точках отрезка, то разность этих функций постоянна: $f(x)-\varphi(x)=C$.*

□ Введем вспомогательную функцию $g(x)$, равную разности функций: $g(x)=f(x)-\varphi(x)$. Найдем производную функции $g(x)$. Поскольку $f'(x)=\varphi'(x)$, имеем $g'(x)=f'(x)-\varphi'(x)=0$. Далее, так как $g'(x)=0$ на $(a,\ b)$, то на отрезке $[a,\ b]$ функция постоянна, т. е. $f(x)-\varphi(x)=C$. ■

2. Исследование функции на монотонность

В предыдущем параграфе исходя из геометрических и физических соображений мы установили связь между знаком производной и характером монотонности функции на промежутке. Теперь соответствующие утверждения будут доказаны.

Теорема 3. *Если функция $f(x)$ непрерывна на отрезке $[a,\ b]$, а ее производная положительна на интервале $(a,\ b)$, то $f(x)$ возрастает на $[a,\ b]$.*

□ Рассмотрим две любые точки $x_1,\ x_2\in[a,\ b]$ такие, что $x_1<x_2$. Для функции $f(x)$ на отрезке $[x_1,\ x_2]$ выполняются условия теоремы Лагранжа: поэтому

$$f(x_2)-f(x_1)=f'(c)(x_2-x_1), \qquad (1)$$

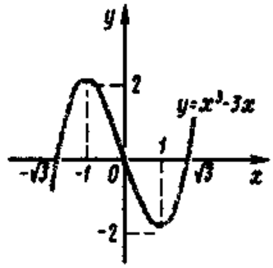


Рис. 116

где точка c лежит между x_1 и x_2. Так как оба множителя в правой части равенства (1) положительны ($f'(c)>0$ по условию, $x_2-x_1>0$ в силу выбора точек), то $f'(c)(x_2-x_1)>0$, а значит, и $f(x_2)-f(x_1)>0$. Итак, $f(x_1)<f(x_2)$ и, следовательно, функция $f(x)$ возрастает на $[a, b]$. ■

Аналогично доказывается следующая теорема.

Теорема 4. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, а ее производная отрицательна на (a, b), то $f(x)$ убывает на $[a, b]$.*

Заметим, что в формулировке обеих теорем вместо отрезка $[a, b]$ можно было взять произвольный промежуток X и потребовать непрерывности функции на X и дифференцируемости ее в любой внутренней точке промежутка X.

Теорема 3 (теорема 4) остается справедливой и в случае, когда $f'(x)\geqslant 0$ ($f'(x)\leqslant 0$), причем производная не обращается тождественно в нуль на каком-либо интервале из рассматриваемого промежутка X. В самом деле, пусть x_1, $x_2\in X$ и $x_1<x_2$. Если $f'(x)\geqslant 0$ всюду в X, то из равенства (1) следует $f(x_1)\leqslant f(x_2)$. Предположим, что $f(x_1)=f(x_2)$. Так как при $x_1<x<x_2$ выполняются неравенства $f(x_1)\leqslant f(x)$ и $f(x)\leqslant f(x_2)$, то на основании равенства $f(x_1)=f(x_2)$ получим $f(x)=f(x_1)$. Тогда функция f постоянна на (x_1, x_2) и, следовательно, ее производная вопреки условию обращается в нуль на целом интервале. Значит, $f(x_1)\neq f(x_2)$, т. е. $f(x_1)<f(x_2)$, и потому функция f возрастает на промежутке X.

Примеры. 1. Исследовать на монотонность функцию x^3-3x и построить ее график.

○ Находим

$$f'(x)=(x^3-3x)'=3x^2-3=3(x-1)(x+1).$$

На луче $(-\infty, -1)$ выполняется неравенство $f'(x)>0$, на интервале $(-1, 1)$ — неравенство $f'(x)<0$, а на луче $(1; +\infty)$ — неравенство $f'(x)>0$. Значит, функция возрастает на лучах $(-\infty, -1]$ и $[1, +\infty)$ и убывает на отрезке $[-1, 1]$.

Для построения графика найдем несколько контрольных точек. Прежде всего отметим, что если $x=-1$, то $y=2$, а если $x=1$, то $y=-2$ (это точки, в которых происходит изменение характера монотонности функции). Кроме того, из уравнения $x^3-3x=0$ найдем точки пересечения графика с осью абсцисс: $x_1=0$, $x_2=-\sqrt{3}$, $x_3=\sqrt{3}$. График изображен на рис. 116. ●

2. Доказать, что функция $\operatorname{arctg} x-x$ убывает на всей числовой прямой.

○ Имеем

$$(\operatorname{arctg} x-x)'=\frac{1}{1+x^2}-1=-\frac{x^2}{1+x^2}.$$

Так как при любом x справедливо неравенство $-\frac{x^2}{1+x^2}\leqslant 0$ и, кроме того, равенство $f'(x)=0$ выполняется только в одной точке $x=0$, то $f'(x)\leqslant 0$ на всей числовой прямой, причем $f'(x)=0$ в одной точке. Значит, функция убывает на всей числовой прямой. ●

3. Исследование функции на экстремум

В предыдущем параграфе были введены понятия максимума и минимума функции, сформулировано и доказано необходимое условие экстремума, согласно которому функция может иметь экстремум только в точках, где производная равна нулю или не существует. Однако необходимое условие не является достаточным для существования экстремума. Рассмотрим одно из достаточных условий экстремума.

Теорема 5. *Пусть функция $f(x)$ определена в точке x_0 и пусть существует $\delta > 0$ такое, что функция $f(x)$ непрерывна на отрезке $[x_0 - \delta, x_0 + \delta]$, дифференцируема на интервалах $(x_0 - \delta, x_0)$ и $(x_0, x_0 + \delta)$, причем производная данной функции сохраняет знак на каждом из этих интервалов. Тогда если на $(x_0 - \delta; x_0)$ и $(x_0, x_0 + \delta)$ знаки производной различны, то x_0 — точка экстремума, а если совпадают, то x_0 не является точкой экстремума. При этом если при переходе через точку x_0 производная меняет знак с плюса на минус*, то x_0 — точка максимума; если же производная меняет знак с минуса на плюс, то x_0 — точка минимума.*

□ Пусть производная $f'(x)$ положительна на интервале $(x_0 - \delta, x_0)$ и отрицательна на $(x_0, x_0 + \delta)$. Докажем, что x_0 — точка максимума функции.

По условию функция $f(x)$ непрерывна на отрезке $[x_0 - \delta, x_0]$, дифференцируема в интервале $(x_0 - \delta, x_0)$ и на этом интервале имеем $f'(x) > 0$. Значит, в силу теоремы 3 функция $f(x)$ возрастает на отрезке $[x_0 - \delta, x_0]$. Поэтому из неравенства $x < x_0$, где $x \in [x_0 - \delta, x_0]$, следует $f(x) < f(x_0)$. Аналогично устанавливаем, что функция $f(x)$ убывает на отрезке $[x_0, x_0 + \delta]$, а потому из неравенства $x_0 < x$, где $x \in [x_0, x_0 + \delta]$, следует $f(x_0) > f(x)$.

Таким образом, в δ-окрестности точки x_0 для точек x, отличных от x_0, выполняется неравенство $f(x) < f(x_0)$. Это и означает, что x_0 — точка максимума функции $f(x)$.

Рассмотрим случай, когда производная не меняет знак при переходе через точку x_0; пусть она отрицательна как слева, так и справа от x_0. Тогда функция убывает как на отрезке $[x_0 - \delta, x_0]$, так и на отрезке $[x_0, x_0 + \delta]$. В этом случае x_0 не является точкой экстремума.

Остальные два случая рассматриваются аналогично. ■

АЛГОРИТМ
исследования функции $y = f(x)$ на экстремум

1°. Найти производную $f'(x)$.

2°. Найти стационарные и критические точки, т. е. точки, в которых функция непрерывна, а производная равна нулю или не существует.

3°. Рассмотреть окрестность каждой из найденных точек, не содержащую других стационарных и критических точек, и исследовать знак производной слева и справа от этой точки.

4°. Используя достаточные условия экстремума, сделать соответствующие выводы.

---

* Т. е. если слева от x_0 выполняется неравенство $f'(x) > 0$, а справа от x_0 — неравенство $f'(x) < 0$.

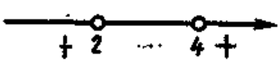


Рис. 117

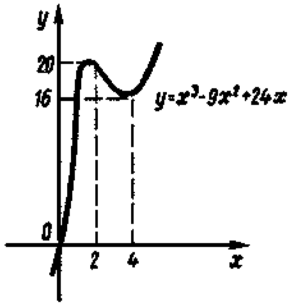


Рис. 118

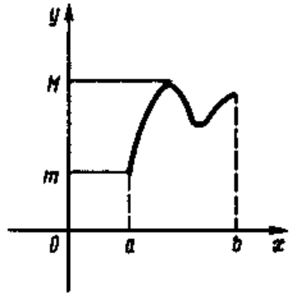


Рис. 119

Пример 1. Исследовать на экстремум функцию $y = x^3 - 9x^2 + 24x$.

○ 1°. Имеем $y' = 3x^2 - 18x + 24 = 3(x-2)(x-4)$.

2°. Приравняв производную нулю, находим $x_1 = 2$, $x_2 = 4$. В данном случае производная определена всюду; значит, кроме двух найденных точек, других стационарных и критических точек нет.

3°. Знак производной $y' = 3(x-2)(x-4)$ изменяется в зависимости от промежутка так, как показано на рис. 117. При переходе через точку $x = 2$ производная меняет знак с плюса на минус, а при переходе через точку $x = 4$ — с минуса на плюс.

4°. В точке $x = 2$ функция имеет максимум $y_{max} = 20$, а в точке $x = 4$ — минимум $y_{min} = 16$. На рис. 118 изображен эскиз графика функции $y = x^3 - 9x^2 + 24x$. ●

Отметим еще одно достаточное условие экстремума функции.

Теорема 6. *Пусть $f'(x_0) = 0$ и в точке x_0 существует $f''(x_0)$. Тогда если $f''(x_0) > 0$, то x_0 — точка минимума, а если $f''(x_0) < 0$, то x_0 — точка максимума функции $f(x)$.*

□ Рассмотрим случай, когда $f''(x_0) > 0$. Так как $f''(x)$ есть производная для функции $f'(x)$, то по лемме о знаке приращения функции (см. п. 2 § 33) найдется такая окрестность точки x_0, в которой знак приращения функции $f'(x)$, т. е. знак разности $f'(x) - f'(x_0)$, совпадает со знаком Δx. Но $f'(x_0) = 0$ и, значит, в указанной окрестности знак $f'(x)$ совпадает со знаком Δx т. е. при переходе через точку x_0 производная $f'(x)$ меняет знак с минуса на плюс. Тогда, согласно теореме 5, x_0 — точка минимума. Аналогично доказывается, что из $f''(x_0) < 0$ следует, что x_0 — точка максимума. ■

Пример 2. Используя теорему 6, исследовать на экстремум функцию $y = x^3 - 9x^2 + 24x$.

○ Имеем $y' = 3x^2 - 18x + 24$. Производная обращается в нуль в точках $x_1 = 2$, $x_2 = 4$. Далее, находим $y'' = 6x - 18$; $y''(2) = -6 < 0$, $y''(4) = 6 > 0$. Значит, $x = 2$ — точка максимума, а $x = 4$ — точка минимума функции (ср. с решением примера 1). ●

4. Отыскание наибольшего и наименьшего значений непрерывной функции на промежутке

Пусть функция $y = f(x)$ непрерывна на отрезке $[a, b]$. Тогда на этом отрезке она достигает своих наибольшего и наименьшего значений (см. теорему 5 § 18).

Наибольшее значение M и наименьшее значение m непрерывной функции могут достигаться как внутри отрезка, так и на его концах (рис. 119). Если своего наибольшего (наименьшего) значения функция достигает во внутренней точке отрезка, то такая точка является точкой экстремума; впрочем, для практических приложений достаточно того, что эта точка — критическая или стационарная.

АЛГОРИТМ
отыскания наибольшего и наименьшего значений непрерывной функции $y = f(x)$ на отрезке $[a, b]$

1°. Найти $f'(x)$.
2°. Найти точки, в которых $f'(x) = 0$ или $f'(x)$ не существует, и отобрать из них те, которые лежат внутри отрезка $[a, b]$.
3°. Вычислить значения функции $y = f(x)$ в точках, полученных в п. 2°, а также на концах отрезка и выбрать из них наибольшее и наименьшее: они и являются соответственно наибольшим ($y_{наиб}$) и наименьшим ($y_{наим}$) значением функции на отрезке $[a, b]$.

Пример 1. Найти наибольшее и наименьшее значения непрерывной функции $y = x^3 - 3x^2 - 45x + 225$ на отрезке $[0, 6]$.

○ 1°. Имеем $y' = 3x^2 - 6x - 45$.

2°. Производная y' существует при всех x. Найдем точки, в которых $y' = 0$; получим $3x^2 - 6x - 45 = 0$; $x^2 - 2x - 15 = 0$; $x_1 = -3$; $x_2 = 5$. Отрезку $[0, 6]$ принадлежит лишь точка $x = 5$.

3°. Вычислим значения функции в точках $x = 0$, $x = 5$ и $x = 6$:

x	0	5	6
y	225	50	63

Наибольшим из найденных значений функции является число 225, а наименьшим — число 50. Итак, $y_{наиб} = 225$, $y_{наим} = 50$. ●

Рассмотрим теперь задачу об отыскании наибольшего или наименьшего значения непрерывной функции на незамкнутом промежутке. Эта задача не всегда имеет решение. Так, на рис. 120 изображены графики непрерывных на интервале (a, b) функций. Функция $y = f_1(x)$ достигает как наибольшего, так и наименьшего значений (рис. 120, *а*); функция $y = f_2(x)$ достигает наибольшего значения, но не имеет наименьшего значения на (a, b) (рис. 120, *б*); функция $y = f_3(x)$ на (a, b) не имеет ни наибольшего, ни наименьшего значений (рис. 120, *в*).

Если для непрерывной на интервале (a, b) функции $y = f(x)$ требуется найти наибольшее и наименьшее значения, то поступают так же, как и при нахождении этих значений на отрезке $[a, b]$. Однако вместо вычисления значений функции на концах отрезка находят односторонние пределы функции при приближении к концам интервала.

Пример 2. Найти наименьшее значение функции $y = \left(\frac{2+\cos x}{\sin x}\right)^2$ на интервале $(0, \pi)$.

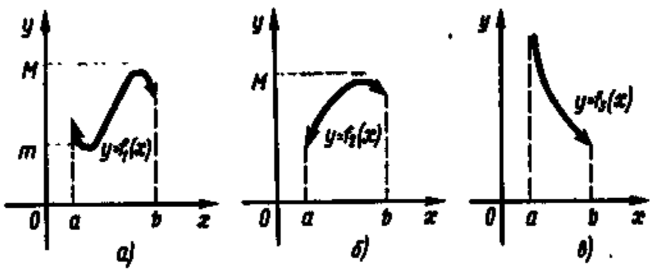


Рис. 120

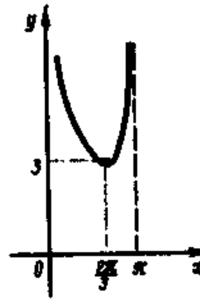


Рис. 121

○1°. Найдем производную

$$y' = 2\left(\frac{2+\cos x}{\sin x}\right)\left(\frac{2+\cos x}{\sin x}\right)' = 2\frac{2+\cos x}{\sin x}\times$$

$$\times\ \frac{(2+\cos x)'\sin x-(2+\cos x)(\sin x)'}{\sin^2 x} = 2\cdot\frac{2+\cos x}{\sin x}\cdot\frac{(-\sin x)\sin x-(2+\cos x)\cos x}{\sin^2 x} =$$

$$= \frac{2(2+\cos x)(-\sin^2 x-\cos^2 x-2\cos x)}{\sin^3 x} = \frac{-2(2+\cos x)(1+2\cos x)}{\sin^3 x}.$$

2°. Имеем $y'=0$, если $1+2\cos x=0$ или $2+\cos x=0$. Второе уравнение не имеет решений, так как $|\cos x|\leqslant 1$, а из первого находим $\cos x=-\frac{1}{2}$, т. е. $x=$ $=\pm\frac{2\pi}{3}+2\pi k$. Из этих значений интервалу $(0,\ \pi)$ принадлежит лишь значение $x=$ $=2\pi/3$.

Производная y' не существует, если $\sin^3 x=0$. Однако на $(0,\ \pi)$ это уравнение не имеет решений.

Итак, внутри интервала $(0,\ \pi)$ функция имеет лишь одну стационарную точку $x=2\pi/3$.

3°. Если $x=2\pi/3$, то

$$y=\left(\frac{2+\cos\frac{2\pi}{3}}{\sin\frac{2\pi}{3}}\right)^2=\left(\frac{2-\frac{1}{2}}{\frac{\sqrt{3}}{2}}\right)^2=3.$$

При приближении x к концам интервала, т. е. при $x\to 0$ или при $x\to\pi$, знаменатель дроби $\frac{(2+\cos x)^2}{\sin^2 x}$ стремится к 0, а числитель соответственно к 9 или к 1. Значит, и в том, и в другом случае $y\to+\infty$.

Поскольку при приближении к концам интервала $(0,\ \pi)$ значения функции неограниченно увеличиваются, функция достигает наименьшего значения в единственной стационарной точке, т. е. в точке $x=2\pi/3$. Итак, $y_{наим}=3$ (рис. 121). ●

Иногда для отыскания наибольшего или наименьшего значений непрерывной функции $y=f(x)$ на промежутке X оказываются полезными следующие два утверждения.

1°. *Если функция $y=f(x)$ имеет в промежутке X только одну точку экстремума $x=a$, причем это точка максимума, то $f(a)$ — наибольшее значение функции на промежутке X.*

2°. *Если функция $y=f(x)$ имеет в промежутке X только одну точку экстремума $x=a$, причем это точка минимума, то $f(a)$ — наименьшее значение функции на промежутке X.*

Так, в рассмотренном выше примере функция имеет в интервале $(0,\ \pi)$ лишь одну стационарную точку $x = 2\pi/3$. При переходе через эту точку производная изменяет знак с минуса на плюс. Значит, $x = 2\pi/3$ — точка минимума, а потому $f(2\pi/3) = 3$ — наименьшее значение функции на интервале $(0, \pi)$.

5. Задачи на отыскание наибольших или наименьших значений величин

Задачи на отыскание наибольших и наименьших значений величин удобно решать по следующему плану:

1^0. Выделяют оптимизируемую величину (т. е. величину, наибольшее или наименьшее значение которой требуется найти) и обозначают ее через y (или S, p, r, R и т. д. в зависимости от содержания задачи).

2^0. Одну из неизвестных величин (сторону, угол и т. д.) считают независимой переменной и обозначают через x; устанавливают реальные границы изменения x в соответствии с условиями задачи.

3^0. Исходя из конкретных условий данной задачи выражают y через x и известные величины.

4^0. Для полученной на предыдущем этапе функции $y = f(x)$ находят наибольшее или наименьшее значение (в зависимости от требований задачи) в промежутке реального изменения x, найденном в п. 2^0.

5^0. Интерпретируют результат п. 4^0 для данной конкретной задачи.

На этапах $1^0 - 3^0$ составляют, как принято говорить, *математическую модель* задачи. Здесь часто успех решения зависит от удачного выбора независимой переменной. Важно, чтобы было сравнительно нетрудно выразить y через x. На этапе 4^0 составленную математическую модель чаще всего исследуют с помощью производной, реже элементарным способом. При таком исследовании содержание самой задачи не имеет значения. Лишь после окончания решения задачи в рамках составленной математической модели полученный результат интерпретируют для исходной задачи (этап 5^0).

Примеры. 1. Шоссе пересекает местность с запада на восток. В 9 км к северу от шоссе находится поисковая партия, а в 15 км к востоку от ближайшей на шоссе к поисковой партии точки расположен райцентр. Поисковая партия отправляет курьера-велосипедиста в райцентр. Каков должен быть маршрут следования курьера, чтобы он прибыл в райцентр в кратчайший срок, если известно, что по полю он едет со скоростью 8 км/ч, а по шоссе — со скоростью 10 км/ч?

○ Сделаем чертеж. На рис. 122 точка P означает местонахождение поисковой партии, прямая l — шоссе, B — райцентр, $PA = 9$ км, $AB = 15$ км, PMB — маршрут следования курьера, где положение точки M между A и B пока неизвестно.

1^0. Оптимизируемая величина — время t движения курьера из P в B; надо найти $t_{\text{наим}}$.

2^0. Положим $AM = x$. По смыслу задачи точка M может занять любое положение между A и B, не исключая самих точек A и B. Значит, реальные границы изменения x таковы: $0 \leqslant x \leqslant 15$.

3^0. Выразим t через x. Имеем $PM = \sqrt{PA^2 + AM^2} = \sqrt{81 + x^2}$. Этот путь велосипедист едет со скоростью 8 км/ч, т. е. время t_1, затраченное на этот путь, выражается формулой $t_1 = \frac{\sqrt{81 + x^2}}{8}$. Далее, $MB = 15 - x$. Этот путь велосипедист едет со скоростью 10 км/ч, т. е. время t_2, затраченное на этот путь, выражается формулой $t_2 = \frac{15 - x}{10}$. Суммарное время t, затраченное на весь путь, равно $t_1 + t_2$, т. е. $t = \frac{\sqrt{81 + x^2}}{8} + \frac{15 - x}{10}$.

4^0. Для функции $t = \frac{\sqrt{81 + x^2}}{8} + \frac{15 - x}{10}$ надо найти наименьшее значение на отрезке $[0, 15]$. Используем для этого алгоритм из п. 4.

Рис. 122

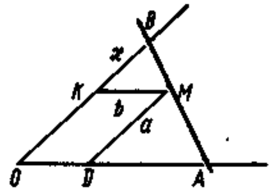


Рис. 123

1°) Находим

$$t' = \frac{1}{8}\cdot\frac{1}{2}(81+x^2)^{-1/2}\cdot 2x + \frac{1}{10}(-1) = \frac{x}{8\sqrt{81+x^2}} - \frac{1}{10}.$$

2°) Производная t' существует при всех x. Найдем точки, в которых $t' = 0$. Имеем

$$\frac{x}{\sqrt{81+x^2}} - \frac{1}{10} = 0;\ 5x = 4\sqrt{81+x^2};$$

$$25x^2 = 16(81+x^2);\ 9x^2 = 16\cdot 81;\ x^2 = 16\cdot 9;\ x = 4\cdot 3 = 12.$$

Значение $x = 12$ принадлежит отрезку $[0, 15]$.

3°) Составим таблицу значений функции, куда включим значения функции на концах отрезка и в найденной стационарной точке:

x	0	12	15
t	105/40	87/40	$5\sqrt{306}/40$

Следовательно, $t_{наим} = 87/40$ (поскольку $87 < 5\sqrt{306}$).

Четвертый этап решения задачи закончен; остается интерпретировать полученный результат применительно к исходной задаче.

5°. Так как $t_{наим}$ достигается при $x = 12$, то велосипедисту надо ехать по такому маршруту PMB, чтобы расстояние между точками A и M по шоссе было равно 12 км. ●

2. Через фиксированную точку M внутри угла провести прямую, отсекающую от угла треугольник наименьшей площади (рис. 123).

○ 1°. Оптимизируемая величина — площадь S треугольника AOB.

2°. Проведем $DM \parallel OB$, $MK \parallel OA$. Положим $KB = x$; реальные границы изменения x таковы: $0 < x < +\infty$.

3°. Поскольку M — фиксированная точка, отрезки DM и KM также фиксированы; положим $DM = a$, $KM = b$ и выразим S через x, a, b.

Из подобия треугольников MKB и AOB следует, что $\frac{MK}{AO} = \frac{KB}{OB}$, т. е. $\frac{b}{AO} = \frac{x}{a+x}$. Отсюда находим $AO = \frac{b(a+x)}{x}$.

Далее, имеем $S = \frac{1}{2}AO\cdot OB\cdot\sin\alpha$, где $\alpha = \angle AOB$. Значит,

$$S = \frac{1}{2}\cdot\frac{b(a+x)}{x}(a+x)\sin\alpha = \frac{b\sin\alpha}{2}\cdot\frac{(a+x)^2}{x}.$$

Математическая модель задачи составлена.

4°. Рассмотрим функцию

$$S = k\frac{(a+x)^2}{x},\ 0 < x < +\infty,$$

где $k = \frac{b\sin\alpha}{2}$. Найдем ее наименьшее значение.

1°) Имеем $S' = k\dfrac{2(a+x)x-(a+x)^2}{x^2} = k\dfrac{(a+x)(x-a)}{x^2}$.

2°) Производная не существует в точке $x=0$ и обращается в нуль в точках $x=-a$, $x=a$. Из этих трех точек промежутку $(0, +\infty)$ принадлежит лишь точка $x=a$.

3°) Как при $x \to 0$, так и при $x \to +\infty$ имеем $S \to +\infty$. Значит, наименьшее значение функции достигается в точке $x=a$.

5°. Вернемся к исходной геометрической задаче. Так как $x = KB = a$ и $OK = a$, то MK — средняя линия треугольника AOB; значит, M — середина AB. Таким образом, чтобы от сторон угла отсечь треугольник наименьшей площади, надо провести через точку M прямую так, чтобы ее отрезок, заключенный между сторонами угла, делился в точке M пополам. ●

6. Применение производной для доказательства неравенств

Доказательство неравенств с помощью производной связано чаще всего с исследованием функций на монотонность и экстремумы.

Примеры. 1. Доказать, что если $\alpha<\beta$, то $\alpha+\cos\alpha<\beta+\cos\beta$.

○ Рассмотрим функцию $f(x)=x+\cos x$ и найдем ее производную: $f'(x)=1-\sin x$. Так как $f'(x)\geqslant 0$, то функция $f(x)$ возрастает на всей числовой прямой. Значит, из $\alpha<\beta$ вытекает $f(\alpha)<f(\beta)$, т. е. $\alpha+\cos\alpha<\beta+\cos\beta$. ●

2. Доказать, что при всех x справедливо неравенство

$$x^5+(1-x)^5 \geqslant \frac{1}{16}.$$

○ Рассмотрим функцию $f(x)=x^5+(1-x)^5$ и исследуем ее на экстремум. Имеем

$$f'(x)=5x^4-5(1-x)^4=5(x^2-(1-x)^2)\,(x^2+(1-x)^2)=5(2x-1)(2x^2-2x+1);$$
$$f'(x)=0 \text{ при } x=1/2.$$

Других критических точек у функции нет (уравнение $2x^2-2x+1=0$ не имеет корней). Если $x<1/2$, то $f'(x)<0$, а если $x>1/2$, то $f'(x)>0$; значит, $x=1/2$ — точка минимума. Так как других точек экстремума данная непрерывная функция не имеет, то $f(1/2)$ — наименьшее значение функции (см. утверждение 2° из п. 4). Но $f\left(\frac{1}{2}\right) = (1/2)^5+(1/2)^5=1/16$. Итак, $f(x)\geqslant 1/16$, т. е. $x^5+(1-x)^5\geqslant 1/16$. ●

7. Исследование функции на выпуклость

На рис. 124 изображены графики трех функций. Первый из них обращен выпуклостью вверх, второй — выпуклостью вниз, а третий на одних участках обращен выпуклостью вверх, а на других — выпуклостью вниз. В этом пункте мы дадим строгие определения этих понятий и выясним, как определить направление выпуклости графика функции на данном отрезке.

Определение 1. Функция $f(x)$, непрерывная на отрезке $[a, b]$ и дифференцируемая в интервале (a, b), называется *выпуклой вверх* (*вниз*) на этом отрезке, если при $a\leqslant x\leqslant b$ ее график лежит не выше (не ниже) касательной, проведенной в любой точке $M(x_0; f(x_0))$, где $a\leqslant x_0\leqslant b$ (рис. 124).

Теорема 7. *Если функция $f(x)$ выпукла вверх (вниз) на отрезке $[a, b]$, то ее график при $a\leqslant x\leqslant b$ расположен не ниже (не выше) хорды AB, где $A(a; f(a))$, $B(b; f(b))$.*

Эту теорему иллюстрирует рис. 125. Заметим, что a и b можно заменить любыми x_1 и x_2 такими, что $a\leqslant x_1\leqslant x_2\leqslant b$ (на $[x_1, x_2]$ функция имеет то же направление выпуклости, что и на всем отрезке $[a, b]$).

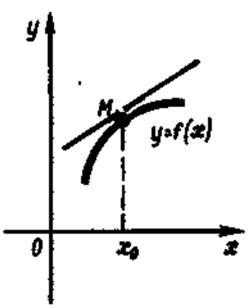


Рис. 124

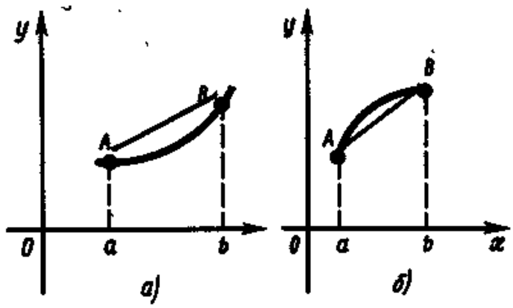


Рис. 125

□ Выберем любое $x_0(a \leqslant x_0 \leqslant b)$ и проведем касательную к графику функции $f(x)$ в точке $M(x_0;\ f(x_0))$. Так как функция выпукла вверх на отрезке $[a,\ b]$, то ее график при $a \leqslant x \leqslant b$ расположен не выше этой касательной. В частности, точки A и B расположены не выше касательной. Тогда и вся хорда AB расположена не выше касательной. Аналогично рассматривается случай, когда функция $f(x)$ выпукла вниз на $[a,\ b]$. ■

Если функция $f(x)$ дифференцируема дважды, то достаточные условия выпуклости вверх (вниз) для $f(x)$ можно сформулировать с помощью второй производной.

Теорема 8. *Пусть функция $f(x)$ имеет вторую производную в любой внутренней точке x отрезка $[a,\ b]$ и непрерывна на концах этого отрезка. Тогда если на интервале $(a,\ b)$ выполняется неравенство $f''(x) \geqslant 0$, то функция $f(x)$ выпукла вниз на отрезке $[a,\ b]$; если на $(a,\ b)$ выполняется неравенство $f''(x) \leqslant 0$, то функция $f(x)$ выпукла вверх на $[a,\ b]$.*

□ Рассмотрим случай, когда $f''(x) \geqslant 0$. Возьмем произвольную точку $x_0 \in (a,\ b)$ и запишем уравнение касательной к графику функции $y = f(x)$ в точке с абсциссой x_0:

$$y_{\text{кас}} = f'(x_0)(x - x_0) + f(x_0).$$

Здесь $y_{\text{кас}}$ — ордината точки, лежащей на касательной и имеющей абсциссу x. Если ординату точки, лежащей на кривой и имеющей ту же абсциссу x, обозначить $y_{\text{кр}}$, то получим

$$y_{\text{кр}} - y_{\text{кас}} = f(x) - f'(x_0)(x - x_0) - f(x_0). \qquad (2)$$

Так как по условию функция f непрерывна на $[a,\ b]$ и дифференцируема в $(a,\ b)$, то к разности $f(x) - f(x_0)$ можно применить формулу Лагранжа. Тогда равенство (2) примет вид

$$y_{\text{кр}} - y_{\text{кас}} = f'(c_1)(x - x_0) - f'(x_0)(x - x_0) = (f'(c_1) - f'(x_0))(x - x_0),$$

где c_1 — точка, лежащая между x и x_0.

Теперь применим формулу Лагранжа к разности $f'(c_1) - f'(x_0)$:

$$f'(c_1) - f'(x_0) = f''(c)(c_1 - x_0)$$

где точка c лежит между c_1 и x. Значит,

$$y_{\text{кр}} - y_{\text{кас}} = f''(c)(c_1 - x_0)(x - x_0). \qquad (3)$$

Точки c_1 и x лежат по одну сторону от точки x_0 и потому $(c_1-x_0)(x-x_0)>0$. Кроме того, по условию $f''(c)\geqslant 0$. В итоге получаем $y_{\text{кр}}\geqslant y_{\text{кас}}$, а это означает, что на отрезке $[a, b]$ график обращен выпуклостью вниз.

Аналогично доказывается, что в случае $f''(x)\leqslant 0$ функция выпукла вверх. ■

С соответствующими изменениями теорема верна для любого промежутка X.

Пример. Исследовать на выпуклость функцию $f(x)=x^{\alpha}$, $x>0$.

○Имеем $f''(x)=\alpha(\alpha-1)x^{\alpha-2}$. Если $\alpha=1$ или $\alpha=0$, то $f''(x)=0$; если $\alpha>1$ или $\alpha<0$, то $f''(x)>0$, а если $0<\alpha<1$, то $f''(x)<0$. Значит, график степенной функции x^{α} на луче $(0, +\infty)$ обращен выпуклостью вниз при $\alpha>1$ и при $\alpha<0$ и выпуклостью вверх при $0<\alpha<1$; при $\alpha=1$ или $\alpha=0$ имеем линейную функцию x или постоянную функцию x^0, которые могут считаться как выпуклыми вверх, так и выпуклыми вниз.●

8. Точки перегиба

До сих пор, говоря о направлении выпуклости непрерывной и дифференцируемой функции $f(x)$ на отрезке $[a, b]$, мы исследовали расположение графика функции относительно касательной к графику глобально, на всем отрезке $[a, b]$. Теперь взаимное расположение графика и касательной в некоторой окрестности точки касания мы изучим локально.

В большинстве случаев график локально располагается по одну сторону от касательной; так, на рис. 126, *а* он в достаточно малой окрестности точки касания лежит выше касательной, а на рис. 126, *б* — ниже касательной. Однако могут существовать точки, слева от которых в достаточно малой окрестности график лежит по одну сторону от касательной, а справа — по другую (рис. 126, *б*). Такие точки называются точками перегиба.

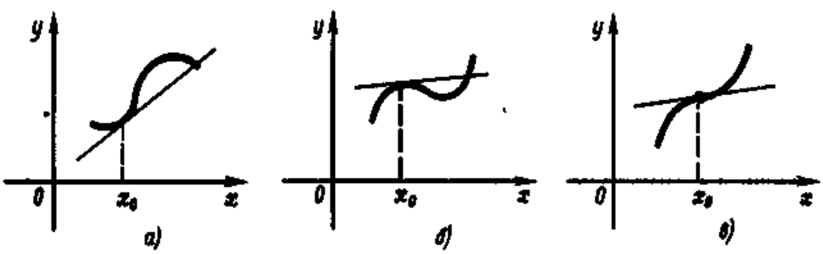


Рис. 126

Определение 2. Точка M линии Γ называется *точкой перегиба,* если существует дуга AB этой линии, содержащая точку M, такая, что дуги AM и MB расположены по разные стороны от касательной, проведенной в точке M к линии Γ.

Обычно точка перегиба отделяет участок, где функция выпукла вниз, от участка, где функция выпукла вверх.

Для отыскания точек перегиба полезна лемма, аналогичная лемме о знаке приращения функции (см. § 33, п. 2). В лемме о знаке приращения функции речь шла о знаке разности $f(x)-f(x_0)$, т. е. о знаке разности между ординатой кривой и ординатой горизонтальной прямой $y=f(x_0)$. В лемме, которая формулируется ниже, речь идет о знаке разности между ординатой графика и ординатой касательной.

Лемма. *Пусть функция $f(x)$ определена на промежутке X и имеет вторую производную в некоторой внутренней точке $x_0\in X$. Если $f''(x_0)>0$, то график функции $f(x)$ в некоторой окрестности точки x_0 расположен не ниже касатель-*

ной к графику в точке x_0; если $f''(x_0)<0$, то график в некоторой окрестности точки x_0 расположен не выше касательной.

□ Рассмотрим случай, когда $f''(x_0)>0$. Введем вспомогательную функцию

$$\Phi(x)=y_{кр}-y_{кас}=f(x)-f'(x_0)(x-x_0)-f(x_0).$$

Ясно, что $\Phi(x_0)=0$. Далее, имеем $\Phi'(x)=f'(x)-f'(x_0)$, $\Phi''(x)=f''(x)$. Значит, $\Phi'(x_0)=0$, $\Phi''(x_0)=f''(x_0)$. Так как по условию $f''(x_0)>0$, то и $\Phi''(x_0)>0$. Но если $\Phi'(x_0)=0$ и $\Phi'(x_0)>0$, то по теореме 6 § 34 x_0 — точка минимума функции $\Phi(x)$. Тогда в некоторой окрестности точки x_0 выполняется неравенство $\Phi(x)\geqslant\Phi(x_0)$, т. е. $\Phi(x)\geqslant 0$ (напомним, что $\Phi(x_0)=0$). Так как $\Phi(x)=y_{кр}-y_{кас}$, то получаем $y_{кр}\geqslant y_{кас}$. Это и означает, что в указанной окрестности график расположен не ниже касательной. Вторая часть леммы доказывается аналогично. ■

Теорема 9 (необходимое условие точки перегиба). *Для того чтобы график дифференцируемой в точке x_0 функции $f(x)$ имел перегиб в точке с абсциссой x_0, необходимо, чтобы в точке x_0 вторая производная либо не существовала, либо была равна нулю.*

□ Возможны четыре случая: 1) $f''(x_0)>0$; 2) $f''(x_0)<0$; 3) $f''(x_0)=0$; 4) $f''(x_0)$ не существует. Однако в случаях 1 и 2 согласно лемме график расположен по одну сторону от касательной, т. е. в этих случаях точка x_0 не может быть точкой перегиба. Значит, возможны лишь случаи 3 и 4, что и требовалось доказать. ■

Подчеркнем, что эта теорема дает необходимое, но не достаточное условие точки перегиба. Например, для функции $f(x)=x^4$ имеем $f''(x)=12x^2$. Вторая производная обращается в нуль в точке $x=0$, но это точка минимума функции, а не точка перегиба.

Теорема 10 (достаточное условие точки перегиба). *Пусть функция $f(x)$ дифференцируема в точке x_0 и пусть найдется окрестность $(x_0-\delta,\ x_0+\delta)$ этой точки такая, что в интервалах $(x_0-\delta,\ x_0)$ и $(x_0;\ x_0+\delta)$ существует вторая производная $f''(x)$, причем она сохраняет знак на каждом из этих интервалов. Тогда если на $(x_0-\delta;\ x_0)$ и $(x_0;\ x_0+\delta)$ знаки второй производной различны, то точка графика, имеющая абсциссу x_0, является точкой перегиба; если же знаки одинаковы, то перегиба нет.*

□ Воспользуемся равенством (3), полученным при доказательстве теоремы 8:

$$\Phi(x)=y_{кр}-y_{кас}=f''(c)(c_1-x_0)(x-x_0).$$

Так как точки x и c_1 лежат по одну сторону от точки x_0, то $(x-x_0)(c_1-x_0)>0$. Отсюда следует, что если знаки второй производной слева и справа от точки x_0 различны, то и знаки функции $\Phi(x)=y_{кр}-y_{кас}$ различны, а это означает, что график переходит с одной стороны касательной на другую, т. е. точка с абсциссой x_0 — точка перегиба. Если же по обе стороны от x_0 выполняется неравенство $f''(x)>0$, то и слева, и справа от этой точки получим $y_{кр}>y_{кас}$, а это означает, что график лежит не ниже касательной и перегиба нет. Случай, когда $f''(x)<0$, рассматривается аналогично. ■

Доказанные теоремы приводят к следующему правилу.

ПРАВИЛО
нахождения точек перегиба графика функции $f(x)$

1°. Найти точки, в которых $f''(x)$ не существует или обращается в нуль.

2°. Пусть x_0 — одна из найденных точек. Исследовать знак $f''(x)$ слева и справа от точки x_0 (разумеется, в достаточно малой окрестности точки x_0). На основании теоремы 10 сделать соответствующие выводы.

Пример. Найти точки экстремума и точки перегиба графика функции $3x^4-8x^3+6x^2+12$ и построить ее график.

○ Имеем

$$f'(x)=12x^3-24x^2+12x=12x(x-1)^2.$$

Очевидно, что $f'(x)=0$ при $x_1=0$ и $x_2=1$. Производная при переходе через точку $x=0$ изменяет знак с минуса на плюс, а при переходе через точку $x=1$ не меняет знака. Значит, $x=0$ — точка минимума ($y_{\min}=12$), а в точке $x=1$ экстремума нет.

Далее, находим

$$f''(x)=36x^2-48x+12=36(x-1)\left(x-\frac{1}{3}\right).$$

Вторая производная обращается в нуль в точках $x_2=1$ и $x_3=1/3$. Знаки второй производной изменяются следующим образом: на луче $(-\infty, 1/3)$ имеем $f''(x)>0$, на интервале $(1/3, 1)$ имеем $f''(x)<0$, на луче $(1, +\infty)$ имеем $f''(x)>0$. Следовательно, $x=1/3$ — точка перегиба графика функции (переход с выпуклости вниз на выпуклость вверх) и $x=1$ — также точка перегиба (переход с выпуклости вверх на выпуклость вниз). Если $x=1/3$, то $y=12\frac{11}{27}$; если $x=1$, то $y=13$.

Используя найденные контрольные точки $(0; 12)$, $\left(\frac{1}{3}; 12\frac{11}{27}\right)$ и $(1; 13)$, строим график функции (рис. 127). ●

9. Общая схема исследования свойств функции и построения ее графика

На практике для построения графика функции $y=f(x)$ иногда поступают так: из уравнения $y=f(x)$ находят несколько точек графика функции $y=f(x)$ и соединяют эти точки плавной кривой. Однако при таком методе легко пропустить какие-то важные особенности графика и допустить ошибку в построении.

Для построения графика функции нужно исследовать ее свойства. Прежде всего найти область определения функции, а затем исследовать функцию на четность и периодичность. Так как график четной функции симметричен относительно оси y, а график нечетной — относительно начала координат, то для четных и нечетных функций можно ограничиться исследованием их свойств лишь при $x\geqslant 0$. Если функция периодическая и T — ее основной период, то достаточно провести исследование ее свойств на промежутке длины T. Далее, полезно найти точки пересечения графика с осями координат и определить интервалы знакопостоянства функции. Если, скажем, на интервале (a, b) функция $y=f(x)$ принимает только положительные значения, то на этом интервале ее график лежит выше оси x. После этого нужно изучить поведение функции на границах области определения, установить характер точек разрыва (если они имеются), найти асимптоты. Наконец, следует найти точки экстремума и перегиба.

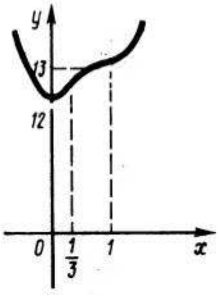


Рис. 127

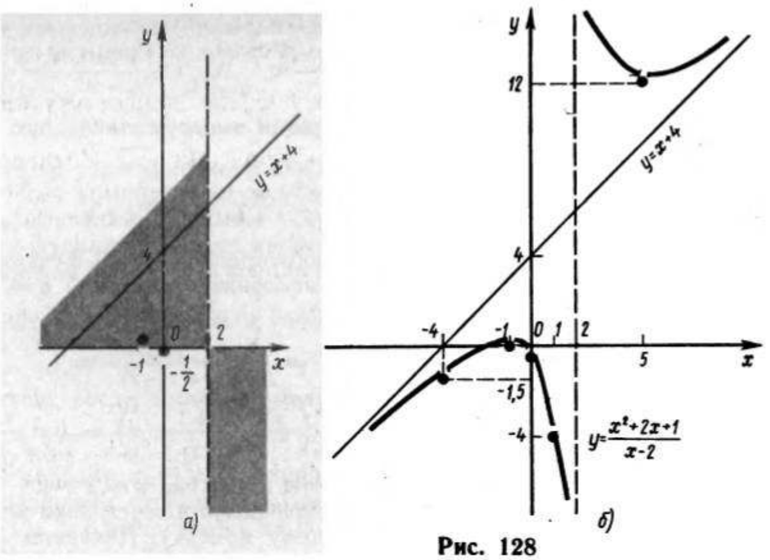


Рис. 128

Подводя итоги сказанному выше, получаем следующую схему.

СХЕМА
исследования свойств функции и построения ее графика

1^0. Найти область определения функции.
2^0. Исследовать функцию на периодичность.
3^0. Исследовать функцию на четность.
4^0. Найти точки пересечения графика с осями координат и определить интервалы знакопостоянства функции; найти точки разрыва и установить характер разрыва.
5^0. Исследовать поведение функции на границах области определения, найти асимптоты.
6^0. Найти промежутки возрастания и убывания функции, точки экстремума.
7^0. Исследовать направление выпуклости графика функции, найти точки перегиба.
8^0. Составить таблицу значений функции для некоторых значений ее аргумента.
9^0. Используя все полученные результаты, построить график функции.

Примеры. 1. Исследовать функцию $y=\dfrac{x^2+2x+1}{x-2}$ и построить ее график.

○ 1^0. Область определения функции: $x \neq 2$.
2^0. Функция не является периодической.
3^0. Функция не является ни четной, ни нечетной.
4^0. График пересекает оси в точках $(-1;\ 0)$ и $(0;\ -1/2)$. Точка разрыва $x=2$.
Отметим интервалы знакопостоянства функции: $(-\infty,\ -1)$ и $(-1,\ 2)$ — здесь функция принимает только отрицательные значения; $(2,\ +\infty)$ — здесь функция принимает только положительные значения.

5^0. Находим

$$\lim_{x\to 2-0}\frac{x^2+2x+1}{x-2}=-\infty,\quad \lim_{x\to 2+0}\frac{x^2+2x+1}{x-2}=+\infty.$$

Значит, график рассматриваемой функции имеет вертикальную асимптоту $x=2$.

Далее, имеем

$$\lim_{x\to\infty}\frac{x^2+2x+1}{x-2}=\infty.$$

Так как степень числителя данной дробно-рациональной функции на единицу больше степени знаменателя, то существует наклонная асимптота $y=kx+b$. Найдем коэффициенты k и b:

$$k=\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\frac{x^2+2x+1}{x(x-2)}=1,$$

$$b=\lim_{x\to\infty}(f(x)-kx)=\lim_{x\to\infty}\left(\frac{x^2+2x+1}{x-2}-x\right)=\lim_{x\to\infty}\frac{4x+1}{x-2}=4.$$

Значит, уравнение наклонной асимптоты имеет вид $y=x+4$.

На рис. 128,*а* представлена геометрическая иллюстрация тех сведений о графике, которыми мы располагаем к настоящему моменту. Показаны асимптоты, заштрихованы те участки координатной плоскости, где графика нет, отмечены известные точки графика. Это заготовка для будущего графика.

6^0. Найдем промежутки возрастания и убывания функции. Имеем

$$f'(x)=\left(\frac{x^2+2x+1}{x-2}\right)'=\left(\frac{(x+1)^2}{x-2}\right)'=\frac{2(x+1)(x-2)-(x+1)^2}{(x-2)^2}=\frac{(x+1)(x-5)}{(x-2)^2};$$

$f'(x)>0$ при $x>5$, $x<-1$; $f'(x)<0$ при $-1<x<2$; $2<x<5$. Значит, функция возрастает на лучах $(-\infty, -1]$, $[5, +\infty)$ и убывает на промежутках $[-1, 2)$ и $(2, 5]$.

Найдем точки экстремума функции: $x=-1$ — точка максимума, $y_{\max}=0$; $x=5$ — точка минимума, $y_{\min}=12$.

7^0. Исследуем функцию на выпуклость. Для этого найдем вторую производную данной функции:

$$f''(x)=\left(\frac{x^2-4x-5}{(x-2)^2}\right)'=\frac{(2x-4)(x-2)^2-2(x-2)(x^2-4x-5)}{(x-2)^4}=\frac{18}{(x-2)^3};$$

$f''(x)>0$ при $x>2$ — здесь функция выпукла вниз; $f''(x)<0$ при $x<2$ — здесь функция выпукла вверх.

8^0. Составим таблицу значений функции для некоторых значений аргумента:

x	-4	-1	0	1	5
$f(x)$	$-1{,}5$	0	$-0{,}5$	-4	12

9^0. Воспользовавшись полученными результатами, построим график функции (рис. 128, *б*).●

2. Исследовать функцию $\sin x-\frac{1}{2}\sin 2x$ и построить ее график.

○ 1^0. Область определения функции: $(-\infty, +\infty)$.

2^0. Функция является периодической, ее основной период равен 2π.

3^0. Функция нечетная.

Поскольку период функции равен 2π, достаточно провести исследование только от $-\pi$ до π, построить график функции на отрезке $[-\pi, \pi]$ и продолжить его на всю числовую прямую, пользуясь периодичностью. При этом так как функция является нечетной, то достаточно исследовать функцию и построить ее график на отрезке $[0, \pi]$, затем отобразить его симметрично относительно начала координат, а далее уже воспользоваться периодичностью.

Итак, дальнейшие исследования проведем для отрезка $[0, \pi]$.

4^0. Найдем точки пересечения графика с осью x. Для этого решим уравнение $\sin x - \frac{1}{2}\sin 2x = 0$. Имеем

$$\sin x - \sin x \cos x = 0,\ \sin x (1 - \cos x) = 0.$$

На отрезке $[0, \pi]$ последнее уравнение имеет два корня; $x_1 = 0$ и $x_2 = \pi$. Значит, график функции не пересекает ось абсцисс ни в какой внутренней точке отрезка $[0, \pi]$.

В интервале $(0, \pi)$ функция принимает только положительные значения.

5^0. Так как данная функция является непрерывной и периодической, то ее график не имеет асимптот. Значения функции на концах рассматриваемого отрезка $[0, \pi]$ таковы: $f(0) = 0$, $f(\pi) = 0$.

6^0. Найдем промежутки возрастания и убывания функции. Имеем

$$f'(x) = \cos x - \cos 2x.$$

Приравняв $f'(x)$ нулю, последовательно получим

$$\cos x - (1 + \cos 2x) + 1 = 0;\ 2\cos^2 x - \cos x - 1 = 0;\ \cos x = 1;\quad \cos x = -1/2.$$

Из первого уравнения находим $x_1 = 0$, из второго $x_2 = 2\pi/3$.

Точка $x_2 = 2\pi/3$ разбивает отрезок $[0, \pi]$ на отрезки $[0, 2\pi/3]$ и $[2\pi/3, \pi]$, причем неравенство $\cos x - \cos 2x \geqslant 0$ выполняется на первом из них. Значит, функция возрастает на $[0, 2\pi/3]$ и убывает на $[2\pi/3, \pi]$.

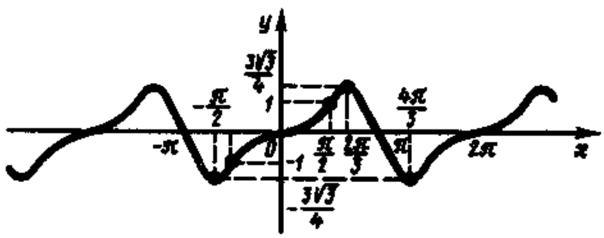


Рис. 129

Так как в точке $x = 2\pi/3$ возрастание функции сменяется ее убыванием, то при $x = 2\pi/3$ имеем точку максимума. Найдем значение функции в этой точке:

$$y_{max} = f\left(\frac{2\pi}{3}\right) = \sin\frac{2\pi}{3} - \frac{1}{2}\sin\frac{4\pi}{3} = \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{4} = \frac{3\sqrt{3}}{4}.$$

7^0. Исследуем функцию на выпуклость. Находим

$$f''(x) = -\sin x + 2\sin 2x.$$

Приравняв $f''(x)$ нулю и решив полученное уравнение, находим $x = \arccos(1/4)$ (остальные корни уравнения $f''(x) = 0$ не принадлежат интервалу $(0, \pi)$).

При переходе через точку $x = \arccos(1/4)$ вторая производная меняет знак с плюса на минус; следовательно, эта точка является точкой перегиба графика функции с выпуклости вниз на выпуклость вверх.

8^0. Составим таблицу значений функции для некоторых значений аргумента:

x	0	$\arccos(1/4)$	$\pi/2$	$2\pi/3$	π
$f(x)$	0	$\approx 3/4$	1	$3\sqrt{3}/4$	0

9^0. Пользуясь полученными результатами, построим график функции сначала на отрезке $[0, \pi]$, а затем на всей числовой прямой (рис. 129). ●

§ 35. Раскрытие неопределенностей при вычислении пределов по правилу Лопиталя

Выше мы уже говорили о различных способах раскрытия неопределенностей при вычислении пределов. Здесь мы рассмотрим еще один, связанный с применением производной. Его обычно называют *правилом Лопиталя**. Краткая формулировка этого правила такова: *в случае неопределенности вида* $0/0$ *или* ∞/∞ *предел отношения функций равен пределу отношения их производных.* На самом деле эта формулировка включает в себя четыре случая: неопределенность вида $0/0$ при $x \to a$ и при $x \to \infty$; неопределенность вида ∞/∞ при $x \to a$ и при $x \to \infty$. Ниже мы дадим строгие формулировки соответствующих теорем и доказательства первых двух из них.

Теорема 1. *Пусть функции $f(x)$ и $g(x)$ дифференцируемы в проколотой окрестности точки a (т. е. дифференцируемы во всех точках этой окрестности, за исключением, быть может, самой точки a), причем $g'(x)$ отлична от нуля, и пусть $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$. Тогда если существует $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}$, то существует и $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$, причем эти пределы равны:*

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

В случае, когда $f'(a) \neq 0$ и $g'(a) \neq 0$, геометрический смысл этой теоремы состоит в следующем. Графики функций $f(x)$ и $g(x)$ пересекают ось абсцисс в точке $M(a; 0)$, и потому уравнения касательных к этим графикам в точке M имеют вид $y = f'(a)(x-a)$ и $y = g'(a)(x-a)$. Теорема 1 фактически утверждает, что предел отношения функций $f(x)$ и $g(x)$ при $x \to a$ равен пределу отношения ординат касательных при $x \to a$, а этот предел и равен $\frac{f'(a)}{g'(a)}$.

Докажем теорему 1 в общем случае.

□ Доопределим функции $f(x)$ и $g(x)$, считая значение каждой из них в точке a равным нулю, т. е. $f(a) = g(a) = 0$. Так как $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$, то после этого функции $f(x)$ и $g(x)$ станут непрерывными в точке a и потому к ним можно применить теорему Коши (см. § 33, п. 5). Учитывая, что $f(a) = g(a) = 0$, получим

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

где точка c лежит между a и x. При $x \to a$ имеем $c \to a$ и, следовательно,

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

Теорема 2. *Пусть функции $f(x)$ и $g(x)$ дифференцируемы на луче $(a, +\infty)$, причем $g'(x) \neq 0$, и пусть $\lim\limits_{x\to+\infty} f(x) = \lim\limits_{x\to+\infty} g(x) = 0$. Тогда если существует $\lim\limits_{x\to+\infty}\frac{f'(x)}{g'(x)}$, то существует и $\lim\limits_{x\to+\infty}\frac{f(x)}{g(x)}$, причем эти пределы равны:*

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

* Г. Лопиталь (1661—1704) — французский математик.

□ Положим $x = 1/t$. Тогда $f(x) = f(1/t)$, $g(x) = g(1/t)$, $(f(1/t))' = f'(1/t) \times$ $\times(-1/t^2)$, $(g(1/t))' = g'(1/t)(-1/t^2)$. Далее, имеем

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{t \to +0} \frac{f(1/t)}{g(1/t)}.$$

Для вычисления предела $\lim\limits_{t \to +0} \frac{f(1/t)}{g(1/t)}$ воспользуемся теоремой 1:

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{t \to +0} \frac{f(1/t)}{g(1/t)} = \lim_{t \to +0} \frac{f'(1/t)(-1/t^2)}{g'(1/t)(-1/t^2)} =$$

$$= \lim_{t \to +0} \frac{f'(1/t)}{g'(1/t)} = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)}. \blacksquare$$

В теоремах 1 и 2 мы рассмотрели случаи раскрытия неопределенности вида $0/0$, когда $x \to a$ и когда $x \to +\infty$. Аналогично обстоит дело и в случае, когда $x \to -\infty$ или $x \to \infty$.

Теорема 3. *Пусть функции $f(x)$ и $g(x)$ дифференцируемы в некоторой проколотой окрестности точки a, причем $g'(x) \neq 0$, и пусть $\lim\limits_{x \to a} f(x) = \infty$, $\lim\limits_{x \to a} g(x) = \infty$. Тогда если существует $\lim\limits_{x \to a} \frac{f'(x)}{g'(x)}$, то существует и $\lim\limits_{x \to a} \frac{f(x)}{g(x)}$, причем $\lim\limits_{x \to a} \frac{f(x)}{g(x)} = \lim\limits_{x \to a} \frac{f'(x)}{g'(x)}$.*

Теорема 4. *Пусть функции $f(x)$ и $g(x)$ дифференцируемы на луче $(a, +\infty)$, причем $g'(x) \neq 0$, и пусть $\lim\limits_{x \to +\infty} f(x) = \infty$, $\lim\limits_{x \to +\infty} g(x) = \infty$. Тогда если существует $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$, то существует и $\lim\limits_{x \to +\infty} \frac{f(x)}{g(x)}$, причем $\lim\limits_{x \to +\infty} \frac{f(x)}{g(x)} = \lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$.*

Примеры. 1. Вычислить $\lim\limits_{x \to 2} \frac{e^x - e^2}{x - 2}$.

○ Это неопределенность вида $0/0$. Применяя правило Лопиталя, получим

$$\lim_{x \to 2} \frac{e^x - e^2}{x - 2} = \lim_{x \to 2} \frac{(e^x - e^2)'}{(x - 2)'} = \lim_{x \to 2} \frac{e^x}{1} = e^2.$$

Разумеется, используя здесь и в дальнейшем подобную краткую запись, мы предполагаем, что все условия соответствующей данному случаю теоремы выполнены и, в частности, что предел отношения производных существует. ●

2. Вычислить $\lim\limits_{x \to +\infty} \frac{x^2 + 6x}{e^{3x}}$.

○ Здесь имеем неопределенность вида ∞/∞. Воспользуемся правилом Лопиталя:

$$\lim_{x \to +\infty} \frac{x^2 + 6x}{e^{3x}} = \lim_{x \to +\infty} \frac{(x^2 + 6x)'}{(e^{3x})'} = \lim_{x \to +\infty} \frac{2x + 6}{3e^{3x}}.$$

Снова получили неопределенность вида ∞/∞. Так как условия теоремы 3 выполнены, то к полученному выражению применим еще раз правило Лопиталя:

$$\lim_{x \to +\infty} \frac{2x + 6}{3e^{3x}} = \lim_{x \to +\infty} \frac{(2x + 6)'}{(3e^{3x})'} = \lim_{x \to +\infty} \frac{2}{9e^{3x}} = 0.$$

Итак, $\lim\limits_{x \to +\infty} \frac{x^2 + 6x}{e^{3x}} = 0$ ●

Заметим, что, раскрывая неопределенности по правилу Лопиталя, следует использовать и те способы раскрытия неопределенностей, с которыми мы познакомились выше. Например, во многих случаях применение правила Лопиталя приводит к более простым выражениям, если предварительно заменить бесконечно малую эквивалентной ей или выполнить необходимые упрощения.

Примеры. 3. Вычислить $\lim\limits_{x\to 0}\dfrac{\sin x-\operatorname{tg} x}{\sin^3 x}$.

○ Так как $\sin x\sim x$ при $x\to 0$, то $\sin^3 x\sim x^3$ и, следовательно,

$$\lim_{x\to 0}\frac{\sin x-\operatorname{tg} x}{\sin^3 x}=\lim_{x\to 0}\frac{\sin x-\operatorname{tg} x}{x^3}.$$

Это неопределенность вида $0/0$. Применив правило Лопиталя, получим

$$\lim_{x\to 0}\frac{\sin x-\operatorname{tg} x}{x^3}=\lim_{x\to 0}\frac{(\sin x-\operatorname{tg} x)'}{(x^3)'}=\lim_{x\to 0}\frac{\cos x-\dfrac{1}{\cos^2 x}}{3x^2}=\lim_{x\to 0}\frac{\cos^3 x-1}{3x^2\cos^2 x}.$$

Это также неопределенность вида $0/0$ и для ее раскрытия снова применим правило Лопиталя. Однако прежде чем перейти к повторному дифференцированию, воспользуемся тем, что $\lim\limits_{x\to 0}3\cos^2 x=3$. Тогда получим

$$\lim_{x\to 0}\frac{\cos^3 x-1}{3x^2\cos^2 x}=\frac{1}{3}\lim_{x\to 0}\frac{\cos^3 x-1}{x^2}=\frac{1}{3}\lim_{x\to 0}\frac{(\cos^3 x-1)'}{(x^2)'}=$$
$$=\frac{1}{3}\lim_{x\to 0}\frac{3\cos^2 x(-\sin x)}{2x}=\frac{1}{3}\left(-\frac{3}{2}\right)\lim_{x\to 0}\frac{\sin x}{x}=-\frac{1}{2}.\ ●$$

4. Вычислить $\lim\limits_{x\to 1}(2-x)^{\operatorname{tg}\frac{\pi x}{2}}$.

○ Этот предел представляет собой неопределенность вида 1^∞. Для его вычисления поступим так: сначала найдем предел логарифма данной функции, а затем воспользуемся тем, что для непрерывной и принимающей только положительные значения функции $f(x)$ справедливо равенство

$$\lim_{x\to a}f(x)=e^{\ln\lim\limits_{x\to a}f(x)}=e^{\lim\limits_{x\to a}\ln f(x)}.$$

Итак, найдем предел логарифма данной функции:

$$\lim_{x\to 1}\ln(2-x)^{\operatorname{tg}\frac{\pi x}{2}}=\lim_{x\to 1}\operatorname{tg}\frac{\pi x}{2}\ln(2-x).$$

Здесь имеем неопределенность вида $0\cdot\infty$. Чтобы свести ее к виду $0/0$, воспользуемся тем, что $\operatorname{tg}\dfrac{\pi x}{2}=\dfrac{1}{\operatorname{ctg}\frac{\pi x}{2}}$. Тогда получим

$$\lim_{x\to 1}\frac{\ln(2-x)}{\operatorname{ctg}\frac{\pi x}{2}}=\lim_{x\to 1}\frac{(\ln(2-x))'}{\left(\operatorname{ctg}\frac{\pi x}{2}\right)'}=\lim_{x\to 1}\frac{\dfrac{1}{2-x}(-1)}{-\dfrac{1}{\sin^2\frac{\pi x}{2}}\cdot\dfrac{\pi}{2}}=\frac{2}{\pi}\lim_{x\to 1}\frac{\sin^2\frac{\pi x}{2}}{2-x}=\frac{2}{\pi}.$$

Значит, $\lim\limits_{x\to 1}(2-x)^{\operatorname{tg}\frac{\pi x}{2}}=e^{\frac{2}{\pi}}$. ●

6

Определенный интеграл и его приложения

§ 36. Неопределенный интеграл и его свойства

1. Первообразная и неопределенный интеграл

В гл. 5 были рассмотрены примеры физических и геометрических задач, сводящихся к нахождению производных от функций. Наряду с этим имеются и задачи, сводящиеся к обратной операции, — отысканию функции по ее производной. Такая операция называется *интегрированием*, а раздел математики, изучающий способы нахождения функции по ее производной и приложения этой операции, — *интегральным исчислением*.

Одним из основных понятий интегрального исчисления является понятие первообразной.

Определение 1. Функцию $y = F(x)$, заданную на промежутке X, называют *первообразной* для функции $y = f(x)$, заданной на том же промежутке, если для всех $x \in X$ выполняется равенство $F'(x) = f(x)$.

Примеры. 1. Пусть $f(x) = x^3$. Тогда первообразная имеет вид $F(x) = \frac{x^4}{4}$, так как $F'(x) = \left(\frac{x^4}{4}\right)' = x^3 = f(x)$.

2. Пусть $f(x) = \sin 3x$. Тогда первообразная имеет вид $F(x) = -\frac{1}{3}\cos 3x$, поскольку $F'(x) = \left(-\frac{1}{3}\cos 3x\right)' = -\frac{1}{3}(-\sin 3x)\cdot 3 = \sin 3x = f(x)$.

Для функции $f(x) = x^3$ из примера 1 мы нашли первообразную $F(x) = \frac{x^4}{4}$. Это не единственное решение задачи. Так, в качестве первообразной можно взять и функцию $F_1(x) = \frac{x^4}{4} + 3$ (поскольку $\left(\frac{x^4}{4} + 3\right)' = x^3$), и функцию $F_2(x) = \frac{x^4}{4} - 5$ (поскольку $\left(\frac{x^4}{4} - 5\right)' = x^3$), и вообще любую функцию вида $\frac{x^4}{4} + C$. Точно так же обстоит дело в примере 2, где в качестве первообразной можно было взять любую функцию вида $-\frac{1}{3}\cos 3x + C$.

Справедлива следующая теорема.

Теорема 1 (о множестве всех первообразных). *Пусть для функции $f(x)$ на промежутке X существует первообразная $F(x)$. Тогда функция $f(x)$ имеет бесконечное множество первообразных, и все эти первообразные имеют вид $F(x)+C$.*

□ Пусть $F(x)$ — первообразная для $f(x)$. Тогда $(F(x)+C)'=F'(x)+C'=f(x)+0=f(x)$, а это означает, что $F(x)+C$ также является первообразной для $f(x)$.

Пусть теперь $F(x)$ и $\Phi(x)$ — две первообразные для функции $y=f(x)$, т. е. $F'(x)=\Phi'(x)=f(x)$. Тогда по теореме 2 § 34 разность этих функций $\Phi(x)-F(x)$ постоянна, т. е. $\Phi(x)-F(x)=C$, откуда получаем $\Phi(x)=F(x)+C$. ■

Определение 2. Если функция $y=f(x)$ имеет первообразную $F(x)$ на промежутке X, то множество всех первообразных называется *неопределенным интегралом* от функции $f(x)$.

Неопределенный интеграл обозначается $\int f(x)dx$ (читается «интеграл эф от икс дэ икс»). Таким образом,

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

Знак $\int$ называют *знаком интеграла*, функцию $f(x)$ — *подынтегральной функцией*, $f(x)dx$ — *подынтегральным выражением*, C — *постоянной интегрирования*.

Остановимся на двух свойствах неопределенного интеграла, легко вытекающих из определения.

Теорема 2. *Дифференциал от неопределенного интеграла равен подынтегральному выражению, а производная от неопределенного интеграла равна подынтегральной функции:*

$$d\left(\int f(x)dx\right)=f(x)dx\,;\quad \left(\int f(x)dx\right)'=f(x).$$

□ Так как $\int f(x)dx=F(x)+C$, где $F'(x)=f(x)$, то $\left(\int f(x)dx\right)'=(F(x)+C)'=F'(x)+C'=f(x)$. Но тогда

$$d\left(\int f(x)dx\right)=\left(\int f(x)dx\right)'dx=f(x)dx\,. \blacksquare$$

Теорема 3. *Неопределенный интеграл от производной некоторой функции равен этой функции, сложенной с произвольной постоянной:*

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

□ Так как $(F(x)+C)'=F'(x)$, то по определению неопределенного интеграла имеем $\int F'(x)dx=F(x)+C$. ■

Учитывая, что $F'(x)dx=dF(x)$, доказанное равенство можно записать в виде

$$\int dF(x)=F(x)+C.$$

2. Таблица основных интегралов

Из каждой формулы дифференцирования вытекает соответствующая формула интегрирования. Так, например, из того, что $\left(\frac{x^{\alpha+1}}{\alpha+1}\right)'=x^{\alpha}$ (где $\alpha\neq-1$), следует равенство $\int x^{\alpha}dx=\frac{x^{\alpha+1}}{\alpha+1}+C$. Приведем таблицу основных интегралов, которая получается из соответствующих формул дифференцирования элементарных функций:

I. $\int dx = x + C$.

II. $\int x^{\alpha} dx = \frac{x^{\alpha+1}}{\alpha+1} + C \ (\alpha \neq -1)$.

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

IV. $\int \frac{dx}{a^2 + x^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C$.

V. $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C$.

VI. $\int a^x dx = \frac{a^x}{\ln a} + C$; $\int e^x dx = e^x + C$.

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

VIII. $\int \cos x\, dx = \sin x + C$.

IX. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$.

X. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$.

XI. $\int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C$.

XII. $\int \frac{dx}{\sqrt{x^2 \pm a^2}} = \ln \left| x + \sqrt{x^2 \pm a^2} \right| + C$.

XIII. $\int \operatorname{sh} x\, dx = \operatorname{ch} x + C$.

XIV. $\int \operatorname{ch} x\, dx = \operatorname{sh} x + C$.

XV. $\int \frac{dx}{\operatorname{sh}^2 x} = -\operatorname{cth} x + C$.

XVI. $\int \frac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C$.

Установим, например, справедливость формулы III, т. е. докажем, что

$$\int \frac{1}{x} dx = \ln|x| + C.$$

В самом деле, если $x > 0$, то $|x| = x$ и, следовательно,

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

Если же $x < 0$, то $|x| = -x$ и, следовательно,

$$(\ln|x|)' = (\ln(-x))' = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

Итак, $(\ln|x|)' = \frac{1}{x}$, а, значит,

$$\int \frac{1}{x} dx = \ln|x| + C$$

Проверим еще, что верна формула XII. Для простоты ограничимся случаем, когда $x + \sqrt{x^2 \pm a^2} > 0$. Имеем

$$(\ln(x + \sqrt{x^2 \pm a^2})' = \frac{1}{x + \sqrt{x^2 \pm a^2}} \cdot \left(1 + \frac{x}{\sqrt{x^2 \pm a^2}}\right) =$$

$$= \frac{1}{x + \sqrt{x^2 \pm a^2}} \cdot \frac{x + \sqrt{x^2 \pm a^2}}{\sqrt{x^2 \pm a^2}} = \frac{1}{\sqrt{x^2 \pm a^2}}.$$

Итак, $(\ln(x + \sqrt{x^2 + a^2}))' = \frac{1}{\sqrt{x^2 \pm a^2}}$, а потому

$$\int \frac{dx}{\sqrt{x^2 \pm a^2}} = \ln(x + \sqrt{x^2 \pm a^2}) + C$$

Заметим, что переменную x, входящую в формулы I — XVI, можно заменить любой другой. Например, вместо формулы $\int \cos x\, dx = \sin x + C$ можно написать $\int \cos t\, dt = \sin t + C$ и т. д.

Пример. Найти интегралы:

а) $\int \frac{dx}{\sqrt[3]{x}}$; б) $\int \frac{dx}{x^2+16}$; в) $\frac{dx}{x^2-16}$; г) $\int \frac{dx}{\sqrt{3-x^2}}$; д) $\int \frac{dx}{\sqrt{x^2-3}}$.

○ а) Применяя формулу II, находим

$$\int \frac{dx}{\sqrt[3]{x}} = \int x^{-\frac{1}{3}}dx = \frac{x^{-\frac{1}{3}+1}}{-\frac{1}{3}+1} + C = \frac{3}{2}x^{\frac{2}{3}} + C.$$

б) Согласно формуле IV, получим

$$\int \frac{dx}{x^2+16} = \int \frac{dx}{x^2+4^2} = \frac{1}{4}\operatorname{arctg}\frac{x}{4} + C.$$

в) Воспользуемся формулой XI:

$$\int \frac{dx}{x^2-16} = \int \frac{dx}{x^2-4^2} = \frac{1}{8}\ln\left|\frac{x-4}{x+4}\right| + C.$$

г) Используя формулу V, находим

$$\int \frac{dx}{\sqrt{3-x^2}} = \int \frac{dx}{\sqrt{(\sqrt{3})^2 - x^2}} = \arcsin\frac{x}{\sqrt{3}} + C.$$

д) Воспользуемся формулой XII:

$$\int \frac{dx}{\sqrt{x^2-3}} = \ln|x + \sqrt{x^2-3}| + C. \bullet$$

3. Свойства неопределенного интеграла

Вычисление многих интегралов сводится к табличным, если использовать свойства неопределенных интегралов, вытекающие из соответствующих свойств дифференциалов.

1°. *Постоянный множитель можно вынести за знак неопределенного интеграла:*

$$\int \lambda f(x)dx = \lambda\int f(x)dx. \tag{1}$$

□ Продифференцировав правую часть равенства (1), получаем

$$d(\lambda\int f(x)dx) = \lambda d(\int f(x)dx) = \lambda f(x)dx.$$

Таким образом, дифференциал правой части доказываемой формулы равен подынтегральному выражению левой части, а это и означает справедливость формулы (1). ■

2°. *Если существуют интегралы $\int f_1(x)dx$ и $\int f_2(x)dx$, то неопределенный интеграл суммы $f_1(x) + f_2(x)$ равен сумме неопределенных интегралов от этих функций:*

$$\int (f_1(x) + f_2(x))dx = \int f_1(x)dx + \int f_2(x)dx. \tag{2}$$

□ Продифференцируем правую часть равенства (2):

$$d(\int f_1(x)dx + \int f_2(x)dx) = d(\int f_1(x)dx) + d(\int f_2(x)dx) = f_1(x)dx + f_2(x)dx = \\ = (f_1(x) + f_2(x))dx.$$

Мы получили подынтегральное выражение неопределенного интеграла, входящего в левую часть равенства (2), откуда и следует справедливость утверждения. ■

Примеры. 1. Найти $\int \frac{x\sqrt{x}+x^2-5}{x^2\sqrt{x}}\,dx.$

○ Разделив почленно числитель на знаменатель и использовав свойства 1° и 2°, получим линейную комбинацию табличных интегралов:

$$\int \frac{x\sqrt{x}+x^2-5}{x^2\sqrt{x}}\,dx = \int\left(\frac{1}{x}+\frac{1}{\sqrt{x}}-\frac{5}{x^2\sqrt{x}}\right)dx =$$

$$= \int\frac{dx}{x} + \int x^{-\frac{1}{2}}\,dx - 5\int x^{-\frac{5}{2}}\,dx = \ln|x| + \frac{x^{-\frac{1}{2}+1}}{-\frac{1}{2}+1} -$$

$$-5\frac{x^{-\frac{5}{2}+1}}{-\frac{5}{2}+1} + C = \ln|x| + 2\sqrt{x} + \frac{10}{3x\sqrt{x}} + C. \ ●$$

Замечание. Каждый из трех неопределенных интегралов содержит свою произвольную постоянную. В окончательном ответе через C обозначают их сумму.

2. Найти $\int \frac{1}{\sin^2 x\cos^2 x}\,dx.$

○ Записав стоящую в числителе единицу в тригонометрическом виде ($1=\sin^2 x + \cos^2 x$), разделим почленно числитель на знаменатель. Применив затем свойство 2°, получим

$$\int \frac{1}{\sin^2 x\cos^2 x}\,dx = \int\left(\frac{1}{\cos^2 x}+\frac{1}{\sin^2 x}\right)dx =$$

$$= \int\frac{dx}{\cos^2 x} + \int\frac{dx}{\sin^2 x} = \operatorname{tg} x - \operatorname{ctg} x + C. \ ●$$

§ 37. Методы интегрирования

1. Интегрирование по частям

Метод вычисления интегралов, называемый *методом интегрирования по частям*, основан на правиле дифференцирования произведения.

Пусть $u=u(x)$, $v=v(x)$ — функции, дифференцируемые на некотором промежутке X. Тогда, как известно, дифференциал произведения этих функций вычисляется по формуле

$$d(uv) = u\,dv + v\,du.$$

Взяв неопределенный интеграл от обеих частей этого равенства, получим

$$\int d(uv) = \int (u\,dv + v\,du).$$

Но $\int d(uv) = uv + C$, а $\int (u\,dv + v\,du) = \int u\,dv + \int v\,du$; поэтому $uv + C = \int u\,dv + \int v\,du$, откуда

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

Так как $\int v\,du$ уже содержит произвольную постоянную, то в правой части полученного равенства можно опустить C и записать равенство в виде

$$\int u\,dv = uv - \int v\,du. \qquad (1)$$

Эта формула называется *формулой интегрирования по частям*. Ею обычно пользуются в тех случаях, когда подынтегральное выражение $v\,du$ проще, чем подынтегральное выражение $u\,dv$.

Заметим, что одно и то же подынтегральное выражение можно записать в виде $u\,dv$ различными способами. Например,

$$x^4 \sin x\,dx = x^4 d(-\cos x) = \sin x\, d\left(\frac{x^5}{5}\right) = x^2 \sin x\, d\left(\frac{x^3}{3}\right)$$

и т. д. Поэтому иногда приходится пробовать различные формы такой записи, прежде чем метод приведет к успеху. Обычно стараются в подынтегральном выражении выделить части u и dv так, чтобы функция v была не сложнее, чем v', а u' проще, чем u. В частности, полезно иметь в виду, что для таких функций, как $\ln x$, x^n, $\operatorname{arctg} x$, $\operatorname{arcctg} x$ производные имеют более простой вид, нежели сами функции. Поэтому в большинстве случаев эти функции удобно принимать за u.

Примеры. 1. Найти $\int x^3 \ln x\,dx$.

○ Положим $u = \ln x$, $dv = x^3 dv$. Тогда $du = \frac{1}{x}dx$, $v = \frac{x^4}{4}$. Используя формулу (1), получим

$$\int x^3 \ln x\,dx = \frac{x^4}{4}\ln x - \int \frac{x^4}{4}\cdot\frac{1}{x}dx = \frac{x^4}{4}\ln x - \frac{1}{4}\int x^3 dx = \frac{x^4}{4}\ln x - \frac{1}{16}x^4 + C. \bullet$$

2. Найти $\int x^2 \sin x\,dx$.

○ Положим $u = x^2$, $dv = \sin x\,dx$. Тогда $du = 2x\,dx$, $v = -\cos x$. Применяя формулу (1), находим

$$\int x^2 \sin x\,dx = -x^2 \cos x + 2\int x\cos x\,dx. \tag{2}$$

Чтобы вычислить интеграл в правой части равенства (2) снова используем метод интегрирования по частям. Получим

$$\int x\cos x\,dx = x\sin x + \cos x + C_1. \tag{3}$$

Возвращаясь к исходному интегралу и воспользовавшись промежуточным равенством (3), окончательно находим

$$\int x^2 \sin x\,dx = -x^2\cos x + 2(x\sin x + \cos x + C_1) = 2x\sin x + \\ + (2 - x^2)\cos x + C \quad (\text{здесь } C = 2C_1) \bullet$$

2. Интегрирование методом замены переменной

Одним из наиболее мощных методов интегрирования является *метод замены переменной*. Поясним суть этого метода. Пусть $F'(x) = f(x)$. Тогда

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

Но в силу инвариантности формы дифференциала равенство $dF(x) = F'(x)dx = f(x)dx$ остается справедливым и в случае, когда x — промежуточный аргумент, т. е. $x = \varphi(t)$. Это означает, что формула $\int f(x)dx = F(x) + C$ верна при $x = \varphi(t)$. Таким образом,

$$\int f(\varphi(t))d\varphi(t) = F(\varphi(t)) + C \qquad \text{или} \qquad \int f(\varphi(t))\varphi'(t)dt = F(\varphi(t)) + C.$$

Тем самым доказана следующая теорема.

Теорема 1. *Если $F(x)$ является первообразной для $f(x)$ на промежутке X, а $x = \varphi(t)$ — дифференцируемая на промежутке T функция, значения которой принадлежат X, то $F(\varphi(t))$ — первообразная для $f(\varphi(t))\varphi'(t)$, $t \in T$ и, следовательно,*

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx. \tag{4}$$

Эта формула позволяет свести вычисление интеграла $\int f(\varphi(t))\varphi'(t)dt$ к вычислению интеграла $\int f(x)dx$. При этом вместо $\varphi(t)$ мы подставляем переменную x, а вместо $\varphi'(t)dt$ — дифференциал этой переменной, т. е. dx. Поэтому формула (4) называется *формулой замены переменной* под знаком неопределенного интеграла. На практике она применяется как «слева направо», так и «справа налево». Метод замены переменной позволяет сводить многие интегралы к табличным. После вычисления интеграла $\int f(x)dx$ нужно снова заменить x на $\varphi(t)$.

Примеры. 1. Найти $\int \frac{e^{\operatorname{arctg} t}}{1+t^2}dt$.

○ В данное подынтегральное выражение входит множитель $\frac{dt}{1+t^2}$, являющийся дифференциалом функции $\operatorname{arctg} t$. Полагая $\operatorname{arctg} t = x$, получим

$$\int \frac{e^{\operatorname{arctg} t}}{1+t^2}dt = \int e^x dx = e^x + C = e^{\operatorname{arctg} t} + C. \bullet$$

2. Найти $\int \frac{x^2dx}{x^6+9}$.

○ Числитель данного подынтегрального выражения напоминает дифференциал для x^3; в самом деле, $dx^3 = 3x^2dx$. Кроме того, знаменатель подынтегрального выражения легко выражается через x^3: $x^6+9 = (x^3)^2+9$. Это наводит на мысль о целесообразности подстановки $x^3 = u$. Тогда $du = 3x^2dx$, откуда $x^2dx = \frac{1}{3}du$. Таким образом,

$$\int \frac{x^2dx}{x^6+9} = \int \frac{\frac{1}{3}du}{u^2+9} = \frac{1}{3}\int \frac{du}{u^2+3^2} =$$
$$= \frac{1}{3}\cdot\frac{1}{3}\operatorname{arctg}\frac{u}{3} + C = \frac{1}{9}\operatorname{arctg}\frac{x^3}{3} + C. \bullet$$

Частным случаем теоремы 1 является следующая теорема, часто применяемая на практике.

Теорема 2. *Если* $\int f(x)dx = F(x) + C$, *то* $\int f(ax+b)dx = = \frac{1}{a}F(ax+b) + C$.

□ В самом деле, для нахождения интеграла $\int f(ax+b)dx$ положим $u = = ax+b$. Тогда $du = a\,dx$, и, значит,

$$\int f(ax+b)dx = \frac{1}{a}\int f(ax+b)a\,dx = \frac{1}{a}\int f(u)du =$$
$$= \frac{1}{a}F(u) + C = \frac{1}{a}F(ax+b) + C. \blacksquare$$

Пример 3. Найти $\int \frac{dx}{5x-3}$.

○ Так как $\int \frac{dx}{x} = \ln|x| + C$, то, используя теорему 2, получим

$$\int \frac{dx}{5x-3} = \frac{1}{5}\ln|5x-3| + C. \bullet$$

В рассмотренных примерах новая переменная представляла собой функцию переменной интегрирования. Наоборот, иногда бывает целесообразно, переменную интегрирования заменить функцией другой переменной. В частности, при интегрировании некоторых видов иррациональных функций оказываются удобными тригонометрические или гиперболические подстановки.

Примеры. 4. Найти $\int\sqrt{a^2-x^2}\,dx$.

○ Положим $x=a\sin t\;(-\pi/2\leqslant t\leqslant\pi/2,\ a>0)$ и выразим через новую переменную t все множители, входящие в подынтегральное выражение: $a^2-x^2=a^2-a^2\sin^2 t=a^2\cos^2 t$; $dx=a\cos t\,dt$. При этом $\cos t\geqslant 0$, так как $-\pi/2\leqslant t\leqslant\pi/2$.

Переходя к новой переменной под знаком неопределенного интеграла и учитывая, что $\cos t\geqslant 0$ и потому $\sqrt{\cos^2 t}=|\cos t|=\cos t$, получим

$$\int\sqrt{a^2-x^2}\,dx=\int\sqrt{a^2\cos^2 t}\,a\cos t\,dt=\int a\cos t\,a\cos t\,dt=$$
$$=a^2\int\cos^2 t\,dt=a^2\int\frac{1+\cos 2t}{2}\,dt=\frac{a^2}{2}\int dt+\frac{a^2}{2}\int\cos 2t\,dt=$$
$$=\frac{a^2}{2}t+\frac{a^2}{4}\sin 2t+C=\frac{a^2}{2}t+\frac{a^2}{4}2\sin t\cos t+C=$$
$$=\frac{a^2}{2}(t+\sin t\cos t)+C.$$

Так как $x=a\sin t$, то $\sin t=\frac{x}{a}$, откуда $t=\arcsin\frac{x}{a}$ (переход к обратной тригонометрической функции возможен, поскольку по условию $-\pi/2\leqslant t\leqslant\pi/2$). Далее, имеем

$$\cos t=\sqrt{1-\sin^2 t}=\sqrt{1-\frac{x^2}{a^2}}=\frac{\sqrt{a^2-x^2}}{a}$$

(перед радикалом берется знак плюс, так как $\cos t>0$). Значит,

$$\int\sqrt{a^2-x^2}\,dx=\int\frac{a^2}{2}\left(\arcsin\frac{x}{a}+\frac{x}{a}\cdot\frac{\sqrt{a^2-x^2}}{a}\right)+C=$$
$$=\frac{a^2}{2}\arcsin\frac{x}{a}+\frac{x}{2}\sqrt{a^2-x^2}+C.\ \bullet$$

5. Найти $\int\sqrt{a^2+x^2}\,dx$.

○ Положим $x=a\,\mathrm{sh}\,t$. В отличие от предыдущего примера здесь не надо накладывать ограничения на t, так как $\mathrm{sh}\,t$ — монотонная функция (см. § 26). Тогда получим $\sqrt{a^2+x^2}=\sqrt{a^2(1+\mathrm{sh}^2 t)}=\sqrt{a^2\mathrm{ch}^2 t}=a\,\mathrm{ch}\,t$ (можно не писать $|\mathrm{ch}\,t|$, поскольку $\mathrm{ch}\,t>0$ при всех t); $dx=a\,\mathrm{ch}\,t\,dt$. Следовательно,

$$\int\sqrt{a^2+x^2}\,dx=\int a\,\mathrm{ch}\,t\cdot a\,\mathrm{ch}\,t\,dt=\frac{a^2}{2}\int(1+\mathrm{ch}\,2t)\,dt=$$
$$=\frac{a^2}{2}\left(t+\frac{1}{2}\,\mathrm{sh}\,2t\right)+C=\frac{a^2}{2}(t+\mathrm{sh}\,t\cdot\mathrm{ch}\,t)+C.$$

Остается выразить t, $\mathrm{sh}\,t$, $\mathrm{ch}\,t$ через x. Находим $\mathrm{sh}\,t=\frac{x}{a}$; $\mathrm{ch}\,t=\sqrt{1+\mathrm{sh}^2 t}=\sqrt{1+\frac{x^2}{a^2}}=\frac{\sqrt{a^2+x^2}}{a}$. Для отыскания t воспользуемся определением функции $\mathrm{sh}\,t$; имеем $\mathrm{sh}\,t=\frac{x}{a}$, т. е. $\frac{e^t-e^{-t}}{2}=\frac{x}{a}$, и, далее, $e^t-\frac{1}{e^t}=\frac{2x}{a}$. Полагая $e^t=y$, получим $y^2-\frac{2x}{a}y-1=0$, откуда

$$y=\frac{x}{a}+\sqrt{\frac{x^2}{a^2}+1};\quad e^t=\frac{1}{a}\left(x+\sqrt{a^2+x^2}\right);$$
$$t=\ln\left(x+\sqrt{a^2+x^2}\right)-\ln a.$$

Итак,

$$\int\sqrt{a^2+x^2}dx=\frac{a^2}{2}t+\frac{a^2}{2}\operatorname{sh}t\operatorname{ch}t+C=$$

$$=\frac{a^2}{2}(\ln(x+\sqrt{a^2+x^2})-\ln a)+\frac{a^2}{2}\cdot\frac{x}{a}\cdot\frac{\sqrt{a^2+x^2}}{a}+C=$$

$$=\frac{a^2}{2}\ln(x+\sqrt{a^2+x^2})+\frac{1}{2}x\sqrt{a^2+x^2}+C$$

(слагаемое $-\frac{a^2}{2}\ln a$ мы включили в постоянную C). ●

§ 38. Интегрирование некоторых классов функций

В предыдущем параграфе речь шла об общих приемах интегрирования. В этом параграфе мы рассмотрим интегрирование конкретных классов функций с помощью изложенных приемов.

1. Интегрирование рациональных функций

Рассмотрим интеграл вида $\int R(x)dx$, где $y=R(x)$ — рациональная функция. Всякое рациональное выражение $R(x)$ можно представить в виде $P(x)/Q(x)$, где $P(x)$ и $Q(x)$ — многочлены. Если эта дробь неправильная, т. е. если степень числителя больше или равна степени знаменателя, то ее можно представить в виде суммы многочлена (целой части) и правильной дроби; правильную же дробь можно представить в виде суммы *простейших дробей* вида

$$\frac{A}{x-a};\ \frac{A}{(x-a)^n};\ \frac{Ax+B}{x^2+px+q};\ \frac{Ax+B}{(x^2+px+q)^n},$$

где A, B, a, p, q — действительные числа, а квадратный трехчлен x^2+px+q не имеет действительных корней.

Примеры. 1. Найти $I=\int\frac{x^3-2x^2-3x+4}{x^4-4x^3+4x^2}dx.$

○ Прежде всего разложим знаменатель на множители:

$$x^4-4x^3+4x^2=x^2(x^2-4x+4)=x^2(x-2)^2.$$

Заданную рациональную дробь можно представить в виде суммы простейших дробей следующим образом:

$$\frac{x^3-2x^2-3x+4}{x^2(x-2)^2}=\frac{A}{x}+\frac{B}{x^2}+\frac{C}{x-2}+\frac{D}{(x-2)^2}.$$

Умножим обе части этого тождества на знаменатель левой части:

$$x^3-2x^2-3x+4=Ax(x-2)^2+B(x-2)^2+Cx^2(x-2)+Dx^2. \qquad (1)$$

При $x=0$ получаем $4=4B$, т. е. $B=1$; при $x=2$ получаем $-2=4D$, т. е. $D=-1/2$.

К сожалению, в данном примере у нас нет больше таких значений x, с помощью которых сразу вычисляются значения неопределенных коэффициентов A и C. Для отыскания этих коэффициентов воспользуемся известной из курса алгебры теоремой о тождественности двух многочленов, согласно которой два многочлена от одной переменной тождественны тогда и только тогда, когда они имеют одинаковую степень и коэффициенты при одноименных степенях переменной в обоих многочленах совпадают.

Раскрывая скобки в правой части тождества (1) и приводя подобные члены, получим

$$x^3-2x^2-3x+4=(A+C)x^3+(-4A+B-2C+D)x^2+(4A-4B)x+4B.$$

Из сравнения членов, содержащих x в первой степени, получаем $-3 = 4A - 4B$. Но $B = 1$ и, следовательно, $4A = 1$, откуда $A = 1/4$. Из сравнения членов, содержащих x^3, имеем $1 = A + C$. Так как $A = 1/4$, то $C = 3/4$. Итак,

$$\frac{x^3-2x^2-3x+4}{x^2(x-2)^2} = \frac{1/4}{x} + \frac{1}{x^2} + \frac{3/4}{x-2} - \frac{1/2}{(x-2)^2}.$$

Теперь вернемся к нахождению интеграла. Имеем

$$I = \int \frac{x^3-2x^2-3x+4}{x^4-4x^3+4x^2}\,dx = \frac{1}{4}\int\frac{dx}{x} + \int\frac{dx}{x^2} +$$

$$+ \frac{3}{4}\int\frac{dx}{x-2} - \frac{1}{2}\int\frac{dx}{(x-2)^2} = \frac{1}{4}\ln|x| - \frac{1}{x} + \frac{3}{4}\ln|x-2| + \frac{1}{2(x-2)} + C. \bullet$$

2. Найти $I = \int \frac{2x^3-13x^2+29x-20}{x^2-6x+11}\,dx$.

○ Прежде всего выделим целую часть из данной неправильной дроби:

$$\begin{array}{r|l} 2x^3-13x^2+29x-20 & x^2-6x+11 \\ \cline{2-2} -\;2x^3-12x^2+22x & 2x-1 \\ \cline{1-1} -x^2+7x-20 & \\ -\;-x^2+6x-11 & \\ \cline{1-1} x-9 & \end{array}$$

Значит,

$$\frac{2x^3-13x^3+29x-20}{x^2-6x+11} = 2x - 1 + \frac{x-9}{x^2-6x+11},$$

а потому

$$I = \int\frac{2x^3-13x^2+29x-20}{x^3-6x+11}\,dx = \int 2x\,dx - \int dx +$$

$$+ \int\frac{x-9}{x^2-6x+11}\,dx = x^2 - x + \int\frac{(x-9)dx}{x^2-6x+11}.$$

Квадратный трехчлен $x^2-6x+11$ не имеет действительных корней. В подобных случаях часто оказывается полезным метод выделения полного квадрата. Имеем

$$\int\frac{(x-9)dx}{x^2-6x+11} = \int\frac{(x-3)-6}{(x-3)^2+2}\,dx.$$

Положим $t = x - 3$; тогда $dt = dx$. Следовательно,

$$\int\frac{(x-3)-6}{(x-3)^2+2}\,dx = \int\frac{t-6}{t^2+2}\,dx = \frac{1}{2}\int\frac{2t\,dt}{t^2+2} - 6\int\frac{dt}{t^2+2}.$$

Полагая в первом из полученных интегралов $u = t^2 + 2$, находим $du = 2t\,dt$ и, значит,

$$\int\frac{2t\,dt}{t^2+2} = \int\frac{du}{u} = \ln|u| = \ln|t^2+2| = \ln(x^2-6x+11).$$

Второй интеграл $\int\frac{dt}{t^2+2}$ является табличным (см. формулу IV из п. 2 § 36) и равен $\frac{1}{\sqrt{2}}\operatorname{arctg}\frac{t}{\sqrt{2}}$, т. е. $\frac{1}{\sqrt{2}}\operatorname{arctg}\frac{x-3}{\sqrt{2}}$.

Окончательно, получаем

$$I = x^2 - x + \frac{1}{2}\ln(x^2-6x+11) - \frac{6}{\sqrt{2}}\operatorname{arctg}\frac{x-3}{\sqrt{2}} + C. \bullet$$

2. Интегрирование тригонометрических функций

При вычислении интегралов от тригонометрических функций для преобразования подынтегральных выражений часто используются различные формулы тригонометрии, в первую очередь — формулы преобразования произведений в сумму и формулы понижения степени.

Пример 1. Вычислить $\int \sin^2 x \cos^4 x\,dx$.

○Имеем

$$\int \sin^2 x \cos^4 x\,dx = \int (\sin^2 x \cos^2 x)\cos^2 x\,dx = \frac{1}{4}\int \sin^2 2x \cos^2 x\,dx =$$

$$= \frac{1}{4}\int \frac{1-\cos 4x}{2}\cdot\frac{1+\cos 2x}{2}\,dx = \frac{1}{16}\int (1 - \cos 4x + \cos 2x -$$

$$- \cos 4x \cos 2x)dx = \frac{1}{16}\left(x - \frac{\sin 4x}{4} + \frac{\sin 2x}{2} - \frac{1}{2}\int(\cos 6x + \cos 2x)dx\right) =$$

$$= \frac{1}{16}\left(x - \frac{\sin 4x}{4} + \frac{\sin 2x}{2} - \frac{1}{2}\left(\frac{\sin 6x}{6} + \frac{\sin 2x}{2}\right)\right) + C =$$

$$= \frac{1}{16}\left(x + \frac{1}{4}\sin 2x - \frac{1}{4}\sin 4x - \frac{1}{12}\sin 6x\right) + C. \bullet$$

Рассмотрим интегралы вида $\int R(\sin x, \cos x)dx$, где $R(u, v)$ — рациональная функция. Такие интегралы всегда рационализируются подстановкой $t = \operatorname{tg}(x/2)$, $(-\pi < x < \pi)$. В самом деле,

$$\sin x = \frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{2t}{1+t^2},$$

$$\cos x = \frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{1-t^2}{1+t^2}.$$

Далее выразим переменную x через переменную t. Так как $\operatorname{tg}(x/2) = t$, $-\pi/2 < x/2 < \pi/2$, то $x = 2\operatorname{arctg} t$, а потому $dx = \frac{2dt}{1+t^2}$. Значит,

$$\int R(\sin x, \cos x)dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}.$$

Таким образом, задача свелась к вычислению интеграла от рациональной функции. Поскольку подстановка $t = \operatorname{tg}(x/2)$ позволяет рационализировать любой интеграл вида $\int R(\sin x, \cos x)dx$, ее называют *универсальной подстановкой*. Любой интеграл этого вида выражается через элементарные функции.

Пример 2. Вычислить $\int \frac{dx}{2\sin x + 3\cos x + 4}$.

○Воспользуемся универсальной подстановкой $t = \operatorname{tg}(x/2)$, $-\pi < x < \pi$. Имеем

$$2\sin x + 3\cos x + 4 = 2\cdot\frac{2t}{1+t^2} + 3\frac{1-t^2}{1+t^2} + 4 = \frac{t^2+4t+7}{1-t^2},\quad dx = \frac{2dt}{1+t^2}.$$

Заменив переменную под знаком интеграла, получим

$$\int \frac{dx}{2\sin x+3\cos x+4} = \int \frac{\frac{2dt}{1+t^2}}{\frac{t^2+4t+7}{1+t^2}} =$$

$$= 2\int \frac{dt}{t^2+4t+7} = 2\int \frac{dt}{(t+2)^2+3} = \frac{2}{\sqrt{3}}\operatorname{arctg}\frac{t+2}{\sqrt{3}} + C =$$

$$= \frac{2}{\sqrt{3}}\operatorname{arctg}\frac{\operatorname{tg}\frac{x}{2}+2}{\sqrt{3}} + C. \bullet$$

Хотя подстановка $t = \operatorname{tg}(x/2)$ универсальна, она часто приводит к слишком громоздким выкладкам. Во многих случаях удается упростить вычисление интегралов вида $\int R(\sin x, \cos x)dx$, используя другие подстановки.

Примеры. 3. Найти $\int \frac{\cos^3 x\,dx}{1+\sin^2 x}$.

○ Имеем

$$\int \frac{\cos^3 x\,dx}{1+\sin^2 x} = \int \frac{\cos^2 x \cos x\,dx}{1+\sin^2 x} = \int \frac{(1-\sin^2 x)\cos x\,dx}{1+\sin^2 x}.$$

Положим $t = \sin x$; тогда $dt = \cos x\,dx$ и мы получим $\int \frac{1-t^2}{1+t^2}dt$, т.е. интеграл от рациональной функции. Далее, находим

$$\int \frac{1-t^2}{1+t^2}dt = \int \left(\frac{2}{1+t^2} - 1\right)dt = 2\operatorname{arctg} t - t + C = 2\operatorname{arctg}(\sin x) - \sin x + C. \bullet$$

4. Найти $\int \frac{\cos^2 x}{\sin^4 x}dx$.

○ Имеем

$$\int \frac{\cos^2 x}{\sin^4 x}dx = \int \frac{\cos^2 x}{\sin^2 x}\,\frac{dx}{\sin^2 x} = \int \operatorname{ctg}^2 x\,\frac{dx}{\sin^2 x}.$$

Положим $t = \operatorname{ctg} x$; тогда $dt = -\frac{dx}{\sin^2 x}$ и мы получим $-\int \operatorname{ctg}^2 x\left(-\frac{dx}{\sin^2 x}\right) = -\int t^2 dt =$ $= -\frac{t^3}{3} + C = -\frac{1}{3}\operatorname{ctg}^3 x + C. \bullet$

§ 39. Определенный интеграл

1. Задача о вычислении площади криволинейной трапеции

Криволинейной трапецией называют фигуру в плоскости xy, ограниченную прямыми $x = a$, $x = b$ $(a < b)$ и графиками функций $y = f_1(x)$, $y = f_2(x)$, непрерывных на $[a, b]$ и таких, что $f_1(x) \leqslant f_2(x)$ для всех $x \in [a, b]$ (рис. 130). Рассмотрим частный случай такой трапеции, ограниченной прямыми $x = a$, $x = b$, $y = 0$ и графиком непрерывной и неотрицательной на $[a, b]$ функции $y = f(x)$ (рис. 131). Как найти площадь такой фигуры? Правда, само понятие площади также нуждается в определении, но к этому мы вернемся позднее. Пока же будем опираться на интуитивное представление о площади.

Разобьем отрезок $[a, b]$ на ряд мелких участков точками $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$, на каждом участке $[x_k, x_{k+1}]$ найдем наименьшее значение функции $y = f(x)$, обозначим его m_k и рассмотрим прямоугольник с основанием $[x_k, x_{k+1}]$ и высотой m_k (рис. 132); его площадь равна $m_k(x_{k+1} - x_k) = m_k\Delta x_k$.

Объединение этих прямоугольников представляет собой вписанную в данную криволинейную трапецию ступенчатую фигуру (рис. 133); ее площадь обозначим $\underline{S}_T$ (буква T символизирует то разбиение отрезка $[a, b]$, которое мы осуществили). Аналогично, если на каждом участке $[x_k, x_{k+1}]$ выбрать наибольшее значение функции M_k и рассмотреть прямоугольник с высотой M_k, то объединение таких прямоугольников даст описанную около данной криволинейной трапеции ступенчатую фигуру (рис. 134); ее площадь обозначим $\bar{S}_T$.

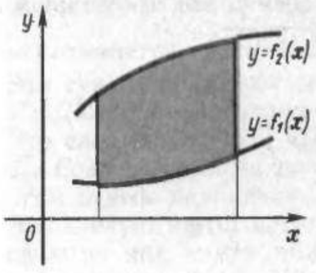


Рис. 130

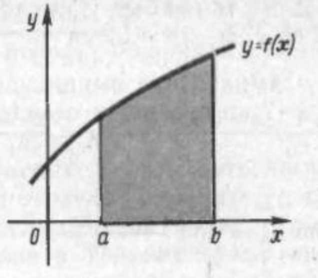


Рис. 131

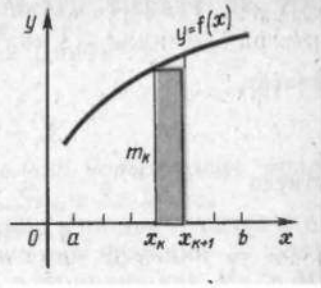


Рис. 132

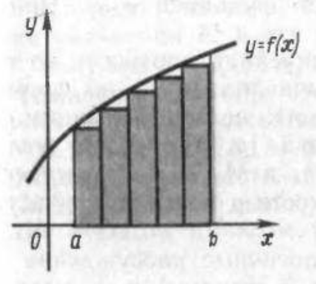


Рис. 133

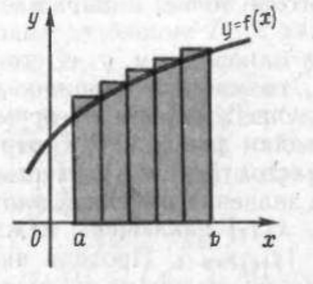


Рис. 134

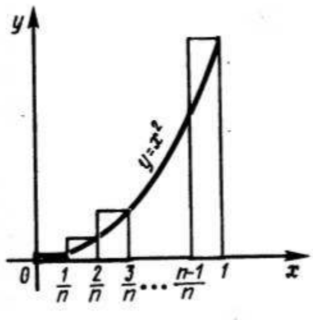


Рис. 135

Очевидно, что для любого разбиения T выполняется неравенство $\underline{S}_T \leqslant S \leqslant \bar{S}_T$, где S — искомая площадь криволинейной трапеции. Эту площадь можно определить как число, которое не меньше площади любой вписанной ступенчатой фигуры и не больше площади любой описанной ступенчатой фигуры, а точнее как число, разделяющее множества $\{\underline{S}_T\}$ и $\{\bar{S}_T\}$ для всевозможных разбиений T отрезка $[a, b]$. Интуитивно ясно, что такое разделяющее число должно быть единственным.

Искомая площадь S приближенно равна площади вписанной или описанной ступенчатой фигуры, т. е. $S \approx \underline{S}_T$ или $S \approx \bar{S}_T$.

На практике делят отрезок $[a, b]$ на n равных частей и вместо $\underline{S}_T$ используют запись $\underline{S}_n$, а вместо $\bar{S}_T$ — запись $\bar{S}_n$. Чем больше n, тем точнее приближенное равенство $\underline{S}_n \approx S$ или $\bar{S}_n \approx S$. Точное равенство получается при переходе к пределу: $S = \lim\limits_{n \to \infty} \underline{S}_n$ или $S = \lim\limits_{n \to \infty} \bar{S}_n$.

Пример. Найти площадь криволинейной трапеции, ограниченной параболой $y = x^2$ и прямыми $x = 0$, $x = 1$, $y = 0$ (рис. 135).

○ Разделим отрезок $[0, 1]$ на n равных частей точками $x_0 = 0$, $x_1 = \frac{1}{n}$, $x_2 = \frac{2}{n}$, ..., $x_{n-1} = \frac{n-1}{n}$, $x_n = \frac{n}{n} = 1$. Тогда $f(x_0) = 0$, $f(x_1) = \frac{1}{n^2}$, $f(x_2) = \frac{2^2}{n^2}$, ..., $f(x_{n-1}) = \frac{(n-1)^2}{n^2}$, $f(x_n) = \frac{n^2}{n^2} = 1$. Составим сумму $\bar{S}_n$ (площадь ступенчатой фигуры на рис. 135):

$$\bar{S}_n = \frac{1}{n}\left(\frac{1}{n^2} + \frac{2^2}{n^2} + \frac{3^2}{n^2} + \ldots + \frac{n^2}{n^2}\right) = \frac{1}{n^3}(1^2 + 2^2 + 3^2 + \ldots + n^2).$$

Ранее (см. пример 2 § 11) мы доказали, что

$$1^2+2^2+3^2+\dots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

Значит,

$$\bar{S}_n=\frac{1}{n^3}\cdot\frac{n(n+1)(2n+1)}{6}=\frac{2n^2+3n+1}{6n^2},$$

откуда

$$S=\lim_{n\to\infty}\bar{S}_n=\lim_{n\to\infty}\frac{2n^2+3n+1}{6n^2}=\frac{2}{6}=\frac{1}{3}. \bullet$$

2. Задача о массе материальной плоской пластины

Пусть дан прямолинейный неоднородный материальный стержень $[a, b]$, линейная плотность которого в точке x выражается функцией $\rho(x)$. Найдем массу μ стержня.

Если бы стержень был однородным, т. е. его линейная плотность во всех точках была бы равна ρ, то масса μ стержня вычислялась бы по формуле $\mu=\rho(b-a)$. В данном случае эту формулу применить нельзя. Поступим следующим образом: произведем разбиение T отрезка $[a, b]$ на ряд мелких участков и рассмотрим участок $[x_k, x_{k+1}]$. Пусть m_k и M_k — соответственно наименьшее и наибольшее значения линейной плотности $\rho(x)$ на этом участке. Тогда масса участка $[x_k, x_{k+1}]$ заключена между числами $m_k\Delta x_k$ и $M_k\Delta x_k$, где Δx_k — длина отрезка $[x_k, x_{k+1}]$. Проведя аналогичные рассуждения для остальных участков разбиения, получим, что масса μ стержня $[a, b]$ удовлетворяет двойному неравенству $\underline{S}_T\leqslant\mu\leqslant\bar{S}_T$, где $\underline{S}_T=m_0\Delta x_0+m_1\Delta x_1+$ $+m_2\Delta x_2+\dots+m_{n-1}\Delta x_{n-1}$ (короче это записывают так: $\sum\limits_{k=0}^{n-1}m_k\Delta x_k$, $\sum$ — «сигма»). Аналогично получим $\bar{S}_T=M_0\Delta x_0+M_1\Delta x_1+\dots+M_{n-1}\Delta x_{n-1}$, или, короче $\sum\limits_{k=0}^{n-1}M_k\Delta x_k$.

Таким образом, масса стержня есть число, разделяющее множества $\{\underline{S}_T\}$ и $\{\bar{S}_T\}$ для всевозможных разбиений T отрезка $[a, b]$.

3. Определение определенного интеграла

Две различные задачи, рассмотренные в предыдущих пунктах, в процессе решения привели к одной и той же математической модели — к двум определенным образом построенным числовым множествам $\{\underline{S}_T\}$ и $\{\bar{S}_T\}$, разделяющимся единственным числом: в первом случае это число определяет площадь криволинейной трапеции, во втором — массу стержня. Оказывается, многие задачи приводят к такой же математической модели, поэтому есть смысл специально заняться ее изучением. Прежде всего нужно более точно осмыслить процесс решения двух рассмотренных выше задач, отвлекаясь от их конкретного содержания.

Итак, пусть на отрезке $[a, b]$ определена ограниченная функция $y=f(x)$. Произведем разбиение T отрезка $[a, b]$ точками $a=x_0<x_1<x_2<\dots<$ $<x_{n-1}<x_n=b$, на каждом из отрезков разбиения $[x_k, x_{k+1}]$ найдем нижнюю и верхнюю грани значений функции, обозначим их соответственно m_k и

M_k и составим две суммы $\underline{S}_T = \sum_{k=0}^{n-1} m_k \Delta x_k$, $\bar{S}_T = \sum_{k=0}^{n-1} M_k \Delta x_k$. Первая из этих сумм называется *нижней*, а вторая — верхней суммой Дарбу*.

Эти суммы обладают следующими свойствами:

1^0. *Для любого T справедливо неравенство $\underline{S}_T \leqslant \bar{S}_T$.*

Это следует из того, что $m_k \leqslant M_k$.

2^0. *Если к данному разбиению T_1 добавить несколько новых точек, получив тем самым разбиение T_2 отрезка $[a, b]$, то $\underline{S}_{T_1} \leqslant \underline{S}_{T_2}$, а $\bar{S}_{T_1} \geqslant \bar{S}_{T_2}$.*

Это следует из того, что если отрезок $[x_k, x_{k+1}]$ разбить на два отрезка и на каждом из них найти нижние и верхние грани значений функции — соответственно m_k', m_k'', M_k', M_k'', — то $m_k \leqslant m_k'$ и $m_k \leqslant m_k''$, в то время как $M_k \geqslant M_k'$, $M_k \geqslant M_k''$.

3^0. *Для любых разбиений T_1 и T_2 отрезка $[a, b]$ выполняется неравенство $\underline{S}_{T_1} \leqslant \bar{S}_{T_2}$.*

Это следует из того, что, составив разбиение T, включающее в себя все точки разбиения T_1 и все точки разбиения T_2, а затем используя свойства 1^0 и 2^0, получим $\underline{S}_{T_1} \leqslant \underline{S}_T \leqslant \bar{S}_T \leqslant \bar{S}_{T_2}$.

Последнее свойство означает, что множество X нижних сумм Дарбу расположено левее множества Y верхних сумм Дарбу, построенных для ограниченной на отрезке $[a, b]$ функции $f(x)$. Тогда найдется хотя бы одно число I, разделяющее множества X и Y, т. е. такое, что для любого разбиения отрезка $[a, b]$ выполняется двойное неравенство

$$\underline{S}_T = \sum_{k=0}^{n-1} m_k \Delta x_k \leqslant I \leqslant \sum_{k=0}^{n-1} M_k \Delta x_k = \bar{S}_T.$$

Определение. Функция $y = f(x)$, ограниченная на отрезке $[a, b]$, называется *интегрируемой* на этом отрезке, если существует единственное число I, разделяющее множества нижних и верхних сумм Дарбу, образованных для всевозможных разбиений отрезка $[a, b]$. Если функция интегрируема на отрезке $[a, b]$, то единственное число, разделяющее эти множества, называют *определенным интегралом* этой функции по отрезку $[a, b]$ и обозначают символом $\int_a^b f(x)\,dx$.

Знак $\int_a^b$ читается: «интеграл от a до b»; числа a и b называются соответственно *нижним* и *верхним пределами интегрирования*. Позднее мы установим связь между $\int f(x)dx$ и $\int_a^b f(x)dx$, которая сделает оправданным использование знака интеграла и в случае определенного интеграла.

Мы определили интеграл $\int_a^b f(x)dx$ для случая, когда $a < b$. Если $a > b$, то положим $\int_a^b f(x)dx = -\int_b^a f(x)dx$. Это определение естественно, так как при изменении направления промежутка интегрирования каждая разность $\Delta x_k = x_{k+1} - x_k$ изменяет знак, а тогда изменят знаки и суммы Дарбу и, следовательно, разделяющее их число, т. е. интеграл.

* Г. Дарбу (1842—1917) — французский математик.

Так как при $a=b$ все Δx_k обращаются в нуль, то положим

$$\int\limits_a^a f(x)dx = 0.$$

Рассматривая в п. 1 задачу о площади криволинейной трапеции, мы получили, что площадь есть число, разделяющее площади вписанных и описанных ступенчатых фигур, а эти площади являются нижними и верхними суммами Дарбу для заданной неотрицательной и непрерывной на отрезке $[a, b]$ функции $y=f(x)$. Опираясь на интуицию, мы предположили, что это число единственно. Значит, $S=\int\limits_a^b f(x)dx$, т. е. *определенный интеграл выражает площадь криволинейной трапеции, ограниченной прямыми $x=a$, $x=b$ $(a<b)$, $y=0$ и графиком непрерывной и неотрицательной на отрезке $[a, b]$ функции $y=f(x)$*. В этом состоит геометрический смысл определенного интеграла.

Рассматривая в п. 2 задачу о массе стержня, мы получили, что масса есть число, разделяющее множества нижних и верхних сумм Дарбу для функции $\rho(x)$, задающей плотность стержня. По смыслу задачи это число единственно. Значит, $\mu=\int\limits_a^b \rho(x)dx$, т. е. *масса стержня есть интеграл от плотности*. В этом состоит физический смысл определенного интеграла.

Приведем пример, показывающий, что существуют неинтегрируемые функции. Напомним, что функцией Дирихле называют функцию $D(x)$, определяемую на отрезке $[0, 1]$ равенствами

$$D(x)=\begin{cases} 0, & \text{если } x \text{ — иррациональное число;} \\ 1, & \text{если } x \text{ — рациональное число.} \end{cases}$$

Какой бы отрезок $[x_k, x_{k+1}]$ мы ни взяли, на нем найдутся как рациональные, так и иррациональные точки, т. е. точки, где $D(x)=0$, и точки, где $D(x)=1$. Поэтому для любого разбиения отрезка $[0, 1]$ все значения m_k равны нулю, а все значения M_k равны единице. Тогда все нижние суммы Дарбу $\sum\limits_{k=0}^{n-1} m_k\Delta x_k$ равны нулю, а все верхние суммы Дарбу $\sum\limits_{k=0}^{n-1} M_k\Delta x_k$ равны единице, поскольку $\sum\limits_{k=0}^{n-1} m_k\Delta x_k = \sum\limits_{k=0}^{n-1} 0\cdot\Delta x_k = 0$, $\sum\limits_{k=0}^{n-1} M_k\Delta x_k = \sum\limits_{k=0}^{n-1} 1\cdot\Delta x_k$, а $\sum\limits_{k=0}^{n-1} \Delta x_k$ — длина отрезка $[0, 1]$. Итак, в рассматриваемом случае $X=\{0\}$, $Y=\{1\}$ и любое число из промежутка $[0, 1]$ разделяет множества X и Y. Значит, функция Дирихле не является интегрируемой на отрезке $[0, 1]$.

Теорема 1 (необходимое и достаточное условие интегрируемости функции). *Для того чтобы функция $y=f(x)$, определенная и ограниченная на отрезке, была интегрируема на этом отрезке, необходимо и достаточно, чтобы для любого $\varepsilon>0$ существовало такое разбиение T, что $\bar{S}_T-\underline{S}_T<\varepsilon$.* Короче:

$$(\forall\, \varepsilon>0)\,(\exists\, T)\, \bar{S}_T-\underline{S}_T<\varepsilon.$$

Это следует из критерия единственности разделяющего числа (см. § 10, п. 3) и свойств 1^0 и 2^0 сумм Дарбу.

Поскольку

$$\bar{S}_T - \underline{S}_T = \sum_{k=0}^{n-1} M_k \Delta x_k - \sum_{k=0}^{n-1} m_k \Delta x_k = \sum_{k=0}^{n-1} (M_k - m_k) \Delta x_k,$$

условие $\bar{S}_T - \underline{S}_T < \varepsilon$ можно записать и так:

$$\sum_{k=0}^{n-1} (M_k - m_k) \Delta x_k < \varepsilon. \tag{1}$$

Разность $M_k - m_k$ будем обозначать через ω_k и называть *колебанием* функции $f(x)$ на отрезке $[x_k, x_{k+1}]$. Тогда неравенство (1) можно записать следующим образом:

$$\sum_{k=0}^{n-1} \omega_k \Delta x_k < \varepsilon.$$

4. Интегрируемость непрерывной функции

В предыдущем пункте мы ввели понятие интегрируемой функции и установили необходимое и достаточное условие интегрируемости. Теперь докажем, что всякая непрерывная на отрезке $[a, b]$ функция интегрируема на нем, т. е. для непрерывной на $[a, b]$ функции $y = f(x)$ всегда существует интеграл $\int\limits_a^b f(x)dx$.

Для доказательства этого утверждения понадобится следующая лемма.

Лемма. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то для любого $\varepsilon > 0$ найдется хотя бы одно разбиение T этого отрезка такое, что все ω_k меньше ε, т. е. $(\forall k)\ \omega_k \leqslant \varepsilon$.*

□ Проведем доказательство методом от противного. Предположим, что для какого-то $\varepsilon > 0$ разбиение, при котором любое ω_k меньше ε, невозможно.

Рис. 136

Рис. 137

Это значит, что для любого разбиения T отрезка $[a, b]$ найдется такой номер k, что $M_k - m_k \geqslant \varepsilon$.

Разделим отрезок $[a, b]$ пополам. Тогда для выбранного $\varepsilon > 0$ хотя бы одну из половин отрезка $[a, b]$ нельзя разбить требуемым образом (так как если бы обе его половины можно было разбить этим образом, то и весь отрезок $[a, b]$ был бы разбит тем же образом). Выбрав ту из его половин, которую нельзя разбить требуемым образом, например $[a_1, b_1]$, снова разобьем ее пополам (рис. 136). Тогда одна из вновь полученных половин, например $[a_2, b_2]$, в соответствии со сделанным предположением, не может быть разбита указанным образом.

Продолжая дальше процесс разбиения отрезков, получим стягивающуюся последовательность отрезков такую, что ни один из отрезков этой системы не может быть для данного $\varepsilon > 0$ разбит требуемым образом. Согласно принципу стягивающихся отрезков, существует одна и только одна точка c, принадлежащая всем этим отрезкам. По условию, в точке c функция $f(x)$ непрерывна. Значит, найдется такая окрестность $(c - \delta, c + \delta)$ этой точки, в которой выполняется неравенство $|f(x) - f(c)| < \varepsilon/2$. При достаточно большом значении k отрезок $[a_k, b_k]$ целиком лежит внутри δ-окрестности точки c (рис. 137).

Пусть M_k и m_k — соответственно наибольшее и наименьшее значения функции $f(x)$ на отрезке $[a_k, b_k]$. Так как точки, в которых функция $f(x)$ принимает значения M_k и m_k, принадлежит и промежутку $(c-\delta, c+\delta)$, то имеют место неравенства $|M_k - f(c)| < \varepsilon/2$ и $|f(c) - m_k| < \varepsilon/2$. Поэтому

$$\omega_k = M_k - m_k = M_k - f(c) + f(c) - m_k = |M_k - f(c)| + |f(c) - m_k| < \\ < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Таким образом, вопреки сделанному предположению, мы получили отрезок $[a_k, b_k]$, на котором выполнено неравенство $\omega_k < \varepsilon$. Значит, это предположение было неверным, и, следовательно, справедливо утверждение леммы. ■

Теорема 2. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке.*

□ Возьмем произвольное $\varepsilon > 0$. В силу леммы существует такое разбиение T, что для всех k имеем $\omega_k < \varepsilon/(b-a)$. Тогда

$$\sum_{k=0}^{n-1} \omega_k \Delta x_k < \frac{\varepsilon}{b-a} \sum_{k=0}^{n-1} \Delta x_k = \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

Но неравенство $\sum\limits_{k=0}^{n-1} \omega_k \Delta x_k < \varepsilon$ и означает, что функция интегрируема на отрезке $[a, b]$ (см. п. 3). ■

5. Аддитивность определенного интеграла

Теорема 3. *Если функция $y = f(x)$ интегрируема на отрезках $[a, c]$ и $[c, b]$, $a < c < b$, то она интегрируема и на отрезке $[a, b]$, причем выполняется равенство*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \qquad (2)$$

(аддитивное свойство интеграла).

□ Возьмем произвольное $\varepsilon > 0$. Так как по условию функция интегрируема на отрезке $[a, c]$, то в силу теоремы 1 существует разбиение T_1 отрезка $[a, c]$ такое, что $\bar{S}_{T_1} - \underline{S}_{T_1} < \varepsilon/2$. Аналогично, функция $f(x)$ интегрируема на отрезке $[c, b]$ и, значит, существует разбиение T_2 отрезка $[c, b]$ такое, что $\bar{S}_{T_2} - \underline{S}_{T_2} < \varepsilon/2$. Эти разбиения T_1 и T_2 в совокупности образуют разбиение T отрезка $[a, b]$, причем

$$\bar{S}_T - \underline{S}_T = (\bar{S}_{T_1} + \bar{S}_{T_2}) - (\underline{S}_{T_1} + \underline{S}_{T_2}) = \\ = (\bar{S}_{T_1} - \underline{S}_{T_1}) + (\bar{S}_{T_2} - \underline{S}_{T_2}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Итак, для произвольного $\varepsilon > 0$ нам удалось построить разбиение T отрезка $[a, b]$ такое, что $\bar{S}_T - \underline{S}_T < \varepsilon$. Это и означает, что функция $f(x)$ интегрируема на отрезке $[a, b]$.

Из неравенств $\underline{S}_{T_1} \leqslant \int_a^c f(x)dx \leqslant \bar{S}_{T_1}$, $\underline{S}_{T_2} \leqslant \int_c^b f(x)dx \leqslant \bar{S}_{T_2}$ следует, что

$$\underline{S}_T = \underline{S}_{T_1} + \underline{S}_{T_2} \leqslant \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \leqslant \bar{S}_{T_1} + \bar{S}_{T_2} = \bar{S}_T\,.$$

Таким образом, как $\int_a^b f(x)\,dx$, так и $\int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ разделяют множества $\{\underline{S}_T\}$ и $\{\bar{S}_T\}$ сумм Дарбу для отрезка $[a, b]$. Поскольку эти множества разделяются лишь одним числом, справедливость равенства (2) доказана.■

Отметим, что если $b < c$, то

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

Значит, и в этом случае

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

Равенство (2) имеет наглядный геометрический смысл: оно выражает *свойство аддитивности площади плоской фигуры*. Так, площадь S криволинейной трапеции $aABb$, изображенной на рис. 138, равна $S_1 + S_2$, где S_1 —

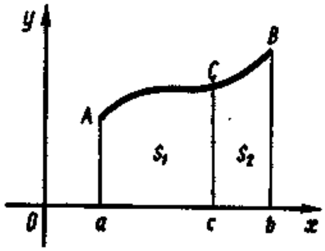


Рис. 138

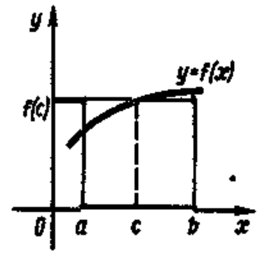


Рис. 139

площадь трапеции $aACc$, а S_2 — площадь трапеции $cCBb$. Но $S_1 = \int_a^c f(x)\,dx$, $S_2 = \int_c^b f(x)\,dx$, $S = \int_a^b f(x)\,dx$, откуда следует равенство (2).

С физической точки зрения равенство (2) выражает *свойство аддитивности массы стержня.*

6. Теорема о среднем для определенного интеграла

Теорема 4 (о среднем). *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то существует точка $c \in [a, b]$ такая, что $\int_a^b f(x)\,dx = f(c)(b-a)$.*

Число $f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx$ называется *средним значением функции* $f(x)$ на отрезке $[a, b]$.

□ Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то по теореме 2 она интегрируема на нем. Пусть m и M — соответственно наименьшее и наибольшее значения функции на $[a, b]$. Выражения $m(b-a)$ и $M(b-a)$ являются нижней и верхней суммами Дарбу, соответствующими разбиению отрезка $[a, b]$, который состоит лишь из одной части — самого этого отрезка. Но $\int_a^b f(x)dx$ разделяет суммы Дарбу и потому

$$m(b-a) \leqslant \int_a^b f(x)dx \leqslant M(b-a),$$

откуда

$$m \leqslant \frac{1}{b-a} \cdot \int_a^b f(x)dx \leqslant M.$$

Число $\frac{1}{b-a}\int_a^b f(x)\,dx$ заключено между m и M. Так как $f(x)$ — непрерывная на $[a, b]$ функция, то по теореме о промежуточном значении (см. § 18 п. 1) указанное значение достигается функцией в некоторой точке c отрезка $[a, b]$:

$$\frac{1}{b-a} \cdot \int_a^b f(x)dx = f(c),$$

откуда находим, что $\int_a^b f(x)dx = f(c)\cdot(b-a)$. ■

Геометрический смысл этой теоремы состоит в том, что *площадь криволинейной трапеции равна площади прямоугольника, имеющего то же основание, что и трапеция, причем высота прямоугольника равна ординате $f(c)$ в некоторой точке c, лежащей между a и b* (рис. 139).

§ 40. Формула Ньютона — Лейбница

В этом параграфе мы докажем основную формулу интегрального исчисления, устанавливающую связь между понятиями определенного интеграла и первообразной.

1. Существование первообразной у непрерывной функции

Если функция $f(x)$ интегрируема на отрезке $[a, b]$, то она интегрируема по любой части этого отрезка и потому при любом $x \in [a, b]$ существует интеграл $\int_a^x f(x)dx$. Чтобы не смешивать обозначения верхнего предела и переменной интегрирования, будем записывать этот интеграл в виде $\int_a^x f(t)\,dt$. Рассмотрим функцию $\Phi(x) = \int_a^x f(t)\,dt$. Докажем справедливость следующей теоремы.

Теорема 1. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то функция $\Phi(x) = \int_a^x f(t)dt$ дифференцируема в любой внутренней точке x этого отрезка, причем $\Phi'(x) = f(x)$.*

Иными словами, *интеграл с переменным верхним пределом является одной из первообразных для непрерывной подынтегральной функции.*

□ Найдем производную функции $\Phi(x)=\int\limits_a^x f(t)\,dt$. Выберем Δx столь малым, чтобы точка $x+\Delta x$ лежала внутри отрезка $[a, b]$; тогда $\Phi(x+\Delta x)=$ $=\int\limits_a^{x+\Delta x} f(t)\,dt$. Далее,

$$\Delta\Phi=\Phi(x+\Delta x)-\Phi(x)=\int\limits_a^{x+\Delta x} f(t)\,dt-\int\limits_a^{x} f(t)\,dt=$$

$$=\int\limits_a^{x} f(t)\,dt+\int\limits_x^{x+\Delta x} f(t)\,dt-\int\limits_a^{x} f(t)\,dt=\int\limits_x^{x+\Delta x} f(t)\,dt$$

(здесь было использовано аддитивное свойство интеграла).

Теперь к полученному интегралу применим теорему о среднем значении (см. теорему 4 § 39): $\Delta\Phi=\int\limits_x^{x+\Delta x} f(t)\,dt=f(c)\,\Delta x$, где $c\in[x, x+\Delta x]$ (или $c\in[x+\Delta x, x]$, если $\Delta x<0$). Итак, $\Delta\Phi=f(c)\Delta x$, а $\frac{\Delta\Phi}{\Delta x}=f(c)$.

Так как функция $f(x)$ непрерывна и $c\to x$ при $\Delta x\to 0$, то $\lim\limits_{\Delta x\to 0} f(c)=f(x)$. Поэтому

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

Из доказанного утверждения вытекает, что если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она имеет на этом отрезке первообразную, а именно функцию Φ, где $\Phi(x)=\int\limits_a^x f(t)\,dt$.

Поэтому доказанная теорема называется **теоремой о существовании первообразной для непрерывной функции.**

2. Основная формула интегрального исчисления (формула Ньютона — Лейбница)

Теорема 2. *Если функция $y=f(x)$ непрерывна на отрезке $[a, b]$, то*

$$\int\limits_a^b f(x)\,dx=F(b)-F(a),\qquad (1)$$

где $F(x)$ — первообразная для функции $f(x)$.

□ Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на нем и, значит, существует $\int\limits_a^b f(x)\,dx$. Далее, в силу непрерывности функции $f(x)$ на $[a, b]$, на этом отрезке существует ее первообразная.

Согласно теореме 1, функция $\Phi(x)=\int_a^x f(t)dt$ является одной из первообразных для функции $f(x)$; следовательно, для любой первообразной $F(x)$ имеем

$$\Phi(x)=F(x)+C. \tag{2}$$

Заметим, что $\Phi(a)=\int_a^a f(t)dt=0$. Из равенства (2) заключаем, что $\Phi(a)=F(a)+C$, т. е. $0=F(a)+C$; значит, $C=-F(a)$. Итак, $\Phi(x)=F(x)-F(a)$, в частности, $\Phi(b)=F(b)-F(a)$. Но $\Phi(b)=\int_a^b f(t)dt=\int_a^b f(x)dx$. Итак,

$$\int_a^b f(x)dx=F(b)-F(a). \quad \blacksquare$$

Равенство (1) называется *формулой Ньютона — Лейбница.*

Разность $F(b)-F(a)$ записывают в виде $F(x)\big|_a^b$; тогда $\int_a^b F(x)dx=F(x)\big|_a^b$.

Например, $\frac{x^3}{3}$ — одна из первообразных для функции x^2. Поэтому $\int_a^b x^2dx=\frac{x^3}{3}\Big|_a^b=\frac{b^3}{3}-\frac{a^3}{3}=\frac{b^3-a^3}{3}$.

3. Свойства определенного интеграла

Из формулы Ньютона — Лейбница легко выводятся основные свойства определенного интеграла. Во всех этих свойствах предполагается, что функции непрерывны на рассматриваемых промежутках.

1^0. *Интеграл от суммы двух функций $f_1(x)$ и $f_2(x)$ по отрезку $[a, b]$ равен сумме интегралов от этих функций по тому же отрезку:*

$$\int_a^b (f_1(x)+f_2(x))dx=\int_a^b f_1(x)dx+\int_a^b f_2(x)dx.$$

□ Из свойств неопределенного интеграла вытекает, что если $F_1(x)$ — первообразная для $f_1(x)$, а $F_2(x)$ — первообразная для $f_2(x)$, то первообразной для $f_1(x)+f_2(x)$ служит функция $F_1(x)+F_2(x)$. Поэтому

$$\int_a^b (f_1(x)+f_2(x))dx=(F_1(b)+F_2(b))-(F_1(a)+F_2(a))=$$

$$=(F_1(b)-F_1(a))+(F_2(b)-F_2(a))=\int_a^b f_1(x)dx+\int_a^b f_2(x)dx. \quad \blacksquare$$

Аналогично доказывается следующее свойство.

2^0. *Постоянный множитель можно вынести за знак интеграла:*

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

Пример. Вычислить $\int\limits_{-2}^{1}(2x^3+3x-4)\,dx$.

○ Имеем

$$\int\limits_{-2}^{1}(2x^3+3x-4)\,dx=\int\limits_{-2}^{1}2x^3dx+\int\limits_{-2}^{1}3x\,dx+\int\limits_{-2}^{1}-4dx=$$

$$=2\int\limits_{-2}^{1}x^3dx+3\int\limits_{-2}^{1}x\,dx-4\int\limits_{-2}^{1}dx=$$

$$=2\cdot\frac{x^4}{4}\Big|_{-2}^{1}+3\cdot\frac{x^2}{2}\Big|_{-2}^{1}-4\cdot x\Big|_{-2}^{1}=$$

$$=\frac{1}{2}(1^4-(-2)^4)+\frac{3}{2}(1^2-(-2)^2)-4(1-(-2))=-24. \bullet$$

3^0. *Если $f(x)\geqslant 0$ на отрезке $[a, b]$, то*

$$\int\limits_a^b f(x)\,dx\geqslant 0. \tag{3}$$

□ В самом деле, если $f(x)\geqslant 0$, то $\underline{S}_\tau\geqslant 0$, а тогда тем более $\int\limits_a^b f(x)\,dx\geqslant 0$. ■

Неравенство (3) допускает простое геометрическое истолкование: *площадь криволинейной трапеции, ограниченной сверху графиком функции, принимающей только неотрицательные значения, есть неотрицательное число.*

4^0. *Если на отрезке $[a, b]$ выполняется неравенство $f(x)\leqslant g(x)$, то*

$$\int\limits_a^b f(x)\,dx\leqslant\int\limits_a^b g(x)\,dx. \tag{4}$$

□ В самом деле, $g(x)-f(x)\geqslant 0$, а тогда согласно свойству 3^0 имеем $\int\limits_a^b(g(x)-f(x))\,dx\geqslant 0$, т. е. $\int\limits_a^b g(x)\,dx-\int\limits_a^b f(x)\,dx\geqslant 0$, откуда и следует неравенство (4). ■

Геометрический смысл неравенства (4) рекомендуем выяснить самостоятельно.

4. Интегрирование по частям в определенном интеграле

Для определенного интеграла формула интегрирования по частям принимает следующий вид:

$$\int\limits_a^b u\,dv=uv\Big|_a^b-\int\limits_a^b v\,du. \tag{5}$$

В самом деле, если $\int u\,dv=F(x)+C_1$, $\int v\,du=\Phi(x)+C_2$, то применяя формулу интегрирования по частям для неопределенного интеграла имеем $\int u\,dv=uv-\int v\,du$, т. е.

$$F(x)=u(x)v(x)-\Phi(x)+C.$$

Поэтому $F(b)=u(b)\,v(b)-\Phi(b)+C$ и $F(a)=u(a)\,v(a)-\Phi(a)+C$. Значит,

$$F(b)-F(a)=(u(b)\,v(b)-u(a)\,v(a))-(\Phi(b)-\Phi(a)),$$

а это и есть формула (5).

Пример. Вычислить $\int\limits_1^2 xe^x dx$.

○ Положим $u = x$, $dv = e^x dx$. Тогда $du = dx$, $v = e^x$. Используя формулу (5), получим

$$\int\limits_1^2 xe^x dx = xe^x \Big|_1^2 - \int\limits_1^2 e^x dx =$$
$$= 2 \cdot e^2 - 1 \cdot e^1 - e^x \Big|_1^2 = 2e^2 - e - (e^2 - e) = e^2. \; \bullet$$

5. Замена переменной в определенном интеграле

Пусть $F(x)$ является первообразной для $f(x)$ на отрезке $[a, b]$ и пусть $x = \varphi(t)$ — дифференцируемая функция, отображающая отрезок $[\alpha, \beta]$ в отрезок $[a, b]$, причем $\varphi(\alpha) = a$, $\varphi(\beta) = b$. В п. 2 § 37 мы установили, что

$$\int f(\varphi(t))\varphi'(t)dt = F(\varphi(t)) + C.$$

Значит,

$$\int\limits_\alpha^\beta f(\varphi(t))\varphi'(t)dt = F(\varphi(t)) \Big|_\alpha^\beta =$$
$$= F(\varphi(\beta)) - F(\varphi(\alpha)) = F(b) - F(a) = \int\limits_a^b f(x)dx.$$

Итак, мы приходим к следующему утверждению.

Теорема 3. *Пусть функция $y = f(x)$ имеет первообразную на отрезке $[a, b]$, а функция $x = \varphi(t)$ определена на отрезке $[\alpha, \beta]$ и дифференцируема внутри этого отрезка, причем $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и $\varphi([\alpha, \beta]) = [a, b]$. Тогда*

$$\int\limits_a^b f(x)dx = \int\limits_\alpha^\beta f(\varphi(t))\varphi'(t)dt. \tag{6}$$

На этом утверждении и основан метод замены переменной под знаком определенного интеграла. Заметим, что на практике формула (6) используется как «слева направо», так и «справо налево».

Условие $\varphi([\alpha, \beta]) = [a, b]$ заведомо выполняется, если функция $x = \varphi(t)$ монотонна на отрезке $[\alpha, \beta]$. Это имеет место, если ее производная сохраняет знак на $[\alpha, \beta]$.

Примеры. 1. Вычислить $\int\limits_0^a \sqrt{a^2 - x^2}\, dx$.

○ Воспользуемся тригонометрической подстановкой $x = \varphi(t) = a \sin t$, $0 \leqslant t \leqslant \pi/2$. Найдем пределы интегрирования α и β для новой переменной t.

Функция $\varphi(t) = a \sin t$ на отрезке $[0, \pi/2]$ определена и дифференцируема внутри него, причем $\varphi(0) = 0$, $\varphi(\pi/2) = a$ и $\varphi([0, \pi/2]) = [0, a]$. Значит, можно применить формулу (6). Используя решение примера 4 § 37, получаем

$$\int\limits_0^a \sqrt{a^2 - x^2}\, dx = \frac{a^2}{2}\int\limits_0^{\pi/2} dt + \frac{a^2}{2}\int\limits_0^{\pi/2} \cos 2t\, dt =$$
$$= \frac{a^2}{2} t \Big|_0^{\pi/2} + \frac{a^2}{2} \frac{\sin 2t}{2} \Big|_0^{\pi/2} = \frac{a^2}{2}(\pi/2 - 0) +$$
$$+ \frac{a^2}{4}(\sin \pi - \sin 0) = \frac{\pi a^2}{4}. \; \bullet$$

2. Вычислить $\int_{0,5}^{1,5}\frac{dx}{4x^2-4x+5}$

○ Выделим полный квадрат в знаменателе: $4x^2-4x+5=(2x-1)^2+4$. Положим $u=2x-1$; тогда $du=2dx$. Если $x=0{,}5$, то $u=2x-1=2\cdot 0{,}5-1=0$; если $x=1{,}5$, то $u=2x-1=2\cdot 1{,}5-1=2$. Таким образом, 0 и 2 — новые пределы интегрирования. Функция $u=2x-1$ на отрезке $[0{,}5;\ 1{,}5]$ определена, дифференцируема и монотонно возрастает; значит, можно воспользоваться формулой (6) (но если в предыдущем примере мы использовали эту формулу «слева направо», то теперь будем идти «справа налево»). Получаем

$$\int_{0,5}^{1,5}\frac{dx}{4x^2-4x+5}=\frac{1}{2}\int_{0,5}^{1,5}\frac{2dx}{(2x-1)^2+4}=\frac{1}{2}\int_{0}^{2}\frac{du}{u^2+4}=$$

$$=\frac{1}{2}\cdot\frac{1}{2}\operatorname{arctg}\frac{u}{2}\Big|_0^2=\frac{1}{4}(\operatorname{arctg}1-\operatorname{arctg}0)=\frac{1}{4}\left(\frac{\pi}{4}-0\right)=\frac{\pi}{16}. \bullet$$

§ 41. Геометрические приложения определенного интеграла

1. Вычисление площадей плоских фигур

В § 39, используя интуитивное представление о площади плоской фигуры и ограничившись нестрогим рассуждением, мы установили связь между определенным интегралом и площадью криволинейной трапеции. Здесь мы придадим указанному рассуждению необходимую строгость.

В курсе геометрии рассматривались площади треугольников, параллелограммов, трапеций и вообще произвольных многоугольников. Теперь введем понятие площади для любой плоской фигуры F, ограниченной непрерывным замкнутым контуром. Хотя площадь — понятие геометрическое, оно изучается в курсе математического анализа, так как и определение, и вычисление площади проводится средствами математического анализа. То же самое относится и к таким понятиям, как длина дуги, объем тела (об этом речь пойдет в следующих пунктах).

Определение 1. Пусть F — плоская фигура, ограниченная непрерывным замкнутым контуром. Рассмотрим всевозможные многоугольники, содержащиеся внутри F (множество их площадей обозначим X), а также всевозможные многоугольники, содержащие F внутри себя (множество их площадей обозначим Y). Числовое множество X расположено левее числового множества Y и, значит, имеется хотя бы одно разделяющее их число. Если это разделяющее число единственно, то оно называется *площадью фигуры* F, а сама фигура F — *квадрируемой.*

Ясно, что любой многоугольник — квадрируемая фигура (тогда множества X и Y имеют общий элемент — площадь самого многоугольника) и к его площади применимо определение 1. Для квадрируемых фигур справедливы такие свойства площади, верные для многоугольников, как равенство площадей равных фигур и свойство аддитивности («площадь суммы равна сумме площадей»).

Теорема 1. *Для того чтобы фигура F была квадрируема, необходимо и достаточно, чтобы для любого $\varepsilon>0$ существовали два многоугольника P и Q: $P\subset F\subset Q$ такие, что разность их площадей меньше ε: $S(Q)-S(P)<\varepsilon$.*

Эта теорема вытекает из критерия единственности разделяющего числа (см. § 10, п. 3).

Теорема 2. *Криволинейная трапеция, ограниченная прямыми $x=a$, $x=b$ $(a<b)$, $y=0$ и графиком непрерывной и неотрицательной на отрезке $[a, b]$ функции $y=f(x)$, квадрируема и ее площадь S вычисляется по формуле*

$$S=\int_a^b f(x)\,dx. \qquad (1)$$

□ Так как функция $y=f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке. Тогда по теореме 1 § 39 для любого $\varepsilon>0$ существует такое разбиение T отрезка $[a, b]$, что $\bar{S}_T-\underline{S}_T<\varepsilon$.

Суммы Дарбу $\underline{S}_T$ и $\bar{S}_T$ геометрически представляют собой площади двух ступенчатых фигур: фигуры P, вписанной в трапецию F, и фигуры Q, описанной около трапеции F (см. рис. 133 и 134). Следовательно, $P\subset F\subset Q$ и поскольку $S(Q)=\bar{S}_T$, а $S(P)=\underline{S}_T$, имеем $S(Q)-S(P)<\varepsilon$. Таким образом, для фигуры F выполняются условия теоремы 1 и, значит, эта фигура квадрируема.

Для любого разбиения T отрезка $[a, b]$ справедливо неравенство $\underline{S}_T\leqslant S\leqslant\bar{S}_T$, где S — площадь квадрируемой фигуры F, т. е. S — число, разделяющее множества нижних и верхних сумм Дарбу. Но таким числом является $\int_a^b f(x)\,dx$; поэтому в силу единственности разделяющегося числа получим $S=\int_a^b f(x)\,dx$. ■

С помощью интеграла можно находить площади не только криволинейной трапеции, ограниченной прямыми $x=a$, $x=b(a<b)$, $y=0$ и графиком непрерывной и неотрицательной на $[a, b]$ функции $y=f(x)$, но и плоских фигур более сложного вида.

Рассмотрим фигуру, ограниченную прямыми $x=a$, $x=b$ $(a<b)$ и графиками функций $y=f_1(x)$, $y=f_2(x)$ непрерывных на отрезке $[a, b]$ и таких, что для всех x из этого отрезка выполняется неравенство $0\leqslant f_2(x)\leqslant f_1(x)$ (на рис. 140 эта фигура заштрихована). Площадь S такой фигуры можно вычислить следующим образом:

$$S=S_{aA'B'b}-S_{aABb}=\int_a^b f_1(x)dx-\int_a^b f_2(x)dx=\int_a^b (f_1(x)-f_2(x))dx.$$

Рассмотрим теперь фигуру такого же вида, как на рис. 140, но не будем требовать, чтобы функции $f_1(x)$ и $f_2(x)$ были неотрицательными (рис. 141, *а*). Пусть m — наименьшее значение функции $y=f_2(x)$ на отрезке $[a, b]$. Тогда функции $h_1(x)=f_1(x)+|m|$, $h_2(x)=f_2(x)+|m|$ непрерывны и неотрицательны на отрезке $[a, b]$ и площадь криволинейной трапеции, ограниченной прямыми $x=a$, $x=b$ и графиками функций $y=h_1(x)$, $y=h_2(x)$ (рис. 141, *б*), вычисляется, как было показано выше, по формуле

$$S=\int_a^b (h_1(x)-h_2(x))dx,$$

т. е.

$$S=\int_a^b ((f_1(x)+|m|)-(f_2(x)+|m|))dx=\int_a^b (f_1(x)-f_2(x))dx.$$

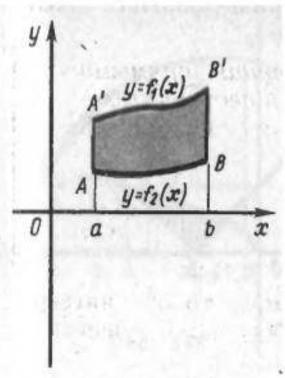


Рис. 140

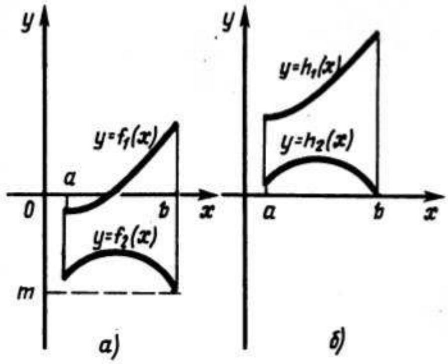


Рис. 141

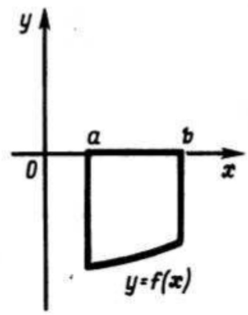


Рис. 142

Так как заданная фигура (рис. 141, *а*) равновелика полученной (рис. 141, *б*), то площадь S вычисляется по формуле

$$S=\int_a^b (f_1(x)-f_2(x))dx. \qquad (2)$$

В частности, для криволинейной трапеции, изображенной на рис. 131, получаем

$$S=\int_a^b (f(x)-0)dx=\int_a^b f(x)dx,$$

а для криволинейной трапеции, изображенной на рис. 142, получаем

$$S=\int_a^b (0-f(x))dx=-\int_a^b f(x)dx.$$

Примеры. 1. Вычислить площадь фигуры, ограниченной линиями

$$y=x,\ y=5-x,\ x=1,\ x=2.$$

○ Фигура, площадь которой надо найти, изображена на рис. 143. Воспользовавшись формулой (2), получим

$$S=\int_1^2 ((5-x)-x)dx=\int_1^2 (5-2x)dx=\int_1^2 5dx-2\int_1^2 x\,dx=$$

$$=5x\Big|_1^2-2\cdot\frac{x^2}{2}\Big|_1^2=5(2-1)-(4-1)=2. \bullet$$

2. Вычислить площадь фигуры, ограниченной линиями $y=x-2$, $y=x^2-4x+2$.

○ Построив прямую $y=x-2$ и параболу $y=x^2-4x+2$, получим фигуру, площадь которой требуется найти (рис. 144). Эта площадь выражается формулой (2), где $f_1(x)=x-2$, $f_2(x)=x^2-4x+2$, а пределы интегрирования a и b равны абсциссам точек пересечения параболы и прямой. Для отыскания абсцисс этих точек решим уравнение $f_1(x)=f_2(x)$. Имеем $x-2=x^2-4x+2$, т. е. $x^2-5x+4=0$, откуда $x_1=1$, $x_2=4$. Итак, $a=1$, $b=4$. Значит,

$$S=\int_1^4 ((x-2)-(x^2-4x+2))dx=\int_1^4 (5x-x^2-4)dx=$$

$$=5\cdot\frac{x^2}{2}\Big|_1^4-\frac{x^3}{3}\Big|_1^4-4x\Big|_1^4=\frac{5}{2}(16-1)-\frac{1}{3}(64-1)-4(4-1)=4{,}5. \bullet$$

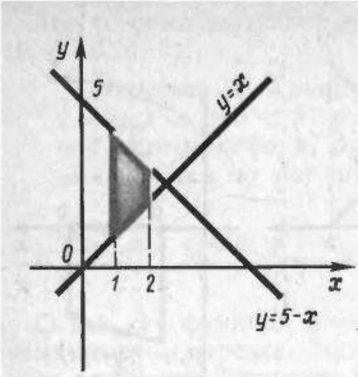


Рис. 143

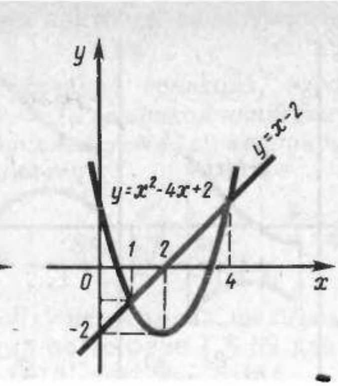


Рис. 144

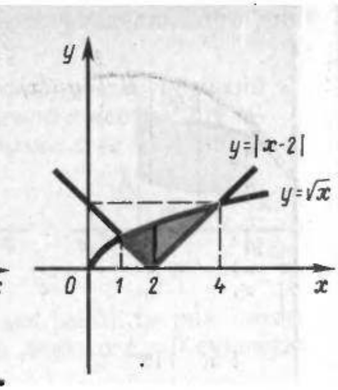


Рис. 145

3. Найти площадь фигуры, ограниченной графиками функций $y=\sqrt{x}$, $y=|x-2|$.

○ Фигура, площадь которой требуется найти, изображена на рис. 145. Эту фигуру можно представить в виде объединения криволинейных трапеций, если провести прямую $x=2$. Тогда площадь S интересующей нас фигуры равны S_1+S_2, где S_1 и S_2 — площади криволинейных трапеций, отмеченных на рис. 145, соответственно горизонтальной и вертикальной штриховкой. Имеем

$$S_1=\int_1^2(\sqrt{x}-(2-x))dx=\int_1^2(x^{1/2}+x-2)dx=\frac{x^{3/2}}{3/2}\Big|_1^2+\frac{x^2}{2}\Big|_1^2-2x\Big|_1^2=$$

$$=\frac{2}{3}(2\sqrt{2}-1)+\frac{1}{2}(4-1)-2(2-1)=\frac{8\sqrt{2}-7}{6};$$

$$S_2=\int_2^4(\sqrt{x}-(x-2))dx=\frac{2x^{3/2}}{3}\Big|_2^4-\frac{x^2}{2}\Big|_2^4+2x\Big|_2^4=$$

$$=\frac{2}{3}(8-2\sqrt{2})-\frac{1}{2}(16-4)+2(4-2)=\frac{10-4\sqrt{2}}{3};$$

$$S=S_1+S_2=\frac{8\sqrt{2}-7}{6}+\frac{10-4\sqrt{2}}{3}=\frac{13}{6}.$$ ●

2. Вычисление длин дуг

В школьном курсе математики рассматривался вопрос о вычислении длин отрезков прямой, длины окружности, а также различных ее частей. В приложениях математики возникает потребность в вычислении длин дуг произвольных кривых. Однако, чтобы вычислить длину произвольной кривой, нужно быть уверенным в том, что рассматриваемая кривая имеет конечную длину.

В средней школе длиной окружности называют предел последовательности периметров вписанных в окружность правильных многоугольников (при неограниченном удвоении числа сторон). Однако это определение неприменимо к произвольным кривым. Дадим общее определение понятия длины кривой.

Определение 2. Пусть дана дуга плоской кривой АВ. Разобьем ее на n частей точками $M_1, M_2, \ldots, M_{n-1}$ и, соединив каждые две соседние точки отрезком прямой, получим ломаную, вписанную в дугу AB. Рассмотрим всевозможные вписанные ломаные и обозначим множество их периметров через X. Если множество X ограничено сверху, то дуга AB называется *спрямляемой* (имеющей длину), а верхняя грань множества X — *длиной дуги* AB.

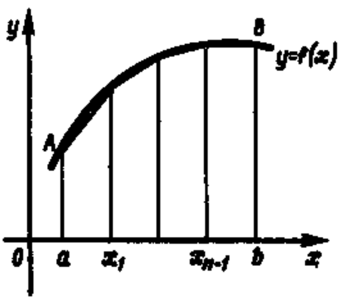


Рис. 146

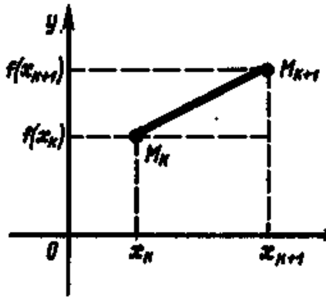


Рис. 147

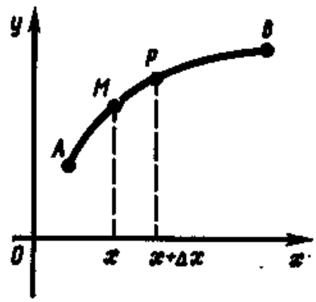


Рис. 148

Ясно, что любая ломаная спрямляема и к ее длине применимо определение 2. Ясно, также, что равные дуги имеют равные длины. Несколько сложнее устанавливается аддитивность длины (соответствующую теорему мы приводим без доказательства).

Теорема 3. *Пусть дуга AB разбита на две части точкой C. Если дуги AC и CB спрямляемы, то и дуга AB спрямляема, причем $\breve{AB} = \breve{AC} + \breve{CB}$.*

Определение 3. Пусть кривая AB представляет собой график функции $y = f(x)$, непрерывной на отрезке $[a, b]$. Если $f'(x)$ непрерывна на $[a, b]$, то AB называют *гладкой кривой.*

Теорема 4. *Гладкая кривая спрямляема.*

□ Произведем разбиение T отрезка $[a, b]$ точками $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$, восставим из этих точек перпендикуляры к оси x до пересечения с графиком функции $y = f(x)$ и рассмотрим ломаную с вершинами в полученных точках пересечения (рис. 146). Периметр этой ломаной обозначим l_T. Нам нужно доказать, что множество $\{l_T\}$ ограничено сверху.

Рассмотрим отдельно звено M_kM_{k+1} этой ломаной (рис. 147) и найдем его длину l_k. Имеем

$$l_k = \sqrt{(f(x_{k+1}) - f(x_k))^2 + (x_{k+1} - x_k)^2}.$$

Применим к разности $f(x_{k+1}) - f(x_k)$ теорему Лагранжа: $f(x_{k+1}) - f(x_k) = f'(c_k)(x_{k+1} - x_k)$, где $x_k < c_k < x_{k+1}$, и обозначим $x_{k+1} - x_k$ через Δx_k. Тогда получим

$$l_k = \sqrt{(f'(c_k))^2(\Delta x_k)^2 + (\Delta x_k)^2} = \sqrt{1 + (f'(c_k))^2}\,\Delta x_k.$$

По условию, функции $f'(x)$, а, значит, и $(f'(x))^2$, непрерывна на отрезке $[a, b]$; следовательно, она достигает на этом отрезке своего наибольшего значения M^2, а потому $\sqrt{1 + (f'(c_k))^2} \leqslant \sqrt{1 + M^2}$. Таким образом, $l_k \leqslant \sqrt{1 + M^2}\,\Delta x_k$.

Используя эти неравенства, для периметра ломаной получаем следующую оценку:

$$l_T = \sum_{k=0}^{n-1} l_k \leqslant \sum_{k=0}^{n-1} \sqrt{1 + M^2} \cdot \Delta x_k = \sqrt{1 + M^2} \sum_{k=0}^{n-1} \Delta x_k = \sqrt{1 + M^2}\,(b - a).$$

Итак, при любом разбиении T имеем $l_T \leqslant \sqrt{1 + M^2}\,(b - a)$, а это и означает, что множество периметров вписанных в дугу AB ломаных ограничено сверху. ■

Теорема 5. *Длина l гладкой кривой удовлетворяет двойному неравенству*

$$\sqrt{1+m^2}(b-a)\leqslant l\leqslant\sqrt{1+M^2}(b-a), \tag{3}$$

где m и M — соответственно наименьшее и наибольшее значения функции $f'(x)$ на отрезке $[a, b]$.

□ В теореме 4 мы доказали, что $\sqrt{1+M^2}(b-a)$ — верхняя граница для множества периметров вписанных в дугу AB ломаных. В то же время длина l дуги AB — верхняя грань указанного множества, т. е. наименьшая из верхних границ. Значит, $l\leqslant\sqrt{1+M^2}(b-a)$.

Проведя относительно l_T те же рассуждения, что и при доказательстве теоремы 4, но не для M, а для m, получим, что при любом разбиении T справедливо неравенство $l_T\geqslant\sqrt{1+m^2}(b-a)$. Так как $l\geqslant l_T$, то тем более $l\geqslant\sqrt{1+m^2}(b-a)$.

Итак, неравенство (3) полностью доказано. ■

Теорема 6. *Пусть на гладкой дуге AB взята точка M с абсциссой x и $l(x)$ — длина дуги AM, т. е. $l(x)=\breve{AM}$. Тогда $l'(x)=\sqrt{1+(f'(x))^2}$.*

□ Найдем производную функции $l(x)=\breve{AM}$. Пусть Δx — приращение аргумента x (предположим для определенности, что $\Delta x>0$), а P — точка дуги AB с абсциссой $x+\Delta x$ (рис. 148). Тогда $l(x+\Delta x)=\breve{AP}$. Следовательно, $\Delta l=l(x+\Delta x)-l(x)=\breve{AP}-\breve{AM}=\breve{MP}$ (в силу теоремы 3 об аддитивности спрямляемой дуги). Далее, используя формулу (3), находим

$$\sqrt{1+m^2}\Delta x\leqslant\Delta l\leqslant\sqrt{1+M^2}\Delta x, \tag{4}$$

где m и M — соответственно наименьшее и наибольшее значения функции $f'(x)$ на отрезке $[x, x+\Delta x]$.

Разделив неравенство (4) почленно на Δx (напомним, что мы условились считать $\Delta x>0$), получим

$$\sqrt{1+m^2}\leqslant\frac{\Delta l}{\Delta x}\leqslant\sqrt{1+M^2}. \tag{5}$$

Устремим Δx к нулю. Тогда вследствие непрерывности функции $f'(x)$ значения m и M будут стремиться к $f'(x)$, т. е.

$$\lim_{\Delta x\to 0}\sqrt{1+m^2}=\sqrt{1+(f'(x))^2} \text{ и } \lim_{\Delta x\to 0}\sqrt{1+M^2}=\sqrt{1+(f'(x))^2}.$$

Отсюда по теореме о пределе промежуточной функции (см. теорему 9 § 16) из неравенства (5) заключаем, что и

$$\lim_{\Delta x\to 0}\frac{\Delta l}{\Delta x}=\sqrt{1+f'(x)^2}, \text{ т. е. } l'(x)=\sqrt{1+(f'(x))^2}. \blacksquare$$

Теорема 7. *Длина l дуги гладкой кривой, служащей графиком функции $y=f(x)$ на отрезке $[a, b]$, вычисляется по формуле*

$$l=\int_a^b\sqrt{1+(f'(x))^2}\,dx. \tag{6}$$

□ Согласно формуле Ньютона—Лейбница, $\int\limits_a^b \sqrt{1+(f'(x))^2}\,dx = F(x)\Big|_a^b$, где $F(x)$ — первообразная для $\sqrt{1+(f'(x))^2}$. Но в силу теоремы 6 такой первообразной является $l(x)$, причем $l(b) = \smile AB = l$, а $l(a) = 0$ (длина дуги от A до A). Значит,

$$\int\limits_a^b \sqrt{1+(f'(x))^2}\,dx = l(b) - l(a) = l - 0 = l. \blacksquare$$

Следствие. *Пусть кривая задана параметрически уравнениями $x = x(t)$, $y = y(t)$, $\alpha \leqslant t \leqslant \beta$ и пусть функции $x(t)$, $y(t)$ непрерывны на $[\alpha, \beta]$, имеют непрерывные производные в (α, β), причем $x'(t)$ сохраняет в (α, β) постоянный знак. Тогда длина этой кривой вычисляется по формуле*

$$l = \int\limits_\alpha^\beta \sqrt{x'(t)^2 + y'(t)^2}\,dt. \tag{7}$$

□ При заданных условиях имеем $y' = \frac{y'(t)}{x'(t)}$ (см. п. 6 § 31). Перейдя в формуле (5) к новой переменной интегрирования t, получим (считаем для определенности, что $x'(t) > 0$)

$$l = \int\limits_\alpha^\beta \sqrt{1+\left(\frac{y'(t)}{x'(t)}\right)^2}\,x'(t)\,dt = \int\limits_\alpha^\beta \sqrt{x'(t)^2 + y'(t)^2}\,dt. \blacksquare$$

Примеры. 1. Найти длину дуги цепной линии $y = \frac{1}{2}(e^x + e^{-x})$, $0 \leqslant x \leqslant 1$.

○ Имеем $y' = \frac{1}{2}(e^x - e^{-x})$; $\quad 1 + (y')^2 = 1 + \frac{1}{4}(e^x - e^{-x})^2 = \frac{4 + e^{2x} - 2 + e^{-2x}}{4} = \left(\frac{e^x + e^{-x}}{2}\right)^2$; $\sqrt{1+(y')^2} = \frac{e^x + e^{-x}}{2}$. Используя формулу (6), находим

$$l = \int\limits_0^1 \sqrt{1+(y')^2}\,dx = \int\limits_0^1 \frac{e^x + e^{-x}}{2}\,dx = \frac{1}{2}(e^x - e^{-x})\Big|_0^1 = \frac{1}{2}\left(e - \frac{1}{e}\right). \bullet$$

2. Найти длину астроиды $\begin{cases} x = a\cos^3 t, \\ y = a\sin^3 t, \ 0 \leqslant t \leqslant 2\pi. \end{cases}$

○ Так как астроида симметрична относительно осей координат (см. рис. 111), то для отыскания длины всей астроиды достаточно найти длину ее дуги в пределах $0 \leqslant t \leqslant \pi/2$ и полученный результат умножить на 4. Воспользуемся формулой (7). Имеем $x'(t) = (a\cos^3 t)' = -3a\cos^2 t \sin t$, $y'(t) = (a\sin^3 t)' = 3a\sin^2 t \cos t$, $x'(t)^2 + y'(t)^2 = 9a^2(\cos^4 t \sin^2 t + \sin^4 t \cos^2 t) = 9a^2 \sin^2 t \cos^2 t$, $\sqrt{x'(t)^2 + y'(t)^2} = 3a\sin t \cos t$ (так как на $[0, \pi/2]$ и $\sin t$, и $\cos t$ неотрицательны). Тогда

$$l = 4\int\limits_0^{\frac{\pi}{2}} 3a\sin t \cos t\,dt = 6a\int\limits_0^{\frac{\pi}{2}} \sin 2t\,dt = 6a\left(-\frac{\cos 2t}{2}\right)\Big|_0^{\frac{\pi}{2}} =$$
$$= -3a(\cos\pi - \cos 0) = 6a. \bullet$$

3. Вычисление объемов тел

В курсе геометрии изучались объемы некоторых многогранников; например, для объема V прямой призмы установлена формула $V = SH$, где S — площадь основания призмы, а H — высота призмы. Нетрудно распространить понятие объема на цилиндрические тела.

Пусть F — плоская фигура. Восставим в каждой точке этой фигуры перпендикуляр к содержащей ее плоскости α и отложим на каждом перпендикуляре от соответствующей точки фигуры F отрезок длины h (все отрезки располагаются по одну сторону от плоскости α). Множество точек этих отрезков образует пространственное тело P, которое называется *прямым цилиндрическим телом* с основанием F и высотой h.

Примерами прямых цилиндрических тел могут служить прямая призма, прямой круговой цилиндр.

Если основание F квадрируемо, то естественно считать объемом прямого цилиндрического тела произведение площади его основания на высоту. Условимся в дальнейшем рассматривать цилиндрические тела только с квадрируемыми основаниями, не оговаривая этого каждый раз особо. Объединение конечного числа прямых цилиндрических тел без общих внутренних точек будем называть *кусочно цилиндрическим телом*, а его объемом естественно считать сумму объемов составляющих его цилиндрических тел.

Теперь мы можем ввести понятие объема для произвольного пространственного тела.

Определение 4. Пусть P — пространственное тело. Рассмотрим всевозможные кусочно цилиндрические тела, содержащиеся внутри P (множество их объемов обозначим X), а также всевозможные кусочно цилиндрические тела, содержащие P внутри себя (множество их объемов обозначим Y). Числовое множество X расположено левее числового множества Y и, значит, имеется хотя бы одно разделяющее их число. Если это разделяющее число единственно, то оно называется *объемом тела P*, а само тело P — *кубируемым*.

Ясно, что любое прямое цилиндрическое тело с квадрируемым основанием кубируемо и к его объему применимо определение 4; кубируемо и кусочно-цилиндрическое тело. Для кубируемых тел справедливы такие же свойства, как и для квадрируемых фигур, например равенство объемов у равных тел и свойство аддитивности: «объем суммы равен сумме объемов».

Как нетрудно убедиться, определение кубируемости аналогично определению квадрируемости из п. 1. Аналогичными являются и основные теоремы о квадрируемости (см. теорему 2) и о кубируемости.

Теорема 8. *Пусть:* 1) *тело P расположено в полосе между плоскостями $x=a$, $x=b$ $(a<b)$;* 2) *сечение тела любой плоскостью $x=x_0$ такой, что $a\leqslant x_0\leqslant b$, есть квадрируемая фигура; ее площадь обозначим $S(x_0)$;* 3) *функция $S(x)$ непрерывна на отрезке $[a, b]$. Тогда тело P кубируемо и его объем V вычисляется по формуле*

$$V=\int_a^b S(x)\,dx. \tag{8}$$

Сопоставим эту теорему с теоремой 2. В теореме 2 фигура расположена между прямыми $x=a$, $x=b$ $(a<b)$, причем точки фигуры F принадлежат как той, так и другой прямой; это аналогично первому условию теоремы 8. Кроме того, в теореме 2 сечение фигуры F любой прямой $x=x_0$ такой, что $a\leqslant x_0\leqslant b$, есть спрямляемая линия — отрезок длины $f(x_0)$; это аналогично второму условию теоремы 8. Далее, в теореме 2 функция непрерывна на $[a, b]$ — это совпадает с третьим условием теоремы 8. Наконец, в теореме 2 утверждается, что фигура F квадрируема и ее площадь S вычисляется по формуле $S=\int_a^b f(x)\,dx$, — это аналогично заключению теоремы 8.

Доказательства квадрируемости фигуры F в теореме 2 и кубируемости тела P в теореме 8 также вполне аналогичны, но последнее несколько сложнее в техническом плане (мы его не приводим). Ограничимся выводом формулы для вычисления объема. Как мы сейчас убедимся, вывод формулы для вычисления объема основан на тех же идеях, которые были использованы в теоремах 6 и 7 при выводе формулы для длины дуги спрямляемой кривой.

□ Рассмотрим часть тела P, заключенную между плоскостями $x=a$ и $x=x_0$, где x_0 — некоторая точка из отрезка $[a, b]$; объем этой части обозначим $V(x_0)$. Дадим аргументу приращение Δx (пусть для определенности $\Delta x>0$), так, чтобы $x_0+\Delta x\in[a, b]$, и проведем плоскость $x=x_0+\Delta x$; объем части тела, заключенного между плоскостями $x=x_0$ и $x=x_0+\Delta x$, обозначим ΔV. Так как функция $S(x)$ непрерывна на отрезке $[x_0, x_0+\Delta x]$, то на этом отрезке она достигает своего наименьшего значения m и наибольшего значения M. Тогда $m\Delta x\leqslant\Delta V\leqslant M\Delta x$. Из этого неравенства следует, что $m\leqslant\frac{\Delta V}{\Delta x}\leqslant M$. Если устремить Δx к нулю, то в силу непрерывности функции $S(x)$ значения m и M также будут стремиться к $S(x_0)$. Следовательно, и $\frac{\Delta V}{\Delta x}$ стремится к тому же пределу: $\lim\limits_{\Delta x\to 0}\frac{\Delta V}{\Delta x}=S(x_0)$, т. е. $V'(x_0)=S(x_0)$.

Так как проведенные выше рассуждения применимы к любой точке x из $[a, b]$, то для любого x справедливо равенство $V'(x)=S(x)$, т. е. $V(x)$ — первообразная для $S(x)$, причем $V(a)=0$, $V(b)=V$.

Используя формулу Ньютона — Лейбница, находим

$$\int_a^b S(x)\,dx= V(x)\Big|_a^b=V(b)-V(a)=V-0=V.$$

Итак, $V=\int\limits_a^b S(x)\,dx.$ ■

Следствие. *Пусть P — тело, образованное вращением вокруг оси x криволинейной трапеции, ограниченной прямыми $x=a$, $x=b$ $(a<b)$ и графиком непрерывной и неотрицательной на отрезке $[a, b]$ функции $y=f(x)$* (рис. 149). *Тогда тело P кубируемо и его объем V вычисляется по формуле*

$$V=\pi\int_a^b f^2(x)\,dx. \qquad (9)$$

□ Сечение тела вращения P плоскостью $x=x_0$ $(a\leqslant x_0\leqslant b)$ есть круг с радиусом $f(x_0)$; площадь этого круга $S(x_0)=\pi(f(x_0))^2$. Функция $\pi f(x)^2$ непрерывна на отрезке $[a, b]$; значит, согласно теореме 8, тело P кубируемо и его объем V вычисляется так:

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

Примеры. 1. Вычислить объем пирамиды, площадь основания которой равна S, а высота равна H (рис. 150).

○ Так как $\frac{S(x)}{S}=\frac{x^2}{H^2}$, то $S(x)=\frac{S}{H^2}x^2$. Следовательно,

$$V=\int_0^H\frac{S}{H^2}x^2dx=\frac{S}{H^2}\cdot\frac{x^3}{3}\Big|_0^H=\frac{S}{H^2}\cdot\frac{H^3}{3}=\frac{1}{3}SH. \ ●$$

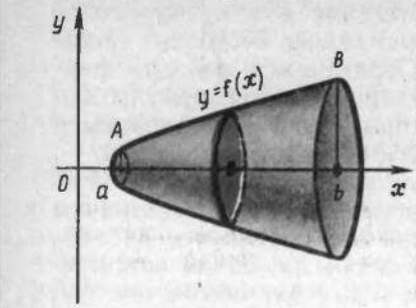


Рис. 149

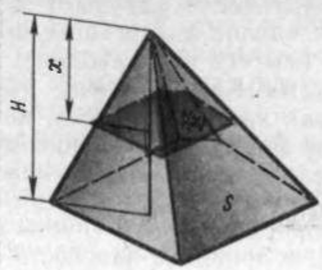


Рис. 150

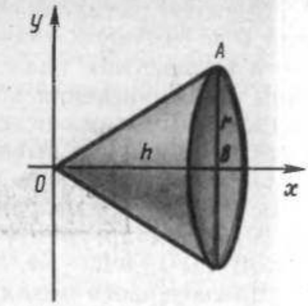


Рис. 151

2. Найти объем шара радиуса R.

○ Рассмотрим круг радиуса R с центром в начале координат. Этот круг при вращении вокруг оси образует шар. Уравнение окружности имеет вид $x^2+y^2=R^2$, поэтому $y^2=R^2-x^2$. Учитывая симметрию круга относительно оси ординат, сначала найдем половину искомого объема:

$$\frac{1}{2}V=\pi\int_0^R y^2dx=\pi\int_0^R (R^2-x^2)dx=$$

$$=\pi\left(R^2x-\frac{x^3}{3}\right)\Big|_0^R=\pi\left(R^3-\frac{R^3}{3}\right)=\frac{2}{3}\pi R^3.$$

Следовательно, объем шара равен $\frac{4}{3}\pi R^3$. ●

3. Вычислить объем конуса, высота которого равна h, радиус основания равен r.

○ Выберем систему координат так, чтобы ось симметрии конуса совпала с осью x (рис. 151), а вершину конуса примем за начало координат. Тогда уравнение прямой OA запишется в виде $y=\frac{r}{h}x$.

Используя формулу (9), получим

$$V=\pi\int_0^h y^2dx=\pi\int_0^h \frac{r^2}{h^2}x^2dx=\frac{\pi r^2}{h^2}\cdot\frac{x^3}{3}\Big|_0^h=\frac{1}{3}\pi r^2h. ●$$

§ 42. Несобственные интегралы

1. Несобственные интегралы первого рода

Для существования определенного интеграла необходимо, чтобы промежуток интегрирования был конечен, а подынтегральная функция ограничена на нем: в противном случае множество сумм Дарбу не будет ограниченным. Возможны случаи, когда одно или оба из этих условий не выполняются, т. е. когда промежуток интегрирования бесконечен или подынтегральная функция не ограничена. Такие интегралы называются *несобственными*. Различают несобственные интегралы первого и второго рода в зависимости от того, имеем ли мы дело с бесконечным промежутком интегрирования или с неограниченной подынтегральной функцией.

Начнем со случая, когда промежутком интегрирования является луч $[a, +\infty)$.

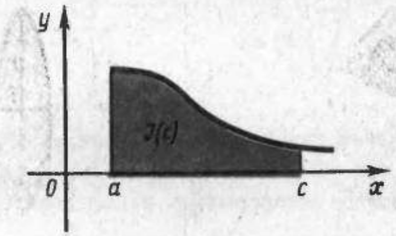


Рис. 152

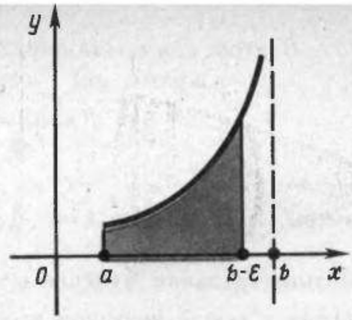


Рис. 153

Определение 1. Пусть функция $y=f(x)$ интегрируема на каждой конечной части луча, т. е. для любого $c>a$ существует интеграл $I(c)=\int\limits_a^c f(x)\,dx$. За значение *несобственного интеграла первого рода* $\int\limits_a^{+\infty} f(x)\,dx$ принимают предел функции $I(c)$, когда c стремится к $+\infty$ (т. е. когда промежуток интегрирования стремится заполнить весь луч $[a, +\infty)$; рис. 152):

$$\int\limits_a^{+\infty} f(x)\,dx = \lim_{c\to+\infty}\int\limits_a^c f(x)\,dx. \qquad (1)$$

Если предел $\lim\limits_{c\to+\infty}\int\limits_a^c f(x)\,dx$ существует и конечен, то несобственный интеграл $\int\limits_a^{+\infty} f(x)\,dx$ называют *сходящимся*, а значение этого предела — значением несобственного интеграла.

Если предел в правой части равенства (1) не существует, то говорят, что несобственный интеграл $\int\limits_a^{+\infty} f(x)\,dx$ *расходится*.

При аналогичных предположениях относительно функции $y=f(x)$ можно рассмотреть случай, когда верхний предел фиксирован, а нижний предел стремится к $-\infty$:

$$\int\limits_{-\infty}^{b} f(x)\,dx = \lim_{c\to-\infty}\int\limits_c^b f(x)\,dx. \qquad (2)$$

Если предел в правой части равенства (2) конечен, то несобственный интеграл $\int\limits_{-\infty}^{b} f(x)\,dx$ называют *сходящимся*; в противном случае его называют *расходящимся*.

Наконец, можно определить и несобственный интеграл вида $\int\limits_{-\infty}^{+\infty} f(x)\,dx$. Будем считать, что функция $y=f(x)$ интегрируема на всей числовой прямой. Выберем на прямой произвольную точку a. Эта точка разобьет прямую на два луча: $(-\infty, a]$ и $[a, +\infty)$. Если существуют несобственные интегралы $\int\limits_{-\infty}^{a} f(x)\,dx$ и $\int\limits_a^{+\infty} f(x)\,dx$, то говорят, что существует и несобственный интеграл

$\int\limits_{-\infty}^{+\infty} f(x)\,dx$. В этом случае полагают

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = \int\limits_{-\infty}^{a} f(x)\,dx + \int\limits_{a}^{+\infty} f(x)\,dx. \tag{3}$$

Примеры. 1. Вычислить $I = \int\limits_{1}^{+\infty} \frac{dx}{1+x^2}$.

○ Подынтегральная функция $y = \frac{1}{1+x^2}$ всюду непрерывна и, следовательно, интегрируема в любом конечном промежутке. Имеем

$$\int\limits_{1}^{+\infty} \frac{dx}{1+x^2} = \lim_{b\to+\infty} \int\limits_{1}^{b} \frac{dx}{1+x^2} = \lim_{b\to+\infty} \operatorname{arctg} x \Big|_{1}^{b} = \lim_{b\to+\infty} (\operatorname{arctg} b - \operatorname{arctg} 1) =$$

$$= \lim_{b\to+\infty} \operatorname{arctg} b - \frac{\pi}{4} = \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4}.$$

Значит, $I = \pi/4$. ●

2. Вычислить $\int\limits_{-\infty}^{\pi/2} \sin x\,dx$.

○ Имеем

$$\int\limits_{-\infty}^{\pi/2} \sin x\,dx = \lim_{a\to-\infty} \int\limits_{a}^{\pi/2} \sin x\,dx = \lim_{a\to-\infty} (-\cos x) \Big|_{a}^{\pi/2} =$$

$$= \lim_{a\to-\infty} \left(-\cos\frac{\pi}{2} + \cos a\right) = \lim_{a\to-\infty} \cos a.$$

Так как $\lim\limits_{a\to-\infty} \cos a$ не существует, то несобственный интеграл $\int\limits_{-\infty}^{\pi/2} \sin x\,dx$ расходится. ●

Запись вычислений несобственных интегралов можно упростить, предварительно найдя первообразную для подынтегральной функции $y = f(x)$. Именно, если $F(x)$ — первообразная функция для $f(x)$, то

$$\int\limits_{a}^{+\infty} f(x)\,dx = \lim_{b\to+\infty} \int\limits_{a}^{b} f(x)dx = \lim_{b\to+\infty} (F(b) - F(a)).$$

Предположим, что существует предел $\lim\limits_{b\to+\infty} F(b)$. Введем обозначение $\lim\limits_{b\to+\infty} F(b) = F(+\infty)$. Тогда

$$\int\limits_{a}^{+\infty} f(x)dx = F(+\infty) - F(a) = F(x)\Big|_{a}^{+\infty}.$$

Аналогично

$$\int\limits_{-\infty}^{b} f(x)dx = F(b) - F(-\infty) = F(x)\Big|_{-\infty}^{b},$$

где $F(-\infty) = \lim\limits_{a\to-\infty} F(a)$. Наконец,

$$\int\limits_{-\infty}^{+\infty} f(x)dx = F(+\infty) - F(-\infty) = F(x)\Big|_{-\infty}^{+\infty}.$$

Примеры. 3. Вычислить $\int\limits_{-\infty}^{+\infty} \frac{dx}{1+x^2}$.

○ Находим

$$\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \operatorname{arctg} x \Big|_{-\infty}^{+\infty} = \lim_{b \to +\infty} \operatorname{arctg} b - \lim_{a \to -\infty} \operatorname{arctg} a =$$

$$= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi. \bullet$$

4. Исследовать на сходимость интеграл $\int_1^{+\infty} \frac{dx}{x^\alpha}$.

○ Если $\alpha = 1$, то

$$\int_1^{+\infty} \frac{dx}{x} = \ln x \Big|_1^{+\infty} = \lim_{b \to +\infty} \ln b - \ln 1 = +\infty.$$

Если $\alpha \neq 1$, то

$$\int_1^{+\infty} \frac{dx}{x^\alpha} = \frac{x^{-\alpha+1}}{-\alpha+1} \Big|_1^{+\infty} = \frac{1}{-\alpha+1} \lim_{b \to +\infty} b^{-\alpha+1} - \frac{1}{-\alpha+1}.$$

Сходимость или расходимость интеграла $\int_1^{+\infty} \frac{dx}{x^\alpha}$ зависит от того, существует или нет предел $\lim_{b \to +\infty} b^{-\alpha+1}$. Если $\alpha > 1$, то $\lim_{b \to +\infty} b^{-\alpha+1} = 0$; если же $\alpha < 1$, то $\lim_{b \to +\infty} b^{-\alpha+1} = \infty$. Таким образом, $\int_1^{+\infty} \frac{dx}{x^\alpha}$ сходится при $\alpha > 1$ и расходится при $\alpha < 1$. ●

2. Несобственные интегралы второго рода

Пусть функция $y = f(x)$ определена во всех точках отрезка $[a, b]$, кроме точки c; пусть $x = c$ — вертикальная асимптота графика функции $y = f(x)$, причем после удаления некоторой достаточно малой ε — окрестности точки c останется два отрезка $[a, c-\varepsilon]$ и $[c+\varepsilon, b]$, на каждом из которых функция интегрируема. Такую точку будем называть *особой*.

Рассмотрим сначала случай, когда особой точкой является точка $x = b$.

Определение 2. Пусть функция $y = f(x)$ не ограничена на отрезке $[a, b]$, но интегрируема на любом отрезке $[a, b-\varepsilon]$, где $\varepsilon > 0$. За значение *несобственного интеграла второго рода принимают* $\lim_{\varepsilon \to 0} \int_a^{b-\varepsilon} f(x)\,dx$ (рис. 153); этот интеграл обозначают $\int_a^b f(x)\,dx$. Если указанный предел существует, то интеграл называют *сходящимся*; если же он не существует, то интеграл называют *расходящимся*.

Аналогично, если функция не ограничена на отрезке $[a, b]$, но интегрируема на любом отрезке $[a+\varepsilon, b]$, то полагают

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \to +0} \int_{a+\varepsilon}^b f(x)\,dx.$$

Наконец, если единственная особая точка c лежит внутри отрезка $[a, b]$, то полагают

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \to +0} \int_a^{c-\varepsilon} f(x)\,dx + \lim_{\varepsilon \to +0} \int_{c+\varepsilon}^b f(x)\,dx.$$

Пусть $F(x)$ — первообразная для функции $f(x)$. Положим $F(a+0) =$

$= \lim\limits_{\varepsilon\to+0} F(a+\varepsilon)$, $F(b-0) = \lim\limits_{\varepsilon\to+0} F(b-\varepsilon)$ (если эти пределы существуют). Тогда для сходящихся интегралов, у которых особыми являются лишь точки a и b, имеем

$$\int_a^b f(x)\,dx = F(b-0) - F(a+0).$$

Если функция $F(x)$ непрерывна в точках a и b, то

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

Аналогично обстоит дело и в случае, когда подынтегральная функция не ограничена в любой окрестности некоторой внутренней точки отрезка $[a, b]$.

Примеры. 1. Вычислить интеграл $\int\limits_0^1 \frac{dx}{\sqrt{1-x^2}}$.

○ Этот интеграл является несобственным, поскольку подынтегральная функция $y = 1/\sqrt{1-x^2}$ не ограничена в любой окрестности точки $x = 1$. Так как первообразная для этой функции равна $\arcsin x$, то, пользуясь определением несобственного интеграла, получаем

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \lim_{\varepsilon\to+0} \arcsin x \Big|_0^{1-\varepsilon} = \lim_{\varepsilon\to+0} \arcsin(1-\varepsilon) - \arcsin 0,$$

откуда, учитывая непрерывность функции $\arcsin x$, находим

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \arcsin x \Big|_0^1 = \frac{\pi}{2} - 0 = \frac{\pi}{2}. \bullet$$

2. Вычислить $\int\limits_2^4 \frac{2x\,dx}{\sqrt[3]{x^2-9}}$.

○ Внутри данного промежутка интегрирования подынтегральная функция имеет одну особую точку $x = 3$. Найдем первообразную для этой функции:

$$\int \frac{2x\,dx}{\sqrt[3]{x^2-9}} = \frac{3}{2}\sqrt[3]{(x^2-9)^2} + C.$$

Так как функция $F(x) = \frac{3}{2}\sqrt[3]{(x^2-9)^2}$ непрерывна в точке $x = 3$, то

$$\int_2^4 \frac{2x\,dx}{\sqrt[3]{x^2-9}} = \frac{3}{2}\sqrt[3]{(x^2-9)^2}\Big|_2^4 = \frac{3}{2}(\sqrt[3]{49} - \sqrt[3]{25}). \bullet$$

3. Вычислить $\int\limits_{-3}^3 \frac{dx}{\sqrt{9-x^2}}$.

○ В данном случае подынтегральная функция имеет две особые точки $x_1 = -3$ и $x_2 = 3$. Пользуясь определением несобственного интеграла и учитывая непрерывность первообразной, получаем

$$\int_{-3}^3 \frac{dx}{\sqrt{9-x^2}} = \arcsin\frac{x}{3}\Big|_{-3}^3 = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi. \bullet$$

4. Исследовать на сходимость интеграл $\int\limits_0^1 \frac{dx}{x}$.

○ Имеем

$$\int_0^1 \frac{dx}{x} = \lim_{\varepsilon\to+0} \int_\varepsilon^1 \frac{dx}{x} = \lim_{\varepsilon\to+0} (\ln 1 - \ln\varepsilon) = -\lim_{\varepsilon\to+0} \ln\varepsilon = +\infty.$$

Значит, данный интеграл расходится. ●

7

Функции нескольких переменных

§ 43. Основные понятия

Определение 1. Пусть D — некоторое множество пар действительных чисел и пусть каждой паре $(x;\ y)$ из D поставлено в соответствие число z. Тогда говорят, что на множестве D задана *функция двух переменных* $z = f(x,\ y)$. Переменные x, y называют *независимыми переменными* (или *аргументами*), z — *зависимой переменной*; говорят также, что $f(x,\ y)$ есть *значение функции* f *в точке* $(x;\ y)$. Множество D называют *областью определения* функции. Все значения, которые принимает функция $f(x,\ y)$ (при $(x;\ y)$, принадлежащих области ее определения), образуют *область значений* функции.

Аналогично можно ввести понятие функции трех переменных $u = f(x,\ y,\ z)$, определенной на множестве D, состоящем не из действительных чисел (как для функции одной переменной) и не из пар действительных чисел (как для функции двух переменных), а из троек действительных чисел $(x;\ y;\ z)$, рассматриваемых в определенном порядке. Можно также ввести понятие функции четырех, пяти и вообще любого конечного числа переменных — все такие функции называют *функциями нескольких переменных*. В дальнейшем мы, говоря о функции нескольких переменных, будем иметь в виду (для большей простоты и наглядности) функцию двух (иногда трех) переменных.

Приведем примеры функций нескольких переменных:

$S = xy$ — площадь прямоугольника со сторонами x, y есть функция двух переменных;

$U = IR$ (закон Ома) — напряжение U на участке электрической цепи есть функция двух переменных: силы тока I и сопротивления R;

$V = xyz$ — объем прямоугольного параллелепипеда со сторонами x, y, z есть функция трех переменных.

Чтобы задать функцию двух (трех) переменных, нужно указать способ, с помощью которого для каждой пары (тройки) значений аргументов можно найти соответствующее значение функции. Наиболее употребительным (как и в случае функций одной переменной) является способ задания функции с помощью формулы $z = f(x,\ y)$, где $f(x,\ y)$ — некоторое выражение с переменными x, y. В таком случае говорят, что функция задана формулой или что функция задана *аналитически*. Выше были приведены примеры функций, заданных аналитически.

Для аналитически заданной функции иногда не указывают явно область

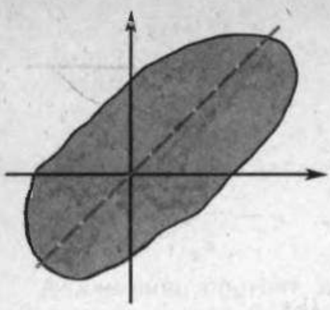


Рис. 154

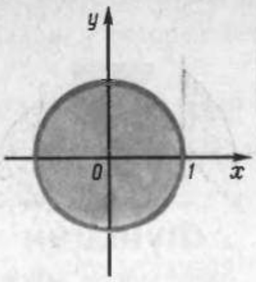


Рис. 155

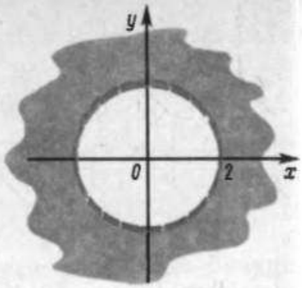


Рис. 156

ее определения. В таком случае подразумевают, что область определения функции $z = f(x, y)$ совпадает с областью определения выражения $f(x, y)$, т. е. с множеством тех значений x, y, при которых выражение (x, y) имеет смысл.

Пример 1. Найти область определения функции:

а) $z = \frac{2x+y}{x-y}$; б) $z = \sqrt[4]{1-x^2-y^2}$; в) $z = \ln(x^2+y^2-4)$; г) $u = \sqrt{x} + \sqrt{y} + \sqrt{z}$.

○ а) Функция не определена лишь в случае, когда $y = x$. Геометрически это означает, что область определения функции состоит из двух полуплоскостей, одна из которых лежит выше, а другая ниже прямой $y = x$ (рис. 154).

б) Функция определена при условии $1 - x^2 - y^2 \geqslant 0$, т. е. $x^2 + y^2 \leqslant 1$. Это круг с центром в начале координат и радиусом 1, включающий свою границу, т. е. окружность $x^2 + y^2 = 1$ (рис. 155).

в) Функция определена при условии $x^2 + y^2 - 4 > 0$, т. е. $x^2 + y^2 > 4$. Это часть плоскости, лежащая вне круга с центром в начале координат и радиусом 2, не включающая границу круга, т. е. окружность $x^2 + y^2 = 4$ (рис. 156).

г) Функция определена при (x, y, z), удовлетворяющих одновременно условиям $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$. ●

Подобно тому, как для множества всех пар действительных чисел геометрической моделью является координатная плоскость, так и для множества всех троек действительных чисел существует геометрическая модель — координатное пространство xyz. Это обычное трехмерное пространство, в котором введена декартова прямоугольная система координат, определяемая тремя взаимно перпендикулярными прямыми с одним началом O и одинаковым масштабом (рис. 157). Точку O называют *началом координат*, а оси x, y, z — соответственно *осью абсцисс, ординат, аппликат*. Плоскости xy, yz, xz называют *координатными плоскостями*; точки плоскости xy удовлетворяют уравнению $z = 0$, точки плоскости yz — уравнению $x = 0$, точки плоскости xz — уравнению $y = 0$.

Если M_0 — точка координатного пространства, то, спроецировав ее на координатные плоскости, найдем координаты проекций x_0, y_0, z_0 (рис. 157). Обратно, если дана тройка чисел, то, отметив точки x_0, y_0, z_0 на осях x, y, z и выполнив соответствующие построения (рис. 157), найдем точку M_0 с координатами x_0, y_0, z_0. Принята запись $M_0(x_0; y_0; z_0)$.

Если отметить в системе координат xyz все точки с абсциссой $x = a$, то получится плоскость, параллельная плоскости yz (рис. 158, *а*); уравнение этой плоскости имеет вид $x = a$.

Аналогично, точки с ординатой $y = b$ образуют плоскость, параллельную плоскости xz (рис. 158, *б*), а точки с аппликатой $z = c$ — плоскость, параллельную плоскости xy (рис. 158, *в*); уравнения построенных плоскостей соответственно имеют вид $y = b$ и $z = c$.

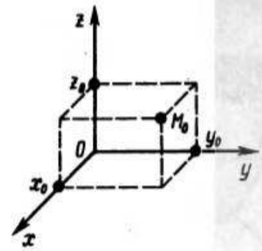


Рис. 157

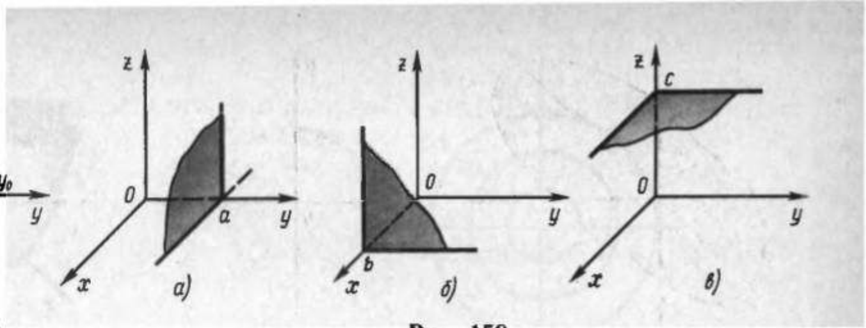


Рис. 158

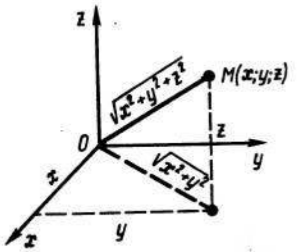


Рис. 159

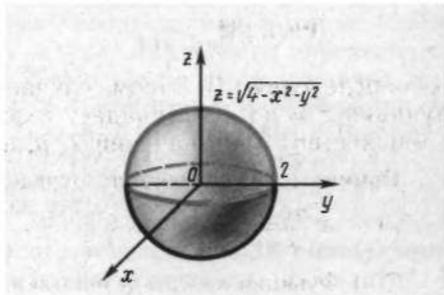


Рис. 160

Определение 2. Пусть функция двух переменных задана аналитически формулой $z=f(x, y)$. Множество всех точек $(x;\ y;\ z)$ пространства таких, что $(x;\ y)$ принадлежит области определения функции, а $z=f(x, y)$, называется *графиком функции* $z=f(x, y)$.

На практике построение графика функции $z=f(x, y)$ осложняется тем, что речь идет об изображении на плоском чертеже пространственной фигуры, а это далеко не всегда удается сделать.

Примеры. 2. Построить график функции $z=\sqrt{4-x^2-y^2}$.

○ Уравнение $z=\sqrt{4-x^2-y^2}$ можно преобразовать к виду $x^2+y^2+z^2=4$. Заметим, что расстояние от точки $O(0;\ 0;\ 0)$ до точки $M(x;\ y;\ z)$ вычисляется по формуле $OM=\sqrt{x^2+y^2+z^2}$ (это легко можно усмотреть из рис. 159). Значит, уравнение x^2+ $+y^2+z^2=4$ описывает множество всех таких точек $(x;\ y;\ z)$, которые удалены от точки $(0;\ 0;\ 0)$ на расстояние 2. Это сфера (шаровая поверхность) с центром в начале координат и радиусом 2 (рис. 160). Из уравнения $x^2+y^2+z^2=4$ находим, что либо $z=\sqrt{4-x^2-y^2}$, либо $z=-\sqrt{4-x^2-y^2}$. В первом случае речь идет о верхней половине сферы (это и есть искомый график), во втором — о ее нижней половине.●

3. Построить график функции $z=x^2+y^2-1$.

○ Здесь для построения графика применим *метод линий уровня*. Он заключается в следующем. Выше мы отмечали, что $z=c$ — уравнение плоскости, параллельной плоскости xy. Для построения графика функции $z=f(x, y)$ его рассекают плоскостями $z=c_1$, $z=c_2$, ..., $z=c_n$, в каждой из плоскостей строят сечения $f(x, y)=c_1$, $f(x, y)=$ $=c_2$, ..., $f(x, y)=c_n$ (линии уровня); совокупность этих сечений и дает представление о графике.

В данном примере положим:

$z=-1$; тогда $x^2+y^2-1=-1$, т. е. $x^2+y^2=0$ — это точка $(0;\ 0)$;

$z=0$; тогда $x^2+y^2-1=0$, т. е. $x^2+y^2=1$ — это окружность с центром в начале координат и радиусом 1;

$z=3$; тогда $x^2+y^2-1=3$, т. е. $x^2+y^2=4$ — это окружность с центром в начале координат и радиусом 2;

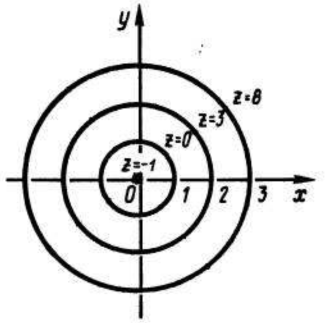


Рис. 161

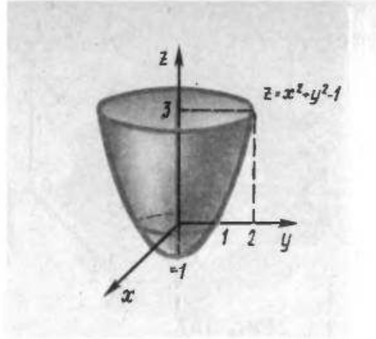


Рис.. 162

$z=8$; тогда $x^2+y-1=8$, т. е. $x^2+y^2=9$ — это окружность с центром в начале координат и радиусом 3; и т. д.

На рис. 161 изображено семейство найденных линий уровня; если теперь каждую из линий поместить в соответствующей плоскости, то получим изображение графика функции $z=x^2+y^2-1$ — это так называемый *параболоид вращения* (рис. 162).●

Определение 3. Функция $z=f(x, y)$, определенная на множестве D, называется *ограниченной сверху (снизу)*, если область значений функции есть множество, ограниченное сверху (снизу). Иными словами, если существует такое число $M(m)$, что для всех $(x; y)$ из D выполняется неравенство $f(x, y) \leqslant M$ $(f(x, y) \geqslant m)$. Функция, ограниченная и сверху, и снизу, называется *ограниченной*.

Например, функция $z=\sqrt{4-x^2-y^2}$ ограничена, так как $0 \leqslant \sqrt{4-x^2-y^2} \leqslant$ $\leqslant 2$ для всех $(x; y)$ из области определения функции (см. рис. 160). Функция $z=x^2+y^2-1$ ограничена снизу, поскольку $x^2+y^2-1 \geqslant -1$, но не ограничена сверху (см. рис. 162).

§ 44. Предел и непрерывность функции нескольких переменных

1. Предел функции нескольких переменных

Выше (см. § 16) мы сформулировали несколько эквивалентных друг другу определений предела функции одной переменной. Согласно одному из определений, равенство $\lim\limits_{x \to a} f(x)=b$ есть краткая запись того, что

$$(\forall \varepsilon>0)(\exists \delta>0)(\forall x: 0<|x-a|<\delta)|f(x)-b|<\varepsilon.$$

Но $|x-a|$ есть расстояние между точками x и a на координатной прямой; обозначив это расстояние $\rho(x, a)$, получим следующее определение: число b называется *пределом функции $y=f(x)$ при $x \to a$*, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что из неравенства $0<\rho(x, a)<\delta$ следует неравенство $|f(x)-$ $-b|<\varepsilon$.

Для функции нескольких переменных $u=f(x, y, \ldots, z)$ часто используют запись $u=f(P)$, где $P=P(x; y, \ldots; z)$. Это позволяет ввести определение предела функции $u=f(P)$ при $P \to P_0$, где $P_0=P_0(x_0; y_0; \ldots; z_0)$ дословно так же, как в случае функции одной переменной.

Определение 1. Число b называется *пределом функции нескольких переменных $u=f(P)$ при $P \to P_0$*, если для любого $\varepsilon>0$ существует $\delta>0$ такое, что из неравенства $0<\rho(P, P_0)<\delta$ следует неравенство $|f(P)-b|<\varepsilon$. Короче:

$$(\forall \varepsilon>0)(\exists \delta>0)(\forall P: 0<\rho(P, P_0)<\delta)|f(P)-b|<\varepsilon.$$

Пишут

$$\lim_{P \to P_0} f(P)=b \text{ или } \lim_{(x, y, \ldots, z) \to (x_0, y_0, \ldots, z_0)} (x, y, \ldots, z)=b \text{ или } \lim_{\substack{x \to x_0 \\ y \to y_0 \\ \cdots \\ z \to z_0}} f(x, y, \ldots, z)=b.$$

Напомним, что δ-окрестностью точки a на координатной прямой называется интервал $(a-\delta, a+\delta)$, т. е. множество таких точек прямой, которые удалены от точки a на расстояние, меньшее δ: $\rho(x, a)<\delta$. Аналогично определяется *δ-окрестность* точки $P_0(x_0; y_0)$ в координатной плоскости: это множество всех таких точек $P(x; y)$ для которых $\rho(P, P_0)<\delta$. Геометрически δ-окрестность точки $P_0(x_0; y_0)$ представляет собой внутренность круга (без его границы) с центром в точке $P_0(x_0; y_0)$ и радиусом δ (рис. 163); такой круг без границы называют *открытым*. В случае функции трех переменных $u=f(x, y, z)$ δ-окрестность точки $P_0(x_0; y_0; z_0)$ представляет собой открытый шар с центром в точке P_0 и радиусом δ. Если из δ-окрестности точки P_0 удалить саму точку P_0, то получится *проколотая δ-окрестность* точки P_0; точки проколотой δ-окрестности удовлетворяют двойному неравенству $0<\rho(P, P_0)<\delta$.

Теперь мы можем сформулировать определение 1 так:

Определение 2. Число b называется *пределом функции $f(P)$ при $P \to P_0$*, если для любого $\varepsilon>0$ можно указать такую проколотую δ-окрестность точки P_0, в которой выполняется неравенство $|f(P)-b|<\varepsilon$.

Отсюда, в частности, следует, что $\lim\limits_{(x, y) \to (x_0, y_0)} c=c$ (предел постоянной функции равен значению этой постоянной), $\lim\limits_{(x, y) \to (x_0, y_0)} x=x_0$, $\lim\limits_{(x, y) \to (x_0, y_0)} y=y_0$.

Определение 1 будем называть определением «на языке расстояний», а определение 2 — «на языке окрестностей».

Поскольку определения 1 и 2 дословно те же, что и соответствующие определения предела в точке для функции одной переменной, все теоремы, полученные логическим путем из определения предела функции одной переменной (см. теоремы 6—10 § 16) распространяются и на случай функции нескольких переменных. Приведем формулировки соответствующих теорем.

Теорема 1. *Если функция $f(P)$ имеет предел при $P \to P_0$, то только один.*

Теорема 2. *Если функция $f(P)$ имеет предел при $P \to P_0$, то она ограничена в некоторой окрестности точки P_0.*

Теорема 3. *Если $\lim\limits_{P \to P_0} f(P)=b$, $\lim\limits_{P \to P_0} g(P)=c$ и в некоторой проколотой окрестности точки P_0 выполняется неравенство $f(P) \leqslant g(P)$, то $b \leqslant c$.*

С л е д с т в и е. *Если $f(P) \geqslant 0$ ($f(P) \leqslant 0$) в некоторой проколотой окрестности точки P_0 и $\lim\limits_{P \to P_0} f(P)$ существует, то он неотрицателен (неположителен).*

Теорема 4. *Если $\lim\limits_{P \to P_0} f(P)=\lim\limits_{P \to P_0} h(P)=b$ и в некоторой проколотой окрестности точки P_0 справедливо неравенство $f(P) \leqslant g(P) \leqslant h(P)$, то и $\lim\limits_{P \to P_0} g(P)=c$.*

Теорема 5. Пусть $\lim\limits_{P\to P_0} f(P)=b$, $\lim\limits_{P\to P_0} g(P)=c$. *Тогда:*

1) $\lim\limits_{P\to P_0}(f(P)+g(P))=b+c$; 2) $\lim\limits_{P\to P_0} f(P)g(P)=bc$; 3) $\lim\limits_{P\to P_0}\dfrac{f(P)}{g(P)}=\dfrac{b}{c}$ *(если $c\neq 0$).*

Если $\lim\limits_{P\to P_0} f(P)=0$, то функция $u=f(P)$ называется *бесконечно малой* при $P\to P_0$.

Примеры. 1. Вычислить $\lim\limits_{(x,\,y)\to(2,\,1)}\dfrac{x^2+y^3}{2x-3y}$.

○ Используя теорему 5 об арифметических операциях над пределами, а также то, что предел постоянной равен значению этой постоянной и $\lim\limits_{(x,\,y)\to(2,\,1)} x=2$, $\lim\limits_{(x,\,y)\to(2,\,1)} y=1$, получим

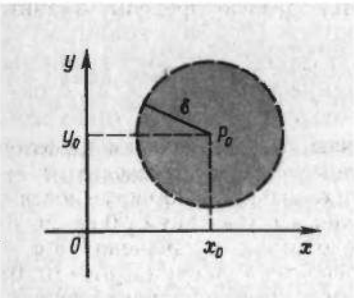


Рис. 163

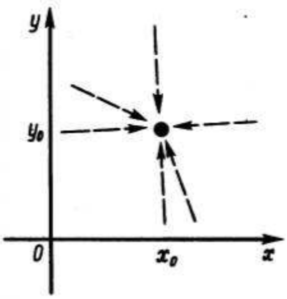


Рис. 164

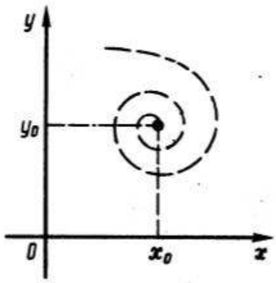


Рис. 165

$$\lim_{(x,y)\to(2,1)}\frac{x^2+y^3}{2x-3y}=\frac{\lim\limits_{(x,y)\to(2,1)}(x^2+y^3)}{\lim\limits_{(x,y)\to(2,1)}(2x-3y)}=\frac{\lim\limits_{(x,y)\to(2,1)} x\cdot x+\lim\limits_{(x,y)\to(2,1)} y\cdot y\cdot y}{\lim\limits_{(x,y)\to(2,1)} 2x-\lim\limits_{(x,y)\to(2,1)} 3y}=\frac{2^2+1^3}{2\cdot 2-3\cdot 1}=5. \bullet$$

2. Вычислить $\lim\limits_{(x,y)\to(0,0)}\dfrac{\ln(1+2xy)}{\sin 3xy}$.

○ Положим $t=xy$; тогда из $(x;\,y)\to(0;\,0)$ следует $t\to 0$ и заданный предел можно переписать в виде $\lim\limits_{t\to 0}\dfrac{\ln(1+2t)}{\sin 3t}$. Но при $t\to 0$ имеем $\ln(1+2t)\sim 2t$, а $\sin 3t\sim 3t$ (см. § 27); значит, $\lim\limits_{t\to 0}\dfrac{\ln(1+2t)}{\sin 3t}=\lim\limits_{t\to 0}\dfrac{2t}{3t}=\dfrac{2}{3}$. Итак, $\lim\limits_{(x,y)\to(0,0)}\dfrac{\ln(1+2xy)}{\sin 3xy}=\dfrac{2}{3}$. ●

Между понятиями предела в точке для функции одной переменной и нескольких переменных имеется, как мы отмечали выше, много общего, но между ними есть и принципиальное отличие, которое делает понятие предела функции нескольких переменных существенно более ограничительным, чем понятие предела функции одной переменной. Дело в следующем. Если $\lim\limits_{x\to a} f(x)=b$, то это означает, что и левосторонний, и правосторонний пределы равны b: $\lim\limits_{x\to a-0} f(x)=b$, $\lim\limits_{x\to a+0} f(x)=b$. Верно и обратное: из существования и совпадения двух односторонних пределов следует и существование предела функции в точке (см. п. 3 § 16). Для функции же двух переменных $z=f(x,y)$ приближаться к точке $(x_0;\,y_0)$ можно бесконечным множеством способов: и справа, и слева, и сверху, и снизу, и под углом 30° к оси x, и под углом 245° к оси x и т. д. (рис. 164). Более того, к точке $(x_0;\,y_0)$ можно приближаться не только по прямой, но и по более сложным траекториям (рис. 165). Ясно, что равенство $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = b$ верно тогда и только тогда, когда предел равен b при приближении к точке $(x_0; y_0)$ по любому направлению и даже по любой траектории. Это существенно более ограничительное требование, чем совпадение двух односторонних пределов в случае функции одной переменной.

Примеры. 3. Доказать, что $\lim\limits_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ не существует.

○ Будем приближаться к точке (0; 0) по прямым $y = kx$. Если $y = kx$, то

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{x\to 0} \frac{x\cdot kx}{x^2+k^2x^2} = \frac{k}{1+k^2}.$$

Замечаем, что значение предела зависит от углового коэффициента прямой: например, при $k=1$ (т. е. при движении по прямой $y=x$) предел равен 1/2; при $k=2$ (т. е. при движении по прямой $y=2x$) предел равен 2/5 и т. д. Итак, если приближаться к точке (0; 0) по различным направлениям, то получаются разные пределы; значит, $\lim\limits_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ не существует. ●

4. Вычислить $\lim\limits_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2}$.

○ Здесь используем прием, часто оказывающийся полезным при вычислении пределов функций двух переменных. Положим $x = \rho\cos\varphi$, $y = \rho\sin\varphi$, где $\rho > 0$. Поскольку $x^2 + y^2 = \rho^2$, это означает, что при фиксированном ρ точка $(x; y)$ лежит на окружности с центром в начале координат и радиусом ρ, а если φ пробегает все значения от 0 до 2π, то точка $(x; y)$ пробегает всю окружность. Давая ρ любые положительные значения, а φ — значения от 0 до 2π, мы можем попасть в любую точку плоскости. Условие $(x; y)\to(0; 0)$ равносильно условию $\rho\to 0$ (здесь отсутствие φ как раз говорит о любом направлении приближения к точке (0; 0)).

Имеем

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2} = \lim_{\rho\to 0} \frac{\rho^2\cos^2\varphi\cdot\rho^2\sin^2\varphi}{\rho^2\cos^2\varphi+\rho^2\sin^2\varphi} = \lim_{\rho\to 0} \frac{\rho^4\cos^2\varphi\sin^2\varphi}{\rho^2(\cos^2\varphi+\sin^2\varphi)} =$$

$$= \lim_{\rho\to 0}\rho^2\cos^2\varphi\sin^2\varphi = 0. \,●$$

2. Непрерывность функции в точке

Непрерывность в точке для функции нескольких переменных определяется так же, как и в случае функции одной переменной, а потому и локальные свойства непрерывной функции нескольких переменных аналогичны соответствующим свойствам функции одной переменной. Это позволяет нам строить изложение данного пункта по образцу п. 1 § 17.

Определение 3. Функция $f(P)$, определенная в некоторой окрестности точки P_0, называется *непрерывной в точке* P_0 если $\lim\limits_{P\to P_0} f(P) = f(P_0)$.

Использовав для предела функции определения 1 и 2, получим определения непрерывности функции в точке «на языке расстояний» и на «языке окрестностей».

Определение 4. Функция $f(P)$, определенная в некоторой окрестности точки P_0, называется *непрерывной в точке* P_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что из неравенства $\rho(P, P_0) < \delta$ следует неравенство $|f(P) - f(P_0)| < \varepsilon$.

Определение 5. Функция $f(P)$, определенная в некоторой окрестности точки P_0, называется *непрерывной в точке* P_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любой точки P из δ-окрестности точки P_0 выполняется неравенство $|f(P) - f(P_0)| < \varepsilon$.

Теорема 6. *Если функция непрерывна в точке, то она ограничена в некоторой окрестности этой точки.*

Теорема 7. *Если функция $f(P)$ непрерывна в точке P_0 и $f(P_0) > 0$ ($f(P_0) < 0$), то в некоторой окрестности точки P_0 выполняется неравенство $f(P) > 0$ ($f(P) < 0$)*

Теорема 8. *Пусть функции $f(P)$ и $g(P)$ непрерывны в точке P_0. Тогда их сумма, произведение и частное (в случае, если $g(P_0) \neq 0$) также непрерывны в точке P_0.*

3. Частные и полные приращения функций. Определение непрерывности «на языке приращений»

Пусть функция $z = f(x, y)$ определена в некоторой окрестности точки $P_0(x_0; y_0)$. Дадим независимым переменным x и y приращения Δx и Δy так, чтобы точка $P(x_0 + \Delta x; y_0 + \Delta y)$ не вышла за пределы указанной окрестности. Тогда и точки $K(x_0 + \Delta x; y_0)$, $M(x_0; y_0 + \Delta y)$ также окажутся в рассматриваемой окрестности (рис. 166).

Разность $f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ называют *полным приращением функции* $f(x, y)$ при переходе от точки $(x_0; y_0)$ к точке $(x_0 + \Delta x; y_0 + \Delta y)$ и обозначают Δz. Разность $f(x_0 + \Delta x, y_0) - f(x_0, y_0)$ называют *частным приращением по* x, а разность $f(x_0, y_0 + \Delta y) - f(x_0, y_0)$ — *частным приращением по* y функции $z = f(x, y)$; их обозначают соответственно $\Delta_x z$ и $\Delta_y z$.

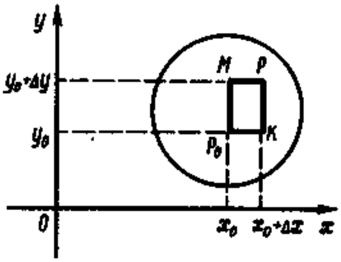


Рис. 166

Итак (см. рис. 166),

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = f(P) - f(P_0),$$
$$\Delta_x z = f(x_0 + \Delta x, y_0) - f(x_0, y_0) = f(K) - f(P_0),$$
$$\Delta_y z = f(x_0, y_0 + \Delta y) - f(x_0, y_0) = f(M) - f(P_0).$$

Аналогично определяются полное приращение Δu функции трех переменных $u = f(x, y, z)$ и частные приращения $\Delta_x u$, $\Delta_y u$, $\Delta_z u$.

Используя понятие полного приращения функции, можно дать следующее определение непрерывности функции в точке «на языке приращений».

Определение 6. Функция $z = f(P)$, определенная в некоторой окрестности точки P_0, называется *непрерывной в точке* P_0, если полное приращение функции Δz бесконечно мало при $P \to P_0$, т. е. при $\Delta x \to 0$ и $\Delta y \to 0$.

Ясно, что из $\Delta z \to 0$ следует $\Delta_x z \to 0$ и $\Delta_y z \to 0$. Обратное, вообще говоря, неверно. Рассмотрим например, функцию

$$z = f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{при } (x; y) \neq (0; 0), \\ 0 & \text{при } (x; y) = (0; 0) \end{cases}$$

и найдем для нее частные приращения в точке $(0; 0)$. Заметив, что $f(x, 0) = 0$ и $f(0, y) = 0$, найдем

$$\Delta_x z = f(x, 0) - f(0,0) = 0, \ \Delta_y z = f(0, y) - f(0,0) = 0.$$

Итак, и $\Delta_x z$, и $\Delta_y z$ — бесконечно малые при $(x; y) \to (0, 0)$. Однако отсюда не следует, что и $\Delta z \to 0$. В самом деле, если приближаться к точке $(0; 0)$ по прямой $y = x$, то $\lim\limits_{(x,y)\to(0,0)} \Delta z = \lim\limits_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \frac{1}{2}$; (см. пример 3 п. 1); при движении по прямой $y = 2x$ получим, что $\Delta z \to \frac{2}{5}$. Здесь $\lim\limits_{(x,y)\to(0,0)} \Delta z$ не существует.

4. Непрерывность сложной функции

Пусть функция $z = f(u, v)$ определена на множестве E, а переменные u и v, в свою очередь, зависят от переменных x и y: $u = u(x, y)$, $v = v(x, y)$, где обе функции $u(x, y)$, $v(x, y)$ определены на множестве D. Если для любого $(x, y) \in D$ значения $u = u(x, y)$ и $v = v(x, y)$ таковы, что $(u, v) \in E$, то говорят, что на множестве D определена *сложная функция* $z = f(u, v)$, где $u = u(x, y)$, $v = v(x, y)$; u, v — *промежуточные переменные*, x, y — *независимые переменные*.

Сложную функцию можно записать в виде $z = f(u(x, y), v(x, y))$. Например, $z = u^2 + v^3$, где $u = \sin(x + y)$, $v = \cos(x - y)$ — сложная функция, определенная на всей координатной плоскости. Ее можно переписать в виде $z = \sin^2(x + y) + \cos^3(x - y)$.

Количество промежуточных и независимых переменных может быть любым. Например, сложная функция $u = \sqrt{x^2 + y - z}$, где $x = t + 1$, $y = 5t$, $z = 1 - 4t$, содержит три промежуточные переменные x, y, z и одну независимую переменную t. Эту функцию можно переписать в следующем виде: $u = \sqrt{(t+1)^2 + 5t - 1 + 4t}$ или $u = \sqrt{t^2 + 11t}$. Область определения сложной функции задается неравенством $t^2 + 11t \geqslant 0$, откуда $t \leqslant -11$; $t \geqslant 0$.

Теорема 9 (о непрерывности сложной функции). *Пусть на множестве D определена сложная функция $z = f(u, v)$, где $u = u(x, y)$, $v = v(x, y)$, и пусть функции $u(x, y)$, $v(x, y)$ непрерывны в точке $(x_0; y_0) \in D$, а функция $f(u, v)$ непрерывна в точке $(u_0; v_0)$*, где $u_0 = u(x_0, y_0)$, $v_0 = v(x_0, y_0)$. *Тогда сложная функция $z = f(u(x, y), v(x, y))$ непрерывна в точке $(x_0; y_0)$.*

□ Для доказательства воспользуемся определением непрерывности «на языке расстояний» и тем, что расстояние между точками S и S_0, где $S = (u; v)$, $S_0 = (u_0; v_0)$, вычисляется по формуле

$$\rho(S, S_0) = \sqrt{(u - u_0)^2 + (v - v_0)^2}. \qquad (1)$$

Возьмем произвольное $\varepsilon > 0$. По условию функция $z = f(u, v)$ непрерывна в точке $(u_0; v_0)$. Это означает, что существует $\lambda > 0$ такое, что из неравенства $\rho(S, S_0) < \lambda$ следует неравенство $|f(S) - f(S_0)| < \varepsilon$.

Аналогично, так как по условию функции $u(x, y)$ и $v(x, y)$ непрерывны в точке $(x_0; y_0)$, то существуют такие $\delta_1 > 0$ и $\delta_2 > 0$, что из неравенств $\rho(P, P_0) < \delta_1$ и $\rho(P, P_0) < \delta_2$ следуют соответственно неравенства

$$|u(P) - u(P_0)| < \lambda/\sqrt{2} \qquad (2)$$

и

$$|v(P) - v(P_0)| < \lambda/\sqrt{2}. \qquad (3)$$

Здесь $P = (x; y)$, $P_0 = (x_0; y_0)$.

Выберем наименьшее из чисел δ_1 и δ_2 и обозначим его через δ. Тогда из $\rho(P, P_0) < \delta$ будут следовать оба неравенства (2) и (3):

$$\rho(P, P_0) < \delta \Rightarrow \begin{cases} |u(P) - u(P_0)| < \lambda/\sqrt{2}, \\ |v(P) - v(P_0)| < \lambda/\sqrt{2}. \end{cases}$$

Теперь в формуле (1) положим $S = (u(P); v(P))$, $S_0 = (u(P_0); v(P_0))$, и воспользуемся неравенствами (2) и (3). Тогда получим

$$\rho(S, S_0) = \sqrt{(u(P) - u(P_0))^2 + (v(P) - v(P_0))^2} < \sqrt{(\lambda/\sqrt{2})^2 + (\lambda/\sqrt{2})^2} = \lambda.$$

Итак, мы доказали, что если $\rho(P, P_0) < \delta$, то $\rho(S, S_0) < \lambda$, а тогда $|f(S) - f(S_0)| < \varepsilon$ или, что то же самое, $|f(u(P), v(P)) - f(u(P_0), v(P_0))| < \varepsilon$. Это и означает непрерывность сложной функции $z = f(u(x, y), v(x, y))$ в точке $P_0 = (x_0; y_0)$. ■

В § 26 был получен фундаментальный результат о непрерывности элементарных функций во всех точках, в которых они определены. Аналогичный результат имеет место и для функции нескольких переменных.

Примеры. 1. Исследовать на непрерывность функцию $u = \sqrt[3]{x^2 + y^3 - z^4}$.

○ Выражение $\sqrt[3]{x^2 + y^3 - z^4}$ определено в любой точке $(x; y; z)$; значит, заданная функция непрерывна в любой точке. ●

2. Исследовать на непрерывность функцию $z = \dfrac{2x + y}{x - y}$.

○ Выражение $\dfrac{2x + y}{x - y}$ не определено в таких точках $(x; y)$, для которых $y = x$. Следовательно, функция $z = \dfrac{2x + y}{x - y}$ имеет разрыв в точках, лежащих на прямой $y = x$, и непрерывна во всех остальных точках (см. рис. 154). ●

5. Замкнутые и открытые множества. Непрерывность функции на множестве

Пусть на плоскости (или в пространстве) дано некоторое множество точек D. Точка P_0 называется *внутренней точкой множества* D, если у этой точки существует окрестность, которая принадлежит D, и внешней, если существует окрестность точки P_0, ни одна точка которой не принадлежит D. Если же в любой окрестности точки P_0 имеются как точки, принадлежащие D, так и точки, не принадлежащие D, то такая точка P_0 называется *граничной точкой множества* D. Происхождение упомянутых терминов наглядно иллюстрирует рис. 167.

Ясно, что множество D содержит все свои внутренние точки. Что же касается граничных точек, то они могут принадлежать множеству D (например, это имеет место для круга $x^2 + y^2 \leqslant 4$), не принадлежать D (в случае открытого круга $x^2 + y^2 < 4$); может быть и так, что одни граничные точки принадлежат, а другие не принадлежат D (например, это имеет место для области определения функции $z = \sqrt{x} + \ln y$; здесь $x \geqslant 0$, $y > 0$, т. е. область определения — I четверть координатной плоскости с вертикальной границей, но без горизонтальной границы).

Определение 7. Если множество D содержит все свои граничные точки, то оно называется *замкнутым*. Если же множество D не содержит ни одной своей граничной точки, то оно называется *открытым* (т. е. открытое множество состоит из одних внутренних точек).

Определение 8. Если любые две точки множества D можно соединить ломаной, целиком лежащей в D, то множество D называется *связным*.

Примеры связных множеств: круг (см. рис. 155), кольцо (рис. 168).

Определение 9. Если все точки P множества D удовлетворяют неравенству $\rho(P, P_0) < M$, где $M > 0$, а P_0 — некоторая точка множества D, то D называется *ограниченным множеством*.

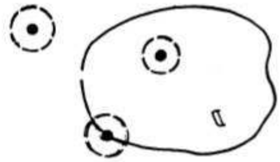


Рис. 167

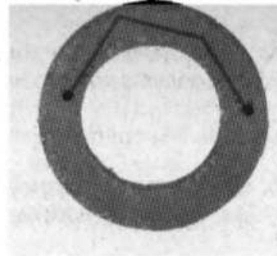


Рис. 168

Геометрически ограниченность множества D означает, что его можно целиком поместить внутри открытого круга (или шара — в случае пространственного множества) с центром в точке P_0 и радиусом M.

Подобно тому, как для функции одной переменной было введено понятие непрерывности на интервале или отрезке, так и для функции нескольких переменных можно ввести понятие непрерывности на открытом или замкнутом множестве. Однако сначала нужно условиться о том, что мы будем понимать под непрерывностью функции в граничной точке (подобно тому, как в п. 3 § 17 было введено понятие непрерывности функции одной переменной $y = f(x)$ на концах отрезка $[a, b]$, т. е. в граничных точках промежутка $[a, b]$).

Пусть P_0 — граничная точка множества D. Будем говорить, что функция $f(P)$ *непрерывна в граничной точке* P_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что из неравенства $\rho(P, P_0) < \delta$, где $P \in D$, следует неравенство $|f(P) - f(P_0)| < \varepsilon$.

Определение 10. Если функция $f(x, y)$ непрерывна в каждой точке множества D (внутренней или граничной), то функция называется *непрерывной на множестве* D.

Так, функция $u = \sqrt[3]{x^2 + y^3 - z^4}$ непрерывна на всем координатном пространстве, а функция $z = \dfrac{2x + y}{x - y}$ непрерывна в полуплоскости $y > x$ и в полуплоскости $y < x$.

6. Свойства функций, непрерывных на ограниченных замкнутых множествах

Свойства функций нескольких переменных, непрерывных на ограниченных замкнутых множествах, аналогичны свойствам непрерывной функции одной переменной на отрезке, о которых шла речь в § 18. Здесь мы сформулируем соответствующие теоремы.

Теорема 10. *Если функция $f(x, y)$ непрерывна на замкнутом ограниченном множестве, то она ограничена на этом множестве.*

Доказательство аналогично доказательству теоремы 4 § 18, но вместо последовательности стягивающихся отрезков на прямой нужно рассматривать последовательность стягивающихся прямоугольников на плоскости.

Теорема 11. *Если функция $f(x, y)$ непрерывна на замкнутом ограниченном множестве, то среди ее значений на этом множестве имеются как наименьшее, так и наибольшее.*

Доказательство почти дословно повторяет доказательство теоремы 5 § 18 (рекомендуем провести его самостоятельно).

7. Свойства функций, непрерывных на связных множествах

Теорема 12 (о нуле непрерывной функции). *Пусть функция $f(x, y)$ непрерывна на связном множестве D и принимает в двух точках A и B этого множества значения разных знаков. Тогда в множестве D найдется такая точка, что в ней функция обращается в нуль.*

□ Положим для определенности $f(A)<0$, $f(B)>0$.

Сначала предположим, что точки A и B можно соединить горизонтальным отрезком, целиком принадлежащим D. Рассмотрим функцию $z=f(x, y)$ на отрезке AB. Здесь $y=y_0$, т. е. мы имеем непрерывную функцию одной переменной $z=f(x, y_0)$, которая в точке x_0 (т. е. в точке $A(x_0; y_0)$) отрицательна, а в точке x_1 (т. е. в точке $B(x_1; y_0)$) положительна. Тогда по теореме 1 § 18 на отрезке $[x_0, x_1]$ (т. е. на отрезке AB) найдется точка c (т. е. точка $C(c; y_0)$), в которой функция обращается в нуль.

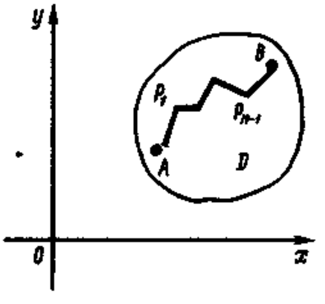


Рис. 169

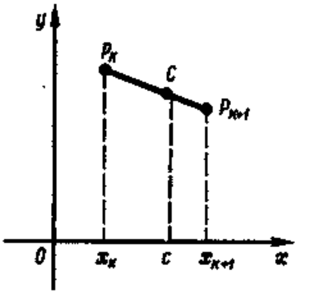


Рис. 170

Аналогично обстоит дело в случае, когда точки A и B можно соединить вертикальным отрезком, целиком принадлежащим D.

Пусть точки A и B нельзя соединить одним горизонтальным или вертикальным отрезком; тогда в силу связности множества D их можно соединить ломаной $AP_1P_2...P_{n-1}B$ (рис. 169). Если в одной из вершин ломаной функция обращается в нуль, то теорема доказана. Если же ни в одной вершине функция не обращается в нуль, то найдется звено ломаной, т. е. отрезок P_kP_{k+1}, такой, что $f(P_k)<0$, а $f(P_{k+1})>0$. Для случая, когда этот отрезок горизонтален или вертикален, теорема доказана.

Пусть отрезок P_kP_{k+1} не горизонтален и не вертикален. Тогда его можно задать уравнением $y=ax+b$, где $a\neq 0$, $x_k\leqslant x\leqslant x_{k+1}$ (рис. 170). Функция $z=f(x, y)$ примет вид $z=f(x, ax+b)$. Эта сложная функция от x непрерывна на отрезке $[x_k, x_{k+1}]$ и на концах его (т. е. в точках P_k, P_{k+1}) принимает значения разных знаков. Следовательно, внутри отрезка найдется точка c, в которой функция обратится в нуль: $f(c, ac+b)=0$. Итак, функция $z=f(x, y)$ обращается в нуль в точке $C(c, ac+b)$, лежащей на отрезке P_kP_{k+1}, т. е. принадлежащей множеству D. ■

Теорема 13 (о промежуточном значении). *Пусть функция $f(x, y)$ непрерывна на связном множестве D и в двух точках A и B этого множества принимает неравные значения $f(A)$ и $f(B)$. Тогда на этом множестве она принимает любое значение μ, лежащее между $f(A)$ и $f(B)$, т. е. существует такая точка $C \in D$, что $f(C) = \mu$.*

Доказательство аналогично доказательству теоремы 2 § 18 (рекомендуем провести его самостоятельно).

§ 45. Дифференцируемость функции нескольких переменных, частные производные, дифференциал

1. Дифференцируемость функции нескольких переменных

Дифференцируемость в точке определяется по сути дела так же, как и для функции одной переменной (см. § 29).

Определение 1. Функция $z = f(x, y)$ называется *дифференцируемой в точке* $(x_0;\ y_0)$, если ее полное приращение Δz можно представить в виде

$$\Delta z = A\Delta x + B\Delta y + \alpha\Delta x + \beta\Delta y, \tag{1}$$

где A, B — числа, а α, β — бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$. Главная линейная часть приращения функции, т. е. $A\Delta x + B\Delta y$, называется *дифференциалом* функции $f(x, y)$ в точке $(x_0;\ y_0)$ и обозначается dz: $dz = A\Delta x + B\Delta y$.

Заметим, что выражение $\alpha\Delta x + \beta\Delta y$, где $\alpha \to 0$, $\beta \to 0$ при $\Delta x \to 0$, $\Delta y \to 0$, можно записать в виде $\gamma\sqrt{(\Delta x)^2 + (\Delta y)^2}$, где $\gamma \to 0$ при $\Delta x \to 0$, $\Delta y \to 0$. В самом деле, имеем

$$\alpha\Delta x + \beta\Delta y = \sqrt{\Delta x^2 + \Delta y^2}\,\frac{\alpha\Delta x + \beta\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}}.$$

Так как $\left|\frac{\Delta x}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}\right| \leqslant 1$ и $\left|\frac{\Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}\right| \leqslant 1$, то $\alpha\frac{\Delta x}{\sqrt{(\Delta x)^2 + (\Delta y)^2}} + \beta\frac{\Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}$ есть величина бесконечно малая при $\Delta x \to 0$, $\Delta y \to 0$; обозначим ее γ. Тогда получаем $\alpha\Delta x + \beta\Delta y = \gamma\sqrt{(\Delta x)^2 + (\Delta y)^2}$. Это позволяет сформулировать определение дифференцируемости следующим образом:

Определение 2. Функция $z = f(x, y)$ называется *дифференцируемой в точке* $(x_0;\ y_0)$, если ее полное приращение Δz можно представить в виде

$$\Delta z = A\Delta x + B\Delta y + \gamma\sqrt{(\Delta x)^2 + (\Delta y)^2},$$

где A, B — числа, а γ — бесконечно малая при $\Delta x \to 0$, $\Delta y \to 0$.

Теорема 1. *Если функция $z = f(x, y)$ дифференцируема в точке $(x_0;\ y_0)$, то существуют пределы $\lim\limits_{\Delta x \to 0}\frac{\Delta_x z}{\Delta x}$ и $\lim\limits_{\Delta y \to 0}\frac{\Delta_y z}{\Delta y}$.*

□ Чтобы из Δz получить $\Delta_x z$, достаточно положить $\Delta y = 0$. Из равенства (1) следует, что если $\Delta y = 0$, то $\Delta_x z = A\Delta x + \alpha\Delta x$, откуда $\lim\limits_{\Delta x \to 0}\frac{\Delta_x z}{\Delta x} = A$.

Аналогично $\lim\limits_{\Delta y \to 0}\frac{\Delta_y z}{\Delta y} = B$. ■

Определение 3. Пусть функция $z = f(x, y)$ определена в точке $(x_0; y_0)$ и в некоторой ее окрестности. Если существует $\lim\limits_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x}$, то он называется *частной производной по x* функции $z = f(x, y)$ в точке $(x_0; y_0)$ и обозначается $\frac{\partial z}{\partial x}$ или z'_x, или $f'_x(x_0, y_0)$. Аналогично, $\lim\limits_{\Delta y \to 0} \frac{\Delta_y z}{\Delta y}$ называется *частной производной по y* функции $z = f(x, y)$ в точке $(x_0; y_0)$ и обозначается $\frac{\partial z}{\partial y}$ или z'_y, или $f'_y(x_0, y_0)$.

Итак, $\lim\limits_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x} = \frac{\partial z}{\partial x}$, $\lim\limits_{\Delta y \to 0} \frac{\Delta_y z}{\Delta y} = \frac{\partial z}{\partial y}$.

Теперь мы можем сформулировать теорему 1 следующим образом:

Теорема 2 (необходимое условие дифференцируемости функции в точке). *Если функция $z = f(x, y)$ дифференцируема в точке $(x_0; y_0)$, то в этой точке существуют частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$.*

Для дифференциала имеем $dz = A\Delta x + B\Delta y = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$. Обычно приращения независимых переменных Δx и Δy записывают в виде dx и dy. Тогда формула для дифференциала примет вид

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy. \qquad (2)$$

Теорема 3. *Если функция $z = f(x, y)$ дифференцируема в точке $(x_0; y_0)$, то она непрерывна в этой точке.*

□ Из равенства (1) следует, что если $\Delta x \to 0$, $\Delta y \to 0$, то $\Delta z \to 0$, а это и означает непрерывность функции $z = f(x, y)$ в точке $(x_0; y_0)$ (см. определение 6 из § 44). ■

Обратное утверждение неверно. Так, функция $z = \sqrt{x^2 + y^2}$ непрерывна в точке (0; 0), но в этой точке не имеет частных производных (в самом деле, если $y = 0$, то $z = \sqrt{x^2}$, т. е. $z = |x|$, а эта функция не имеет производной в точке $x = 0$; аналогично, если $x = 0$, то $z = \sqrt{y^2}$, т. е. $z = |y|$, а эта функция также не имеет производной в точке $y = 0$). Следовательно, в силу теоремы 2 функция $z = \sqrt{x^2 + y^2}$ не может быть дифференцируемой в точке (0; 0).

Аналогичные утверждения о связи между понятиями дифференцируемости и непрерывности в точке имеют место для функции одной переменной (см. § 29).

Из определений частных производных следует, что они находятся по тем же правилам, что и производные функции одной переменной. Нужно лишь помнить, что при нахождении $\frac{\partial z}{\partial x}$ мы считаем y не переменной, а постоянной $\left(\text{поскольку нас интересует лишь } \Delta_x z, \text{ а точнее } \lim\limits_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x}\right)$, а при нахождении $\frac{\partial z}{\partial y}$ мы считаем x не переменной, а постоянной.

Примеры. 1. Найти $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ для функции $z = x^3 y + \sin(x^2 + \sqrt{y}) + \operatorname{tg} x + \ln y$.

○ Найдем $\frac{\partial z}{\partial x}$. Считая, что $\Delta y = 0$, т. е. $y = \text{const}$, получаем

$$\frac{\partial z}{\partial x} = 3x^2 y + \cos(x^2 + \sqrt{y})2x + \frac{1}{\cos^2 x}.$$

При нахождении $\frac{\partial z}{\partial y}$ считаем, что $\Delta x=0$, т. е. $x=\text{const}$. Получаем

$$\frac{\partial z}{\partial y}=x^3+\cos(x^2+\sqrt{y})\frac{1}{2\sqrt{y}}+\frac{1}{y}. \bullet$$

2. Найти dz для функции $z=\frac{x}{y}$ в точке (2; 1), если $\Delta x=0{,}1$, $\Delta y=-0{,}2$.

○ Имеем $\frac{\partial z}{\partial x}=\frac{1}{y}$, $\frac{\partial z}{\partial y}=-\frac{x}{y^2}$. В точке (2; 1) получаем $\frac{\partial z}{\partial x}=1$, $\frac{\partial z}{\partial y}=-2$. Так как $dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy$, а по условию $dx=\Delta x=0{,}1$, $dy=\Delta y=-0{,}2$, то $dz=1\cdot 0{,}1-2(-0{,}2)=0{,}5$. ●

3. Найти функцию $z=f(x, y)$ такую, что $\frac{\partial z}{\partial x}=6x^5y^2+\frac{1}{x}$, $\frac{\partial z}{\partial y}=2x^6y+\cos y$.

○ Так как $\frac{\partial z}{\partial x}$ — это производная искомой функции по переменной x при условии $y=\text{const}$, то $z=f(x, y)$ — первообразная для $\frac{\partial z}{\partial x}$ с точностью до постоянного слагаемого, зависящего от y. Имеем

$$z=\int\frac{\partial z}{\partial x}dx=\int\left(6x^5y^2+\frac{1}{x}\right)dx=6y^2\frac{x^6}{6}+\ln|x|+C(y)=x^6y^2+\ln|x|+C(y).$$

Тогда $\frac{\partial z}{\partial y}=2x^6y+C'(y)$. Но по условию $\frac{\partial z}{\partial y}=2x^6y+\cos y$. Значит, $C'(y)=\cos y$, т. е. $C(y)=\sin y$.

Итак, $z=x^6y^2+\ln|x|+\sin y$. ●

Если функцию $z=f(x, y)$, имеющую частные производные в точке $(x_0; y_0)$, рассматривать при условии $y=y_0$, то геометрически это означает, что поверхность $z=f(x, y)$ пересекается плоскостью $y=y_0$, параллельной координатной плоскости xz; в сечении получается плоская линия. Тогда $f'_x(x_0, y_0)$ *представляет собой угловой коэффициент касательной к указанному сечению в точке* $(x_0; y_0)$, *т. е. тангенс угла наклона этой касательной к положительному направлению оси* x. В этом состоит геометрический смысл частной производной $f'_x(x_0, y_0)$. Аналогично, $f'_y(x_0, y_0)$ есть угловой коэффициент проходящей через точку $(x_0; y_0)$ касательной к плоской кривой, полученной в результате пересечения поверхности $z=f(x, y)$ с плоскостью $x=x_0$.

Выше мы говорили о дифференцируемости, дифференциале, частных производных функции двух переменных. Аналогично определяются соответствующие понятия и для функции любого числа переменных. Например, дифференцируемость функции $u=f(x, y, z)$ в точке $(x_0; y_0; z_0)$ означает возможность представления полного приращения функции в виде

$$\Delta u=A\Delta x+B\Delta y+C\Delta z+\alpha\Delta x+\beta\Delta y+\gamma\Delta z.$$

Здесь α, β, γ — бесконечно малые при $\Delta x\to 0$, $\Delta y\to 0$, $\Delta z\to 0$; A, B, C — числа; как и в случае двух переменных, эти числа равны частным производным функции $u=f(x, y, z)$ в точке $(x_0; y_0; z_0)$: $A=\frac{\partial u}{\partial x}$, $B=\frac{\partial u}{\partial y}$, $C=\frac{\partial u}{\partial z}$. Дифференциал функции $u=f(x, y, z)$ имеет вид

$$du=\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy+\frac{\partial u}{\partial z}dz.$$

Пример 4. Найти $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ и $\frac{\partial u}{\partial z}$ для функции $u = x^2y^3z^4 + e^x + 2^y + \operatorname{arctg} z$.

○ При нахождении $\frac{\partial u}{\partial x}$ считаем, что $y = \text{const}$, $z = \text{const}$; получаем $\frac{\partial u}{\partial x} = 2xy^3z^4 + e^x$. Аналогично находим $\frac{\partial u}{\partial y} = 3x^2y^2z^4 + 2^y \ln 2$, $\frac{\partial u}{\partial z} = 4x^2y^3z^3 + \frac{1}{1+z^2}$. ●

2. Достаточное условие дифференцируемости функции нескольких переменных в точке

Для функции одной переменной дифференцируемость и существование производной являются равнозначными утверждениями. В случае функции нескольких переменных дело обстоит по-другому: существование частных производных является необходимым условием дифференцируемости функции в точке (см. теорему 2), но не является достаточным условием дифференцируемости. Так, для функции

$$z = f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{при } (x;\ y) \neq (0;\ 0), \\ 0 & \text{при } (x;\ y) = (0;\ 0) \end{cases}$$

в точке $(0;\ 0)$ и $\Delta_x z = 0$, и $\Delta_y z = 0$, а значит, $\frac{\partial z}{\partial x} = 0$, $\frac{\partial z}{\partial y} = 0$. Однако эта функция разрывна в точке $(0;\ 0)$, а потому дифференцируемой в ней быть не может (см. п. 3 § 44). Итак, для дифференцируемости функции $z = f(x, y)$ в точке $(x_0;\ y_0)$ недостаточно только существования частных производных; дифференцируемость имеет место, если дополнительно потребовать непрерывность частных производных, как это следует из приводимой ниже теоремы.

Теорема 4 (достаточное условие дифференцируемости). *Если функция $z = f(x, y)$ в некоторой окрестности точки $P_0(x_0;\ y_0)$ имеет непрерывные частные производные, то она дифференцируема в точке $(x_0;\ y_0)$.*

□ Дадим переменным x и y приращения Δx и Δy так, чтобы точка $P(x_0 + \Delta x;\ y_0 + \Delta y)$ не вышла за пределы той окрестности, о которой говорится в условии теоремы. Нужно доказать, что Δz, т. е. $f(P) - f(P_0)$, можно представить в виде (1).

Представим приращение функции в виде

$$\Delta z = f(P) - f(P_0) = (f(P) - f(P_1)) + (f(P_1) - f(P_0)), \tag{3}$$

где $P_1 = P_1(x_0 + \Delta x;\ y)$ (рис. 171). Рассмотрим разность $f(P_1) - f(P_0)$, т. е. $f(x_0 + \Delta x, y_0) - f(x_0, y_0)$. Это разность двух значений функции одной переменной $f(x, y_0)$ на концах отрезка $[x_0, x_0 + \Delta x]$. По условию в любой точке этого отрезка функция $f(x, y_0)$ имеет производную, а, значит, и непрерывна; поэтому к ней на этом отрезке можно применить теорму Лагранжа (см. теорему 4 § 33):

$$f(x_0 + \Delta x, y_0) - f(x_0, y_0) = f'_x(c_1, y_0)\Delta x,$$

где $c_1 \in [x_0, x_0 + \Delta x]$. Таким образом,

$$f(P_1) - f(P_0) = f'_x(S)\Delta x, \tag{4}$$

где $S = S(c_1;\ y_0)$ (рис. 171).

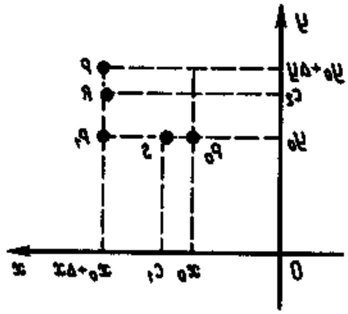


Рис. 171

Рассмотрим разность $f(P)-f(P_1)$, т. е. $f(x_0+\Delta x;\ y_0+\Delta y)-f(x_0+\Delta x, y_0)$. Это разность двух значений функции одной переменной $f(x_0+\Delta x, y)$ на концах отрезка $[\Delta y_0, y_0+\Delta y]$. Используя, как и на предыдущем шаге, теорему Лагранжа, получим $f(x_0+\Delta x, y_0+\Delta y)-f(x_0+\Delta x, y_0)=f'_y(x_0+\Delta x, c_2)\Delta y$, где $c_2 \in [y_0, y_0+\Delta y]$. Следовательно,

$$f(P)-f(P)=f'_y(R)\cdot\Delta y, \tag{5}$$

где $R=R(x_0+\Delta x,\ c_2)$ (рис. 171).

Подставив выражения (4) и (5) в равенство (3), получим

$$\Delta z=f'_x(S)\Delta x+f'_y(R)\Delta y. \tag{6}$$

Устремим теперь Δx и Δy к нулю. В силу непрерывности частных производных имеем $\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_x(S)=f'_x(x_0, y_0)$ и $\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}} f'_y(R)=f'_y(x_0, y_0)$. Это означает, что $f'_x(S)=f'_x(x_0, y_0)+\alpha$, $f'_y(R)=f'_y(x_0, y_0)+\beta$, где $\alpha\to 0$, $\beta\to 0$ при $\Delta x\to 0$, $\Delta y\to 0$. Полагая $A=f'_x(x_0, y_0)$, $B=f'_y(x_0, y_0)$, получим $f'_x(S)=A+\alpha$, $f'_y(R)=B+\beta$. Подставив эти выражения в соотношение (6), приходим к равенству

$$\Delta z=(A+\alpha)\Delta x+(B+\beta)\Delta y=A\Delta x+B\Delta y+\alpha\Delta x+\beta\Delta y,$$

где A, B — числа, а α, β — бесконечно малые при $\Delta x\to 0$, $\Delta y\to 0$, что и означает дифференцируемость функции $z=f(x, y)$ в точке $(x_0; y_0)$. ■

3. Использование дифференциала в приближенных вычислениях

Если функция $z=f(x, y)$ дифференцируема в точке $(x_0; y_0)$, т. е. выполняется равенство (1), то $\Delta z\approx dz$, где $dz=A\Delta x+B\Delta y$. Приближенное равенство $\Delta z\approx dz$ выражает основную роль дифференциала в математическом анализе и его приложениях.

Итак,

$$f(x_0+\Delta x\ , y_0+\Delta y)-f(x_0, y_0)\approx f'_x(x_0, y_0)\Delta x+f'_y(x_0, y_0)\Delta y.$$

Полагая в этом приближенном равенстве $a=x_0+\Delta x$, $b=y_0+\Delta y$, получим

$$f(a, b)\approx f(x_0, y_0)+f'_x(x_0, y_0)(a-x_0)+f'_y(x_0, y_0)(b-y_0). \tag{7}$$

На этой формуле основан следующий алгоритм.

АЛГОРИТМ
использования дифференциала в приближенных вычислениях

Чтобы найти приближенное значение некоторой величины A, нужно:

1^0. Представить A в виде значения некоторой функции в точке $(a;\ b)$: $A=f(a, b)$

2^0. Подобрать x_0, y_0 так, чтобы точка $(x_0;\ y_0)$ была достаточно близка к точке $(a;\ b)$ и чтобы значение $f(x_0, y_0)$ вычислялось легко.

3^0. Вычислить $f(x_0, y_0)$.

4^0. Для функции $z=f(x, y)$ найти $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $f'_x(x_0, y_0)$, $f'_y(x_0, y_0)$.

5^0. Подставить найденные значения a, b, x_0, y_0, $f(x_0, y_0)$, $f'_x(x_0, y_0)$, $f'_y(x_0, y_0)$ в формулу (7).

Рекомендуем сравнить этот алгоритм с аналогичным алгоритмом для функции одной переменной (см. п. 3 § 29).

Пример. Найти приближенное значение величины $\sqrt{1{,}03^2+1{,}98^3}$.

○ 1^0. Здесь $A=\sqrt{1{,}03^2+1{,}98^3}$; положим $f(x, y)=\sqrt{x^2+y^3}$, тогда $A=f(a, b)$, где $a=1{,}03$, $b=1{,}98$.

2^0. Положим $x_0=1$, $y_0=2$.

3^0. Имеем $f(x_0, y_0)=\sqrt{1^2+2^3}=3$.

4^0. Имеем $\frac{\partial z}{\partial x}=\frac{1}{2}(x^2+y^3)^{-\frac{1}{2}}\cdot 2x=\frac{x}{\sqrt{x^2+y^3}}$, $\frac{\partial z}{\partial y}=\frac{1}{2}(x^2+y^3)^{-\frac{1}{2}}\cdot 3y^2=\frac{3y^2}{2\sqrt{x^2+y^3}}$.

$f'_x(x_0, y_0)=f'_x(1, 2)=\frac{1}{3}$; $f'_y(x_0, y_0)=f'_y(1, 2)=2$.

5^0. По формуле (7) находим $f(a,\ b)\approx 3+\frac{1}{3}(1{,}03-1)+2(1{,}98-2)=2{,}97$, т. е. $\sqrt{1{,}03^2+1{,}98^3}\approx 2{,}97$. ●

4. Касательная плоскость

Если в равенстве (7) писать (x, y) вместо (a, b), то получим

$$f(x, y)\approx f(x_0, y_0)+f'_x(x_0, y_0)(x-x_0)+f'_y(x_0, y_0)(y-y_0). \tag{8}$$

Рассмотрим уравнение

$$z=f(x_0y_0)+f'_x(x_0, y_0)(x-x_0)+f'_y(x_0, y_0)(y-y_0). \tag{9}$$

Как известно из курса геометрии, это уравнение плоскости, проходящей через точку $(x_0;\ y_0;\ z_0)$, где $z_0=f(x_0, y_0)$. Геометрический смысл равенства (8) состоит в следующем: поверхность $z=f(x, y)$ в некоторой достаточно малой окрестности точки $(x_0;\ y_0)$ почти неотличима (при условии дифференцируемости функции $z=f(x, y)$ в этой точке) от плоскости, заданной уравнением (9). Плоскость, заданная уравнением (9), называется *касательной плоскостью* к поверхности $z=f(x, y)$ в точке $(x_0;\ y_0)$.

Итак, при условии дифференцируемости функции $f(x, y)$ в точке $(x_0;\ y_0)$ касательная плоскость к поверхности $z=f(x, y)$ в этой точке существует и определяется уравнением

$$z-z_0=A(x-x_0)+B(y-y_0),$$

где $z_0=f(x_0, y_0)$, $A=f'_x(x_0, y_0)$, $B=f'_y(x_0, y_0)$.

Заметим, что поверхность в пространстве чаще задается уравнением вида $F(x, y, z)=0$ (например, $x^2+y^2+z^2-16=0$ — сфера с центром в начале координат и радиусом 4). Тогда уравнение касательной плоскости к поверхности $F(x, y, z)=0$ в точке $(x_0, y_0; z_0)$ имеет вид

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0,$$

где $A=F'_x(x_0, y_0, z_0)$, $B=F'_y(x_0, y_0, z_0)$, $C=F'_z(x_0, y_0; z_0)$.

Пример. Составить уравнение касательной плоскости и нормали к сфере x^2+y^2+ $+z^2-16=0$ в точке $(2; 2; 2\sqrt{2})$.

○ Имеем $(x_0; y_0; z_0)=(2; 2; 2\sqrt{2})$, $F(x, y, z)=x^2+y^2+z^2-16$; $F'_x=2x$, $A=F'_x(2, 2, 2\sqrt{2})=4$; $F'_y=2y$, $B=F'_y(2, 2, 2\sqrt{2})=4$; $F'_z=2z$, $C=F'_z(2, 2, 2\sqrt{2})=4\sqrt{2}$. Значит, уравнение касательной плоскости имеет вид

$$4(x-2)+4(y-2)+4\sqrt{2}(z-2\sqrt{2})=0, \text{ или } x+y+z\sqrt{2}=8.$$

Нормаль к поверхности в точке $(x_0; y_0; z_0)$ — это прямая, проходящая через точку $(x_0; y_0; z_0)$ и перпендикулярная касательной плоскости. Значит, $(F'_x; F'_y; F'_z)$ — направляющий вектор нормали и уравнение нормали в общем виде таково: $(x-x_0)/F'_x=$ $=(y-y_0)/F'_y=(z-z_0)/F'_z$. В данном случае уравнение нормали имеет вид

$$\frac{x-2}{4}=\frac{y-2}{4}=\frac{z-2\sqrt{2}}{4\sqrt{2}}, \text{ или } x-2=y-2=\frac{z-2\sqrt{2}}{\sqrt{2}}. \bullet$$

5. Дифференцирование сложной функции

Теорема 5. *Пусть на множестве D определена сложная функция $z=f(u, v)$, где $u=u(x, y)$, $v=v(x, y)$ и пусть функции $u(x, y)$, $v(x, y)$ имеют в некоторой окрестности точки $(x_0; y_0)\in D$ непрерывные частные производные, а функция $f(u, v)$ имеет непрерывные частные производные в некоторой окрестности точки $(u_0; v_0)$, где $u_0=u(x_0, y_0)$, $v_0=v(x_0, y_0)$. Тогда сложная функция $z=f(u(x, y), v(x, y))$ дифференцируема в точке (x_0, y_0) причем*

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial x}, \quad \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y}.$$

□ По условию функции $u(x, y)$, $v(x, y)$ имеют непрерывные частные производные в некоторой окрестности точки $(x_0; y_0)$; значит, в силу теоремы 4 они дифференцируемы в точке $(x_0; y_0)$, т. е.

$$\Delta u=A_1\Delta x+B_1\Delta y+\alpha_1\Delta x+\beta_1\Delta y, \quad \Delta v=A_2\Delta x+B_2\Delta y+\alpha_2\Delta x+\beta_2\Delta y.$$

Если, в частности, дать приращение только аргументу x, то получим

$$\Delta_x u=A_1\Delta x+\alpha_1\Delta x, \quad \Delta_x v=A_2\Delta x+\alpha_2\Delta x, \tag{10}$$

где $A_1=\frac{\partial u}{\partial x}$, $A_2=\frac{\partial v}{\partial x}$, а α_1, α_2 — бесконечно малые при $\Delta x\to 0$.

Приращения $\Delta_x u$, $\Delta_x v$ приводят, в свою очередь, к приращению функции $z=f(u, v)$, которое обозначим $\Delta_x z$. Так как по условию функция $f(u, v)$ имеет непрерывные частные производные в некоторой окрестности точки $(u_0; v_0)$, то она дифференцируема в этой точке, т. е.

$$\Delta_x z=A\Delta_x u+B\Delta_x v+\alpha\Delta_x u+\beta\Delta_x v, \tag{11}$$

где $A=\frac{\partial z}{\partial u}$, $B=\frac{\partial z}{\partial v}$, а α, β — бесконечно малые при $\Delta_x u\to 0$, $\Delta_x v\to 0$.

Подставив выражения (10) в равенство (11), получим

$$\Delta_x z=A(A_1\Delta x+\alpha_1\Delta x)+B(A_2\Delta x+\alpha_2\Delta x)+\alpha(A_1\Delta x+\alpha_1\Delta x)+\beta(A_2\Delta x+\alpha_2\Delta x)=$$
$$=(AA_1+BA_2)\Delta x+(A\alpha_1+B\alpha_2+\alpha A_1+\alpha\alpha_1+\beta A_2+\beta\alpha_2)\Delta x=(AA_1+BA_2)\Delta x+\gamma\Delta x,$$

где $\gamma=A\alpha_1+B\alpha_2+\alpha A_1+\alpha\alpha_1+\beta A_2+\beta\alpha_2$ — бесконечно малая при $\Delta x\to 0$.

Тогда $\lim\limits_{\Delta x\to 0}\frac{\Delta_x z}{\Delta x}=AA_1+BA_2$, т. е. $\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\cdot\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial x}$.

Аналогично доказывается, что $\frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\cdot\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial y}$.

Полученные выражения показывают, что $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ непрерывны в точке $(x_0;\ y_0)$, а потому сложная функция $z=f(u(x,\ y),\ v(x,\ y))$ дифференцируема в этой точке. ■

Примеры. 1. Найти $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ для функции $z=u^2+v^3$, где $u=\sqrt[3]{xy}$, $v=\sqrt[5]{\frac{x}{y}}$.

○ Имеем $\frac{\partial z}{\partial u}=2u$, $\frac{\partial z}{\partial v}=3v^2$, $\frac{\partial u}{\partial x}=\sqrt[3]{y}\frac{1}{3\sqrt[3]{x^2}}$, $\frac{\partial u}{\partial y}=\sqrt[3]{x}\frac{1}{3\sqrt[3]{y^2}}$, $\frac{\partial v}{\partial x}=\frac{1}{\sqrt[5]{y}}\cdot\frac{1}{5\sqrt[5]{x^4}}$, $\frac{\partial v}{\partial y}=\sqrt[5]{x}\left(y^{-\frac{1}{5}}\right)'=-\frac{1}{5}\sqrt[5]{x}\frac{1}{\sqrt[5]{y^6}}$. Значит,

$$\frac{\partial z}{\partial x}=\frac{2u}{3}\cdot\sqrt[3]{\frac{y}{x^2}}+3v^2\frac{1}{5\sqrt[5]{x^4y}},\quad \frac{\partial z}{\partial y}=\frac{2u}{3}\cdot\sqrt[3]{\frac{x}{y^2}}-\frac{3v^2}{5}\sqrt[5]{\frac{x}{y^6}}.$$

Остается подставить в полученные формулы выражения для u и v через x и y. ●

2. Найти $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ для функции $z=\ln t$, где $t=\sin x+\cos y$.

○ Имеем $\frac{\partial z}{\partial x}=\frac{dz}{dt}\cdot\frac{\partial t}{\partial x}$; $\frac{\partial z}{\partial y}=\frac{dz}{dt}\cdot\frac{\partial t}{\partial y}$ (заметим, что $z=\ln t$ — функция одной переменной, а потому ее производная обозначается не как частная производная $\frac{\partial z}{\partial t}$, а как обыкновенная производная $\frac{dz}{dt}$). Далее, находим $\frac{dz}{dt}=\frac{1}{t}$, $\frac{\partial t}{\partial x}=\cos x$, $\frac{\partial t}{\partial y}=-\sin y$. Следовательно, $\frac{\partial z}{\partial x}=\frac{1}{t}\cos x=\frac{\cos x}{\sin x+\cos y}$, $\frac{\partial z}{\partial y}=\frac{1}{t}(\sin y)=-\frac{\sin y}{\sin x+\cos y}$. ●

6. Производная по направлению. Градиент

Определение 3. Пусть функция $z=f(x,y)$ определена в некоторой окрестности точки $P_0(x_0;\ y_0)$; l — некоторый луч с началом P_0; $P(x;\ y)$ — точка на этом луче, принадлежащая рассматриваемой окрестности точки P_0 (рис. 172); Δl — длина отрезка P_0P. Если существует $\lim\limits_{\Delta l\to 0}\frac{f(P)-f(P_0)}{\Delta l}$, то этот предел называется *производной* функции $z=f(x,y)$ *по направлению* $\vec{l}$ в точке P_0 и обозначается $\frac{\partial z}{\partial l}$; здесь $\vec{l}$ — вектор, имеющий направление луча l.

В частности, $\frac{\partial z}{\partial x}$ есть производная функции $z=f(x,y)$ по положительному направлению оси x, а $\frac{\partial z}{\partial y}$ — производная по положительному направлению оси y.

Производная по направлению $\frac{\partial z}{\partial l}$ характеризует скорость изменения функции $z=f(x,y)$ в точке $P_0(x_0;\ y_0)$ по направлению $\vec{l}$.

Не ограничивая общности рассуждений, можно считать, что $\vec{l}$ — единичный вектор; тогда он имеет координаты $(\cos\alpha;\sin\alpha)$ или, что то же самое, $(\cos\alpha;\ \cos\beta)$, где α, β — углы между $\vec{l}$ и положительными направлениями оси x, оси y соответственно (рис. 173).

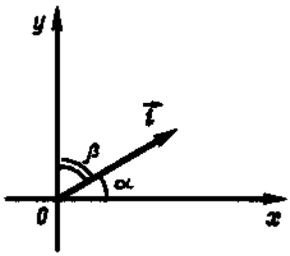


Рис. 172

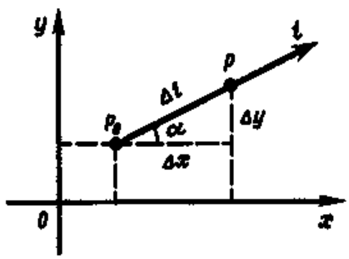


Рис. 173

Теорема 6. *Если функция $z = f(x, y)$ имеет в точке $P_0(x_0; y_0)$ непрерывные частные производные, то в этой точке существует производная $\frac{\partial z}{\partial l}$ по любому направлению $\vec{l} = (\cos\alpha; \cos\beta)$, причем*

$$\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta\,, \qquad (12)$$

где $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ — значения частных производных в точке $(x_0; y_0)$.

□ Так как функция $z = f(x, y)$ имеет в точке $(x_0; y_0)$ непрерывные частные производные, то в силу теоремы 4 она дифференцируема в этой точке:

$$\Delta z = A\Delta x + B\Delta y + \gamma_1\Delta x + \gamma_2\Delta y\,,$$

где $A = \frac{\partial z}{\partial x}$, $B = \frac{\partial z}{\partial y}$, а γ_1, γ_2 — бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$. Тогда $\frac{\Delta z}{\Delta l} = A\frac{\Delta x}{\Delta l} + B\frac{\Delta y}{\Delta l} + \gamma_1\frac{\Delta x}{\Delta l} + \gamma_2\frac{\Delta y}{\Delta l}$. Но $\frac{\Delta x}{\Delta l} = \cos\alpha$, а $\frac{\Delta y}{\Delta l} = \sin\alpha = \cos\beta$ (см. рис. 172). Значит,

$$\frac{\Delta z}{\Delta l} = A\cos\alpha + B\cos\beta + \gamma_1\cos\alpha + \gamma_2\cos\beta\,.$$

Если $\Delta l \to 0$, то $\Delta x \to 0$, $\Delta y \to 0$ и, следовательно, $\gamma_1 \to 0$, $\gamma_2 \to 0$. Поэтому $\lim\limits_{\Delta l \to 0}\frac{\Delta z}{\Delta l} = A\cos\alpha + B\cos\beta$, т. е.

$$\lim_{\Delta l \to 0}\frac{f(P) - f(P_0)}{\Delta l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta\,,$$

а это означает, что $\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta$. ■

Рассмотрим вектор $\vec{a} = \left(\frac{\partial z}{\partial x}; \frac{\partial z}{\partial y}\right)$. Тогда скалярное произведение вектора $\vec{a}$ на вектор $\vec{l} = (\cos\alpha; \cos\beta)$ выражается формулой $\vec{a}\,\vec{l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial l}\cos\beta$. Сопоставляя последнюю формулу с равенством (12), получим $\frac{\partial z}{\partial l} = \vec{a}\,\vec{l}$. С другой стороны, $\vec{a}\,\vec{l} = |\vec{a}||\vec{l}|\cos\varphi$, где φ — угол между векторами $\vec{a}$ и $\vec{l}$. Это скалярное произведение имеет максимальное значение при $\cos\varphi = 1$, т. е. при $\varphi = 0$. Таким образом, наибольшая скорость изменения функции достигается в направлении $\vec{l} = \left(\frac{\partial z}{\partial x}; \frac{\partial z}{\partial y}\right)$.

Определение 4. Вектор с координатами $\left(\frac{\partial z}{\partial x}; \frac{\partial z}{\partial y}\right)$, характеризующий направление максимального роста функции $z = f(x, y)$ в точке $(x_0; y_0)$, называется *градиентом функции* $z = f(x, y)$ в этой точке и обозначается grad z.

Итак, $\operatorname{grad} z = \left(\frac{\partial z}{\partial x}; \frac{\partial z}{\partial y}\right)$.

Производная функции $z = f(x, y)$ в направлении $\vec{l}$ и градиент связаны соотношением

$$\frac{\partial z}{\partial l} = \operatorname{grad} z \cdot \vec{l}.$$

Для функции трех переменных $u = f(x, y, z)$ направление $\vec{l}$ задается вектором $(\cos\alpha; \cos\beta; \cos\gamma)$, где α, β, γ — углы между вектором $\vec{l}$ и положительными направлениями осей x, y, z. Тогда

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma,$$

а grad u есть вектор $\left(\frac{\partial u}{\partial x}; \frac{\partial u}{\partial y}; \frac{\partial u}{\partial z}\right)$.

7. Производные высших порядков

Пусть функция $z = f(x, y)$ имеет частные производные во всех точках множества D. Возьмем любую точку $(x; y) \in D$; в этой точке существуют частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, которые зависят от x и y, т. е. снова являются функциями двух переменных. Значит, можно ставить вопрос об отыскании их частных производных: $\left(\frac{\partial z}{\partial x}\right)'_x, \left(\frac{\partial z}{\partial x}\right)'_y, \left(\frac{\partial z}{\partial y}\right)'_x, \left(\frac{\partial z}{\partial y}\right)'_y$. Если они существуют, то называются *частными производными второго порядка* и обозначаются соответственно $\frac{\partial^2 z}{\partial x^2}$ или $f''_{x^2}(x, y)$ (читается «дэ два зет по дэ икс квадрат»), $\frac{\partial^2 z}{\partial x \partial y}$ или $f''_{xy}(x, y)$ (читается «дэ два зет по дэ икс дэ игрек»), $\frac{\partial^2 z}{\partial y \partial x}$ или $f''_{xy}(x, y)$, $\frac{\partial^2 z}{\partial y^2}$ или $f''_{y^2}(x, y)$.

Итак, в общем случае существует 4 частных производных второго порядка. Аналогично можно говорить о 8 частных производных третьего порядка, 16 — четвертого и т. д.

Примеры. 1. Найти $\frac{\partial^3 z}{\partial x \partial y^2}$ для функции $z = x^2 y^3$.

$$\bigcirc\ \frac{\partial z}{\partial x} = 2xy^3, \frac{\partial^2 z}{\partial x \partial y} = (2xy^3)'_y = 6xy^2; \frac{\partial^3 z}{\partial x \partial y^2} = (6xy^2)'_y = 12xy. \bullet$$

2. Найти $\frac{\partial^2 z}{\partial x \partial y}$ и $\frac{\partial^2 z}{\partial y \partial x}$ для функции $z = \sin(x^2 + y^3)$.

$$\bigcirc\ \frac{\partial z}{\partial x} = \cos(x^2 + y^3)\cdot 2x, \frac{\partial^2 z}{\partial x \partial y} = (\cos(x^2 + y^3)\cdot 2x)'_y = 2x(-\sin(x^2+y^3))3y^2 =$$

$$= -6xy^2 \sin(x^2 + y^3); \frac{\partial z}{\partial y} = \cos(x^2 + y^3)\cdot 3y^2, \frac{\partial^2 z}{\partial y \partial x} = (\cos(x^2 + y^3)\cdot 3y^2)'_x =$$

$$= 3y^2(-\sin(x^2 + y^3))2x = -6xy^2 \sin(x^2 + y^3). \bullet$$

В примере 2 мы получили, что $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$. Оказывается, что это равенство имеет место в достаточно большом числе случаев, что вытекает из следующей теоремы (мы приводим ее без доказательства).

Теорема 7. *Если частные производные второго порядка функции $z = f(x, y)$ непрерывны в точке $(x_0; y_0)$, то в этой точке $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$.*

Аналогичная теорема справедлива для частных производных любого порядка, например $\frac{\partial^3 z}{\partial x \partial y \partial x} = \frac{\partial^3 z}{\partial^2 x \partial y} = \frac{\partial^3 z}{\partial y \partial^2 x}$ (в случае непрерывных частных производных третьего порядка).

§ 46. Исследование функций нескольких переменных на экстремум

1. Максимум и минимум функции нескольких переменных

Определение 1. Пусть функция $z = f(x, y)$ определена в некоторой окрестности точки $(x_0; y_0)$ и непрерывна в этой точке. Если для всех точек $(x; y)$ этой окрестности выполняется неравенство $f(x,y) \leqslant f(x_0, y_0)$ ($f(x,y) \geqslant f(x_0, y_0)$), то $(x_0; y_0)$ называется *точкой максимума (минимума)* функции $f(x,y)$.

Это определение почти дословно повторяет соответствующее определение для функции одной переменной (см. п. 1 § 33). Как и для функции одной переменной, точки максимума и минимума функции нескольких переменных называют *точками экстремума*.

Примеры. 1. Функция $f(x, y) = (x-2)^2 + (y-1)^2 - 1$ имеет точку минимума $(2; 1)$. Действительно, $f(2, 1) = -1$, а если $(x; y) \neq (2; 1)$, то $(x-2)^2 + (y-1)^2 > 0$ и, значит, $f(x, y) > -1$. Здесь $(2; 1)$ — не только точка минимума, но и точка, в которой функция достигает своего наименьшего значения во всей области определения.

2. Функция $f(x, y, z) = 0{,}5 - \sin(x^2 + y^2 + z^2)$ имеет точку максимума $(0; 0; 0)$. В самом деле, $f(0, 0, 0) = 0{,}5$; если взять проколотую окрестность точки $(0; 0; 0)$ такую, чтобы было $x^2 + y^2 + z^2 < \pi/2$, то в этой окрестности $\sin(x^2 + y^2 + z^2) > 0$, а потому $f(x, y, z) < 0{,}5$. Здесь, в отличие от примера 1, $(0; 0; 0)$ — точка максимума, но не точка, в которой функция достигает своего наибольшего значения: это значение равно 1,5 и достигается во всех таких точках $(x; y; z)$, в которых $\sin(x^2 + y^2 + z^2) = -1$ (например, в точках, лежащих на сфере $x^2 + y^2 + z^2 = 3\pi/2$).

Теорема 1 (необходимое условие экстремума). *Если функция $z = f(x, y)$ имеет экстремум в точке $(x_0; y_0)$, то в этой точке частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ либо равны нулю, либо хотя бы одна из них не существует.*

□ Пусть, например, $(x_0; y_0)$ — точка максимума функции $f(x, y)$. Тогда в некоторой окрестности этой точки выполняется неравенство $f(x, y) \leqslant f(x_0, y_0)$ и, в частности, неравенство $f(x, y_0) \leqslant f(x_0, y_0)$. Это означает, что x — точка максимума для функции одной переменной $f(x, y_0)$. Следовательно, производная этой функции, т. е. $\frac{\partial z}{\partial x}$, либо равна нулю, либо не существует (см. теорему 2 § 33). Аналогично устанавливается, что $\frac{\partial z}{\partial y}$ либо равна нулю, либо не существует. ■

Так, для функции $z=(x-2)^2+(y-1)^2-1$ имеем $\frac{\partial z}{\partial x}=2(x-2)$, $\frac{\partial z}{\partial y}=$ $=2(y-1)$. Частные производные обращаются в нуль в точке (2; 1), которая, как было показано в примере 1, является точкой минимума. Для функции $u=0{,}5-\sin(x^2+y^2+z^2)$ имеем $\frac{\partial u}{\partial x}=-2x\cos(x^2+y^2+z^2)$, $\frac{\partial u}{\partial y}=$ $=-2y\cos(x^2+y^2+z^2)$, $\frac{\partial u}{\partial z}=-2z\cos(x^2+y^2+z^2)$; все три частные производные обращаются в нуль в точке (0; 0; 0), которая, как мы установили в примере 2, есть точка максимума функции.

Рассмотрим теперь функцию $z=x^2-y^2$. Имеем $\frac{\partial z}{\partial x}=2x, \frac{\partial z}{\partial y}=-2y$. Обе частные производные обращаются в нуль в точке (0; 0) которая, однако, не является точкой экстремума. В самом деле, $f(0,0)=0$. Если приближаться к точке (0; 0) по оси x (т. е. по прямой $y=0$), то получим $z=x^2>0$; если же приближаться к точке (0; 0) по оси y (т. е. по прямой $x=0$), то получим $z=$ $=-y^2<0$. Следовательно, в любой окрестности точки (0; 0) имеются как точки $(x; y)$, в которых $f(x,y)>f(0,0)$, так и точки, в которых $f(x,y)<f(0,0)$. Значит, (0; 0) не является ни точкой максимума, ни точкой минимума.

Аналогичная ситуация имеет место и в случае функции одной переменной, где условие $y'=0$ или y' не существует есть лишь необходимое, но не достаточное для существования экстремума. Как и для функции одной переменной, нужны достаточные условия, позволяющие установить, имеется ли экстремум в точке, в которой обе частные производные равны нулю, и если имеется, то какой (максимум или минимум). Одно из таких достаточных условий приводится без доказательства в следующей теореме.

Теорема 2 (достаточное условие экстремума функции двух переменных). *Пусть функция $z=f(x,y)$ имеет в некоторой окрестности точки $(x_0; y_0)$ непрерывные частные производные первого и второго порядков, причем $f'_x(x_0,y_0)=0$ и $f'_y(x_0,y_0)=0$. Далее, пусть $f''_{x^2}(x_0,y_0)=$ $=A$, $f''_{xy}(x_0,y_0)=B$, $f''_{y^2}(x_0,y_0)=C$. Тогда если $AC-B^2>0$ и $A<0$, то $(x_0; y_0)$ — точка максимума функции $z=f(x,y)$; если $AC-B^2>0$ и $A>0$, то $(x_0; y_0)$ — точка минимума; если $AC-B^2<0$, то в точке $(x_0; y_0)$ экстремума нет.*

АЛГОРИТМ

исследования функции $z=f(x,y)$ на экстремум

1°. Найти первые частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$.

2°. Найти стационарные точки, т. е. точки, в которых $\frac{\partial z}{\partial x}=0$ и $\frac{\partial z}{\partial y}=0$.

3°. Найти вторые частные производные $\frac{\partial^2 z}{\partial x^2}$, $\frac{\partial^2 z}{\partial x\partial y}$, $\frac{\partial^2 z}{\partial y^2}$.

4°. Вычислить значения вторых производных в стационарных точках.

5°. Для каждой стационарной точки найти значение $\Delta=AC-B^2$, где $A=\frac{\partial^2 z}{\partial x^2}$, $C=\frac{\partial^2 z}{\partial y^2}$ и $B=\frac{\partial^2 z}{\partial x\partial y}$, и сделать выводы на основании теоремы 2.

Пример 3. Исследовать на экстремум функцию $z = x^3 + y^3 - 3xy$.

○ 1^0. Находим $\dfrac{\partial z}{\partial y} = 3x^2 - 3y$, $\dfrac{\partial z}{\partial x} = 3y^2 - 3x$.

2^0. Решив систему $\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$, находим две стационарные точки (0; 0) и (1; 1).

3^0. Находим $\dfrac{\partial^2 x}{\partial x^2} = 6x$, $\dfrac{\partial^2 z}{\partial y^2} = 6y$, $\dfrac{\partial^2 z}{\partial x \partial y} = -3$.

4^0. Для точки (0; 0) имеем $A_1 = 0$, $C_1 = 0$, $B_1 = -3$; для точки (1; 1) имеем $A_2 = 6$, $C_2 = 6$, $B_2 = -3$.

5^0. Находим $\Delta_1 = A_1C_1 - B_1^2 = -9$; $\Delta_2 = A_2C_2 - B_2^2 = 27$. Так как $\Delta_1 < 0$, то в точке (0; 0) экстремума нет; так как $\Delta_2 > 0$ и $A_2 > 0$, то (1; 1) — точка минимума функции, причем $z_{\min} = -1$.●

2. Отыскание наибольшего и наименьшего значений непрерывной функции на замкнутом ограниченном множестве

Выше мы отмечали, что функция, непрерывная на замкнутом ограниченном множестве D, достигает на нем наибольшего и наменьшего значений (см. теорему 13 § 44). Эти значения она может принимать как во внутренних точках множества D (ясно, что каждая такая точка является точкой экстремума функции; в этой точке первые частные производные равны нулю или не существуют), так и на его границе; значит, необходимо специальное исследование граничных точек множества D.

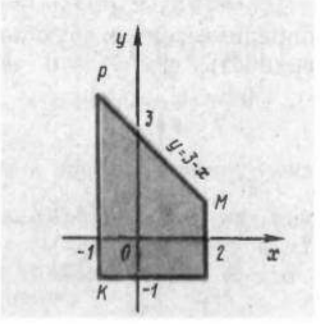


Рис. 174

Пример. Найти наибольшее и наменьшее значения функции $z = x^3 + y^3 - 3xy$ в области, ограниченной прямыми $x = -1$, $x = 2$, $y = -1$, $y = 3 - x$ (рис. 174).

○ Алгоритм решения примерно такой же, как и для отыскания наибольшего и наименьшего значений непрерывной на отрезке функции одной переменной (см. п. 4 § 34).

Внутри заданной области функция имеет две стационарные точки: O (0; 0) и E (1; 1) (см. пример 3 п. 1).

Исследуем поведение функции на границе области.

Отрезок KL имеет уравнение $y = -1$, $-1 \leqslant x \leqslant 2$. Подставив $y = -1$ в заданную функцию, получим $z = x^3 - 1 + 3x$. Нужно найти наибольшее и наименьшее значения этой функции на отрезке $[-1, 2]$.

Имеем $z' = 3x^2 + 3 > 0$; значит, функция возрастает и потому достигает наибольшего и наименьшего значений на концах отрезка, т. е. в точках $K(-1; -1)$ и $L(2; -1)$.

Отрезок LM имеет уравнение $x = 2$, $-1 \leqslant y \leqslant 1$. Подставив $x = 2$ в заданную функцию, получим $z = 8 + y^3 - 6y$. Имеем $z' = 3y^2 - 6 < 0$ на отрезке $[-1, 1]$. Значит, функция $z = 8 + y^3 - 6y$ достигает наибольшего и наименьшего значений на концах отрезка, т. е. в точках L (2; −1) и M (2; 1).

Отрезок PM имеет уравнение $y = 3 - x$, $-1 \leqslant x \leqslant 2$. Подставив $y = 3 - x$ в заданную функцию, получим $z = x^3 + (3-x)^3 - 3x(3-x)$, т. е. $z = 27 - 36x + 12x^2$. Имеем $z' = 24x - 36$, откуда $z' = 0$ при $x = 3/2$. Значит, на отрезке PM функция может достигать наибольшего и наименьшего значений в точках $M\ (2;\ 1)$, $P\ (-1;\ 4)$ и $T\ (3/2;\ 3/2)$.

Отрезок KP имеет уравнение $x = -1$, $-1 \leqslant y \leqslant 4$. Подставив $x = -1$ в заданную функцию, получим $z = -1 + y^3 + 3y$. Имеем $z' = 3y^2 + 3 > 0$; значит, функция достигает наибольшего и наименьшего значений на концах отрезка, т. е. в точках $K\ (-1;\ -1)$ и $P\ (-1;\ 4)$.

Таким образом, функция $z = x^3 + y^3 - 3xy$ может достигать наибольшего и наименьшего значений только в следующих точках: $O\ (0;\ 0)$, $E\ (1;\ 1)$, $K\ (-1;\ -1)$, $L\ (2;\ -1)$, $M\ (2;\ 1)$, $P\ (-1;\ 4)$ и $T\ (3/2;\ 3/2)$.

Находим: $f(0,\ 0) = 0$; $f\ (1;\ 1) = -1$; $f\ (-1;\ -1) = -5$; $f\ (2;\ -1) = 13$; $f\ (2;\ 1) = 3$; $f\ (-1;\ 4) = 75$; $f\ (3/2;\ 3/2) = 0$.

Итак, $z_{\text{наим}} = -5$ и это значение достигается в точке $(-1;\ -1)$; $z_{\text{наиб}} = 75$ и это значение достигается в точке $(-1;\ 4)$. ●

3. Условный экстремум

Как известно, уравнение окружности с центром в точке $(0;\ 0)$ и радиусом 1 имеет вид $x^2 + y^2 = 1$. Это же уравнение в системе координат xyz определяет цилиндрическую поверхность, образующая которой параллельна оси z (рис. 175). Вообще, уравнение $\varphi(x,\ y) = 0$, задающее некоторую кривую на плоскости xy (рис. 176, *а*), определяет в пространстве xyz цилиндрическую поверхность (рис. 176, *б*).

Пусть функция $z = f(x,\ y)$ определяет некоторую поверхность в пространстве xyz и цилиндрическая поверхность $\varphi(x,\ y) = 0$ высекает из поверхности

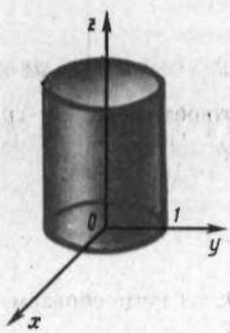


Рис. 175

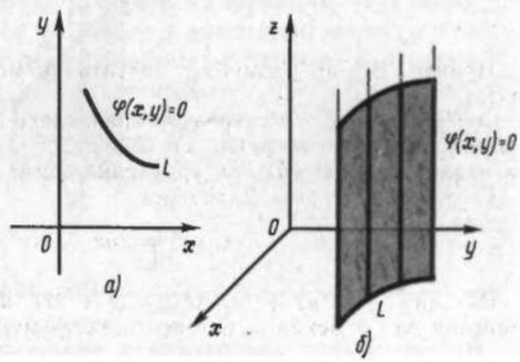


Рис. 176

$z = f(x,\ y)$ некоторую линию. На этой пространственной линии могут быть свои точки максимума и минимума — они называются условными экстремумами. Дадим строгое определение.

Определение 2. Пусть дано уравнение $\varphi(x,\ y) = 0$ и точка $(x_0;\ y_0)$ удовлетворяет этому уравнению. Далее, пусть функция $z = f(x,\ y)$ определена в некоторой окрестности точки $(x_0;\ y_0)$ и непрерывна в этой точке. Тогда если для всех точек $(x;\ y)$ этой окрестности, удовлетворяющих уравнению $\varphi(x,\ y) = 0$, выполняется неравенство $f(x,\ y) \leqslant f(x_0,\ y_0)$ $(f(x,\ y) \geqslant f(x_0,\ y_0))$, то $(x_0;\ y_0)$ называется *точкой условного максимума (минимума)* функции $f(x,\ y)$, а $\varphi(x,\ y) = 0$ — *уравнением связи*.

Для отыскания условных экстремумов используется *метод множителей Лагранжа;* он вытекает из следующей теоремы (мы приводим ее без доказательства) выражающей необходимое условие условного экстремума.

Теорема 3. *Если $(x_0; y_0)$ — точка условного экстремума функции $z = f(x, y)$ при уравнении связи $\varphi(x, y) = 0$, то существует такое число λ, что тройка чисел $(x_0; y_0; \lambda)$ удовлетворяет системе уравнений*

$$\begin{cases} F'_x(x, y) = 0\,, \\ F'_y(x, y) = 0\,, \\ \varphi(x, y) = 0\,, \end{cases}$$

где $F(x, y) = f(x, y) + \lambda\varphi(x, y)$ есть так называемая функция Лагранжа.

Аналогично определяется понятие условного экстремума для функции трех переменных $u = F(x, y, z)$ при уравнении связи $\varphi(x, y, z) = 0$. В этом случае функция Лагранжа имеет вид $F(x, y, z) = f(x, y, z) + \lambda\varphi(x, y, z)$, а система уравнений для отыскания условного экстремума — вид

$$\begin{cases} F'_x(x, y, z) = 0\,, \\ F'_y(x, y, z) = 0\,, \\ F'_z(x; y, z) = 0\,, \\ \varphi(x, y, z) = 0\,. \end{cases}$$

Пример. В шар радиуса r вписать прямоугольный параллелепипед наибольшего объема.

○ Пусть x, y, z — измерения параллелепипеда, тогда его объем u равен xyz. Диагональ параллелепипеда равна $2r$; значит, $x^2 + y^2 + z^2 = 4r^2$. Итак, нужно найти максимум функции $u = xyz$ при уравнении связи $x^2 + y^2 + z^2 - 4r^2 = 0$.

Составим функцию Лагранжа

$$F(x,y,z) = xyz + \lambda(x^2 + y^2 + z^2 - 4r^2)\,.$$

Находим $F'_x = yz + 2\lambda x$, $F'_y = xz + 2\lambda y$, $F'_z = xy + 2\lambda z$. Таким образом, система уравнений для отыскания условного экстремума имеет вид

$$\begin{cases} yz + 2\lambda x = 0\,, \\ xz + 2\lambda y = 0\,, \\ xy + 2\lambda z = 0\,, \\ x^2 + y^2 + z^2 = 4r^2\,. \end{cases}$$

Умножим обе части первого уравнения на x, второго — на y и вычтем второе уравнение из первого: $(xyz + 2\lambda x^2) - (xyz + 2\lambda y^2) = 0$. Получим $\lambda(x^2 - y^2) = 0$, откуда $x = y$. Аналогично $y = z$. Тогда последнее уравнение системы примет вид $x^2 + x^2 + x^2 = 4r^2$, откуда $x = 2r\sqrt{3}\,/3$. По смыслу задачи при $x = y = z = 2r\sqrt{3}\,/3$ функция принимает наибольшее значение.

Итак, из всех прямоугольных параллелепипедов, вписанных в шар радиуса r, наибольший объем имеет куб с ребром $2r\sqrt{3}\,/3$.●

§ 47. Двойной интеграл

1. Задача о массе материальной плоской пластины

Пусть дана плоская неоднородная пластина D, линейная плотность которой в точке $(x; y)$ выражается функцией $\rho(x, y)$. Найдем массу μ пластины.

Выше (см. п. 2 § 39) мы решали задачу о массе материального стержня. Воспользуемся тем же методом: произведем разбиение T фигуры D на n мелких ячеек (рис. 177). Возьмем ячейку D_k; пусть m_k и M_k — соответственно наименьшее и наибольшее значения функции $\rho(x, y)$ на множестве D_k. Тогда масса

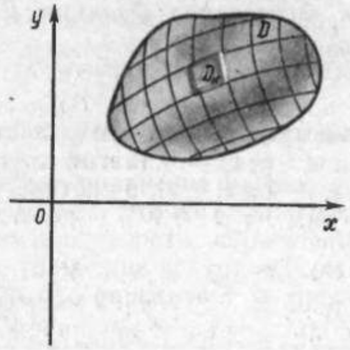


Рис. 177

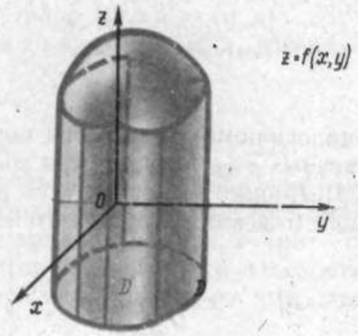


Рис. 178

ячейки D_k заключена между числами $m_k \cdot$ пл D_k и $M_k \cdot$ пл D_k, где пл D_k — площадь* фигуры D_k. Проведя аналогичные рассуждения для остальных ячеек разбиения, получим, что масса μ пластины D удовлетворяет двойному неравенству $\underline{S_T} \leqslant \mu \leqslant \bar{S}_T$, где $\underline{S}_T = \sum\limits_{k=0}^{n-1} m_k \cdot \text{пл} D_k$, а $\bar{S}_T = \sum\limits_{k=0}^{n-1} M_k \cdot \text{пл} D_k$.

Таким образом, *масса пластины есть число, разделяющее множества* $\{\underline{S_T}\}$ *и* $\{\bar{S}_T\}$ *для всевозможных разбиений* T *фигуры* D.

2. Задача об объеме цилиндрического бруса

Пусть в плоскости xy дано замкнутое ограниченное связное множество D, на котором определена неотрицательная функция $z = f(x, y)$. Пространственное тело, ограниченное плоской фигурой D, графиком функции $z = f(x, y)$ и перпендикулярами к D, проведенными из точек контура D до пересечения с поверхностью $z = f(x, y)$, будем называть *цилиндрическим брусом* (рис. 178). Найдем его объем.

Если функция $z = f(x,y)$ постоянна, т. е. $f(x,y) = H$, то объем вычисляется по формуле пл.$D \cdot H$, где пл.D — площадь фигуры D. Если же $f(x,y) \neq \text{const}$, то этой формулой пользоваться нельзя. Поступим следующим образом. Произведем разбиение T фигуры D на n мелких ячеек и рассмотрим ячейку D_k; пусть m_k и M_k — соответственно наименьшее и наибольшее значения функции $f(x, y)$ на множестве D_k. Тогда объем цилиндрического столбика, построенного на D_k и упирающегося сверху в график функции $z = f(x, y)$, заключен между

---

* Всюду в настоящем параграфе мы рассматриваем только квадрируемые (т. е. имеющие площадь; см. § 41) фигуры, не оговаривая этого каждый раз особо.

числами $m_k \cdot$пл.D_k и $M_k \cdot$пл.D_k, где пл. D_k — площадь фигуры D_k. Проведя аналогичные рассуждения для остальных ячеек разбиения, получим, что объем V цилиндрического бруса удовлетворяет двойному неравенству $\underline{S}_T \leqslant V \leqslant \bar{S}_T$, где $\underline{S}_T = \sum_{k=0}^{n-1} m_k \cdot$пл.$D_k$, $\bar{S}_T = \sum_{k=0}^{n-1} M_k \cdot$пл.$D_k$.

Таким образом, *объем цилиндрического бруса есть число, разделяющее множества* $\{\underline{S}_T\}$ *и* $\{\bar{S}_T\}$ *для всевозможных разбиений T фигуры D.*

3. Определение двойного интеграла

Две различные задачи, рассмотренные в предыдущих пунктах, привели в процессе решения к одной и той же математической модели, к двум определенным образом построенным числовым множествам $\{\underline{S}_T\}$ и $\{\bar{S}_T\}$, разделяющимся единственным числом; в первом случае это число характеризует массу неоднородной плоской пластины, во втором — объем цилиндрического бруса. Многие задачи приводят к такой же математической модели, когда для функции $z = f(x, y)$, определенной на множестве D, составляются нижняя и верхняя суммы Дарбу $\underline{S}_T = \sum_{k=0}^{n-1} m_k \cdot$пл.$D_k$ и $\bar{S}_T = \sum_{k=0}^{n-1} M_k \cdot$пл.$D_k$, где T — разбиение фигуры D на n мелких ячеек, пл. D_k — площадь ячейки D_k, а m_k и M_k — соответственно нижняя и верхняя грани значений функции $z = f(x, y)$ на множестве D_k.

Такой схемой мы пользовались и ранее при определении определенного интеграла (см. § 39). Как и в указаном параграфе, устанавливается, что множество $\{\underline{S}_T\}$ расположено левее множества $\{\bar{S}_T\}$, а потому всегда имеется хотя бы одно разделяющее их число.

Определение 1. Функция $f(x, y)$, ограниченная на замкнутом квадрируемом множестве*, называется *интегрируемой* на этом множестве, если существует единственное число I, разделяющее множества нижних и верхних сумм Дарбу, образованных для всевозможных разбиений фигуры D на квадрируемые ячейки. Это единственное число называют *двойным интегралом* функции по множеству D и обозначают $\iint\limits_D f(x, y)\,dx\,dy$.

Теорема 1. *Если функция $f(x, y)$ непрерывна на множестве D, то она интегрируема на D.*

Доказательство аналогично доказательству теоремы 2 § 39 и мы его не приводим.

Опираясь на рассуждения, проведенные в пп. 1 и 2, приходим к следующим выводам:

масса μ плоской пластины D с линейной плотностью $\rho(x, y)$ вычисляется по формуле

$$\mu = \iint\limits_D \rho(x, y)\,dx\,dy$$

(физический смысл двойного интеграла);

объем V цилиндрического бруса, ограниченного снизу фигурой D, лежащей

---

* В дальнейшем в этом параграфе мы всюду понимаем под D замкнутое квадрируемое множество, не оговаривая этого каждый раз специально.

в плоскости xy, а сверху — графиком непрерывной неотрицательной функции $z = f(x, y)$, вычисляется по формуле

$$V = \iint_D f(x, y)dxdy$$

(геометрический смысл двойного интеграла).

4. Двойной интеграл как предел интегральных сумм

В этом пункте мы рассмотрим другой подход к введению двойного интеграла.

Пусть функция $z = f(x, y)$ интегрируема на множестве D, и пусть $\iint_D f(x,y)dxdy = I$. Разобьем фигуру D на n мелких ячеек, затем в каждой ячейке D_k возьмем произвольную точку (ξ_k, η_k) и, наконец, составим сумму $S_T = \sum_{k=0}^{n-1} f(\xi_k, \eta_k)\text{пл.}D_k$, которая называется *интегральной суммой*. Тогда

$$\underline{S}_T \leqslant S_i \leqslant \bar{S}_T . \qquad (1)$$

Назовем *диаметром ячейки* D_k наибольшую из длин всех отрезков, принадлежащих ячейке D_k, а *диаметром разбиения* T — наибольший из диаметров ячеек; обозначим диаметр разбиения d_T. Чем мельче разбиение, т. е. чем меньше d_T, тем меньше отличаются друг от друга $\underline{S}_T$ и $\bar{S}_T$. Справедливо следующее утверждение.

Лемма. *Если функция $f(x, y)$ непрерывна на D, то для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что из $d_T < \delta$ следует неравенство*

$$\bar{S}_T - \underline{S}_T < \varepsilon . \qquad (2)$$

Поскольку

$$\underline{S}_T \leqslant I \leqslant \bar{S}_T , \qquad (3)$$

из неравенств (1), (2) и (3) получаем такой результат: если функция $z = f(x, y)$ непрерывна на множестве D, то для любого $\varepsilon > 0$ существует $\delta > 0$, что из $d_T > \delta$ следует $|I - S_T| < \varepsilon$. Короче:

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall T : d_T < \delta)|I - S_T| < \varepsilon .$$

Эта запись соответствует равенству $\lim\limits_{d_T \to 0} S_T = I$ при нескольких дополнительных предположениях, которые отражены в следующем варианте определения двойного интеграла.

Определение 2. Если для ограниченной на множестве D функции $z = f(x, y)$ существует $\lim\limits_{d_T \to 0} S_T$, причем этот предел не зависит ни от способа разбиения фигуры D на ряд мелких участков, ни от выбора точек $(\xi_k; \eta_k)$ внутри ячеек D_k, то он называется *двойным интегралом* функции $z = f(x, y)$ по множеству D.

Итак, мы имеем два эквивалентных определения двойного интеграла: как единственное число, разделяющее множества нижних и верхних сумм Дарбу, и как предел интегральных сумм при $d_T \to 0$. На практике в равной степени используют оба определения.

Заметим, что и определенный интеграл $\int_a^b f(x)dx$ можно было определить как $\lim\limits_{d_T \to 0} S_T$, где $S_T = \sum_{k=0}^{n-1} f(\xi_k)\Delta x_k$, $\xi_k \in [x_k, x_{k+1}]$, а d_T — длина наибольшего из отрезков $[x_k, x_{k+1}]$.

5. Свойства двойного интеграла

Используя определения 1 и 2, а также по необходимости геометрический или физический смысл двойного интеграла, нетрудно получить ряд его свойств, аналогичных свойствам определенного интеграла, рассмотренным в § 39 и 40. Приведем эти свойства.

1^0. $\iint\limits_D dx\,dy = \text{пл}.D$.

2^0. $\iint\limits_D \lambda f(x, y)dx\,dy = \lambda\iint\limits_D f(x, y)dx\,dy$.

3^0. $\iint\limits_D (f_1(x, y) + f_2(x, y))dx\,dy = \iint\limits_D f_1(x, y)dx\,dy + \iint\limits_D f_2(x, y)dx\,dy$.

4^0. *Если* $f(x, y) > 0$ *на* D, *то* $\iint\limits_D f(x, y)dx\,dy > 0$.

5^0. *Если* $f(x, y) \geqslant g(x, y)$ *на* D, *то* $\iint\limits_D f(x, y)dx\,dy \geqslant \iint\limits_D g(x, y)dx\,dy$.

6^0. *Если* $D = D_1 \cup D_2$, *причем* D_1 *и* D_2 *не имеют общих внутренних точек, то* $\iint\limits_D f(x, y)dx\,dy = \iint\limits_{D_1} f(x, y)dx\,dy + \iint\limits_{D_2} f(x, y)dx\,dy$ *(аддитивное свойство двойного интеграла).*

7^0. *Если функция* $f(x, y)$ *непрерывна на* D, *то существует такая точка* $(\xi, \eta) \in D$, *что*

$$\iint\limits_D f(x, y)dx\,dy = f(\xi, \eta)\,\text{пл}.D.$$

(теорема о среднем для двойного интеграла).

Докажем, например, что справедливо свойство 3^0. Имеем

$$\iint\limits_D (f_1(x,y) + f_2(x,y)dx\,dy = \lim_{d_T \to 0} \sum_{k=0}^{n-1} (f_1(\xi_k, \eta_k) + f_2(\xi_k, \eta_k))\,\text{пл}.D_k =$$

$$= \lim_{d_T \to 0} \sum_{k=0}^{n-1} f_1(\xi_k, \eta_k)\,\text{пл}.D_k + \lim_{d_T \to 0} \sum_{k=0}^{n-1} f_2(\xi_k, \eta_k)\,\text{пл}.D_k =$$

$$= \iint\limits_D f_1(x, y)dx\,dy + \iint\limits_D f_2(x, y)dx\,dy.$$

Мы воспользовались определением интеграла как предела интегральных сумм, предполагая, естественно, что обе функции интегрируемы на D.

Докажем еще справедливость свойства 5^0.

Если $f(x, y) \geqslant g(x, y)$, то для любого разбиения T выполнено неравенство $\underline{S}_T^f \geqslant \bar{S}_T^g$, где $\underline{S}_T^f$ — нижняя сумма Дарбу для функции $f(x, y)$, а $\bar{S}_T^g$ — верхняя сумма Дарбу для $g(x, y)$. Далее, имеем

$$\iint\limits_D f(x, y)dx\,dy \geqslant \underline{S}_T^f \geqslant \bar{S}_T^g \geqslant \iint\limits_D g(x, y)\,dx\,dy.$$

Здесь мы воспользовались определением двойного интеграла как разделяющего числа.

6. Вычисление двойного интеграла

Выше мы отметили, что если функция $f(x, y)$ непрерывна и неотрицательна на D, то $\iint\limits_D f(x, y)dx\,dy$ выражает объем цилиндрического бруса. Пусть D — криволи-

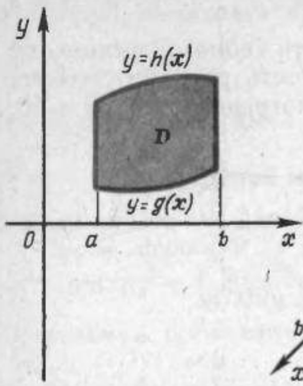


Рис. 179

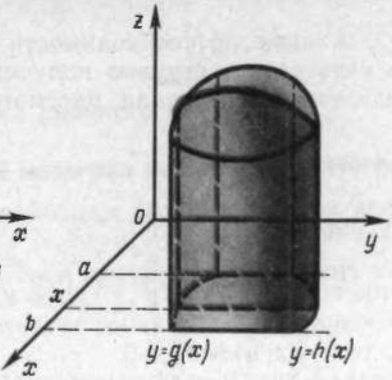


Рис. 180

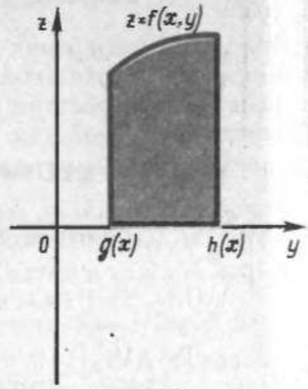


Рис. 181

нейная трапеция, ограниченная прямыми $x=a$, $x=b$ $(a<b)$ и графиками непрерывных на $[a, b]$ функций $y=g(x)$, $y=h(x)$, где $g(x)\leqslant h(x)$ (рис. 179). Построим сечение бруса плоскостью $x=\text{const}$ (рис. 180) и обозначим площадь сечения через $S(x)$. Тогда $V=\int\limits_a^b S(x)dx$ (см. теорему 8 § 41).

Чтобы вычислить $S(x)$, спроецируем сечение на плоскость yz; получим криволинейную трапецию (рис. 181), площадь которой вычисляется по формуле

$$S(x)=\int\limits_{g(x)}^{h(x)} f(x,y)dy$$

(напомним, что $x=\text{const}$). Значит,

$$V=\int\limits_a^b\Bigl(\int\limits_{g(x)}^{h(x)} f(x,y)dy\Bigr)dx.$$

Обычно множители $\int\limits_{g(x)}^{h(x)} f(x,y)dy$ и dx меняют местами.

Таким образом,

$$\iint\limits_D f(x,y)dx\,dy=\int\limits_a^b dx\int\limits_{g(x)}^{h(x)} f(x,y)dy. \qquad (4)$$

Формула (4) справедлива для любой непрерывной на D функции, не обязательно неотрицательной на D. Ее смысл состоит в том, что вычисление двойного интеграла сводится к вычислению двух определенных интегралов: сначала вычисляется внутренний интеграл $\int\limits_{g(x)}^{h(x)} f(x,y)dy$, в результате чего получается некоторая функция $S(x)$, а затем интеграл от полученной функции $S(x)$ по промежутку $[a, b]$.

Пример 1. Вычислить $\iint\limits_D (x^2+y^2)\,dx\,dy$, где D — фигура, ограниченная линиями $y=2$, $y=x$, $xy=1$ (рис. 182).

О Фигура D есть объединение двух криволинейных трапеций: D_1, ограниченной линиями $x=1/2$, $x=1$, $y=1/x$ и $y=2$, и D_2, ограниченной линиями $x=1$, $x=2$, $y=x$ и $y=2$. Сначала вычислим интеграл по D_1, а затем по D_2. Имеем

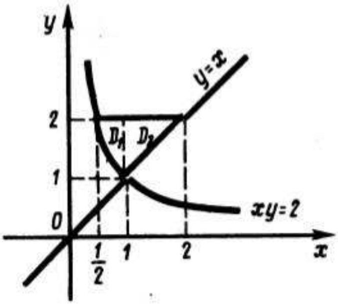


Рис. 182

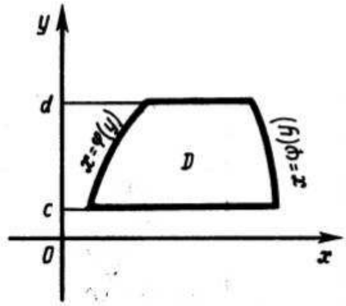


Рис. 183

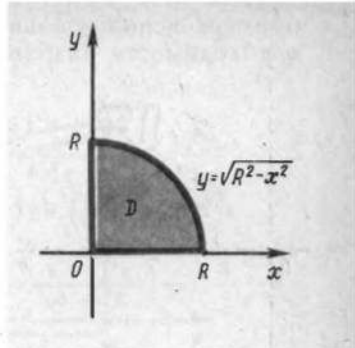


Рис. 184

$$\iint\limits_{D_1} (x^2+y^2)\,dx\,dy = \int\limits_{1/2}^{1} dx \int\limits_{1/x}^{2} (x^2+y^2)\,dy = \int\limits_{1/2}^{1} dx\left(x^2y+\frac{y^3}{3}\right)\Big|_{1/x}^{2}\Big] =$$

$$= \int\limits_{1/2}^{1} dx\left(2x^2+\frac{8}{3}-x^2\cdot\frac{1}{x}-\frac{1}{3}\cdot\frac{1}{x^3}\right) = \int\limits_{1/2}^{1}\left(2x^2+\frac{8}{3}-x-\frac{1}{3x^3}\right)dx =$$

$$=\left(2\cdot\frac{x^3}{3}+\frac{8}{3}x-\frac{x^2}{2}-\frac{1}{3}\cdot\frac{x^{-2}}{(-2)}\right)\Big|_{1/2}^{1} = \left(\frac{2x^3}{3}+\frac{8x}{3}-\frac{x^2}{2}+\frac{1}{6x^2}\right)\Big|_{1/2}^{1} =$$

$$=\left(\frac{2}{3}+\frac{8}{3}-\frac{1}{2}+\frac{1}{6}\right)-\left(\frac{1}{12}+\frac{4}{3}-\frac{1}{8}+\frac{2}{3}\right)=\frac{25}{24}.$$

Далее, находим

$$\iint\limits_{D_2} (x^2+y^2)\,dx\,dy = \int\limits_{1}^{2} dx \int\limits_{x}^{2} (x^2+y^2)\,dy = \int\limits_{1}^{2} dx\left(x^2y+\frac{y^3}{3}\right)\Big|_{x}^{2} =$$

$$=\int\limits_{1}^{2} dx\left(2x^2+\frac{8}{3}-x^3-\frac{x^3}{3}\right) = \int\limits_{1}^{2}\left(2x^2+\frac{8}{3}-\frac{4x^3}{3}\right)dx =$$

$$=\left(\frac{2x^3}{3}+\frac{8x}{3}-\frac{x^4}{3}\right)\Big|_{1}^{2} = \frac{1}{3}(16+16-16)-\frac{1}{3}(2+8-1)=\frac{7}{3}.$$

Наконец, воспользовавшись аддитивным свойством двойного интеграла (см. свойство 6° из п. 5), получим

$$\iint\limits_{D} (x^2+y^2)\,dx\,dy = \iint\limits_{D_1} (x^2+y^2)\,dx\,dy + \iint\limits_{D_2} (x^2+y^2)\,dx\,dy = \frac{25}{24}+\frac{7}{3}=\frac{27}{8}. \bullet$$

Если D — криволинейная трапеция, ограниченная линиями $y=c$, $y=d$ $(c<d)$, $x=\varphi(y)$, $x=\psi(y)$, где $\varphi(y)\leqslant\psi(y)$ на $[c,\ d]$ (рис. 183), то, рассуждая как и при выводе формулы (4), получим

$$\iint\limits_{D} f(x,\ y)\,dx\,dy = \int\limits_{c}^{d} dy \int\limits_{\varphi(y)}^{\psi(y)} f(x,\ y)\,dx. \qquad (5)$$

На практике формулы (4) и (5) используются в равной степени. Обычно предпочтение отдается той из них, которая приводит либо к меньшему числу слагаемых, либо к более простым вычислениям. Так, в рассмотренном выше примере использование формулы (5) предпочтительнее, поскольку тогда нет необходимости разбивать D на составные части. Получим

$$\iint\limits_D (x^2+y^2)\,dx\,dy=\int_1^2 dy\int_{1/y}^{y}(x^2+y^2)\,dx=\int_1^2 dy\left(\frac{x^3}{3}+y^2x\right)\Big|_{1/y}^{y}=$$

$$=\int_1^2 dy\left(\frac{y^3}{3}+y^3-\frac{1}{3y^3}-y\right)=\int_1^2\left(\frac{4y^3}{3}-\frac{1}{3y^3}-y\right)dy=$$

$$=\left(\frac{y^4}{3}+\frac{1}{6y^2}-\frac{y^2}{2}\right)\Big|_1^2=\left(\frac{16}{3}+\frac{1}{24}-2\right)-\left(\frac{1}{3}+\frac{1}{6}-\frac{1}{2}\right)=\frac{27}{8}.$$

Пример 2. Вычислить объем тела, ограниченного координатными плоскостями, цилиндром $x^2+y^2=R^2$ и гиперболическим параболоидом $z=xy$ (в I октанте).

○ Данное тело представляет собой цилиндрический брус, в основании которого лежит фигура D, представляющая собой четверть круга (рис. 184), а сверху брус ограничен поверхностью гиперболического параболоида $z=xy$. Значит, объем V такого тела вычисляется по формуле $V=\iint\limits_D xy\,dx\,dy$. Имеем

$$V=\iint\limits_D xy\,dx\,dy=\int_0^R dx\int_0^{\sqrt{R^2-x^2}} xy\,dy=\int_0^R x\,dx\frac{y^2}{2}\Big|_0^{\sqrt{R^2-x^2}}=$$

$$=\frac{1}{2}\int_0^R x(R^2-x^2)\,dx=\left(\frac{R^2}{4}x^2-\frac{1}{8}x^4\right)\Big|_0^R=\frac{R^4}{4}-\frac{R^4}{8}=\frac{R^4}{8}. \bullet$$

§ 48. Криволинейный интеграл

1. Задача о работе силового поля

Предположим, что в каждой точке плоскости xy действует вектор силы $\vec{F}(x, y)$ с горизонтальной составляющей $P(x,y)$ и вертикальной составляющей $Q(x, y)$. Пусть в этом силовом поле материальная точка прошла криволинейный путь BC (рис. 185). Найдем произведенную при этом работу.

Известно, что работа A, произведенная при движении в направлении вектора $\vec{s}$ под действием силы $\vec{F}$, вычисляется по формуле $A=\vec{F}\vec{s}$ (скалярное произведение векторов). Если при этом $\vec{F}=(P;\ Q)$, а $\vec{s}=(a;\ b)$, то $\vec{F}\vec{s}=Pa+$ $+Qb$.

В данном случае пользоваться этой формулой мы не можем, так как сила $\vec{F}$ меняется от точки к точке, а путь движения не прямолинеен. Поступим следующим образом. Разобьем дугу BC на n мелких участков (рис. 185) и рассмотрим k-й участок разбиения дуги M_kM_{k+1}. Возьмем на нем произвольную точку $(\xi_k;\ \eta_k)$ и будем считать, что на всем участке действует постоянная сила, такая же, как и в этой точке: $\vec{F}(\xi_k,\eta_k)=(P(\xi_k,\eta_k);\ Q(\xi_k,\eta_k))$. Кроме того, сделаем еще одно допущение: будем считать, что движение осуществляется не по участку дуги, а по вектору $\Delta\vec{s}_k$, начальная и конечная точки которого совпадают с начальной и конечной точками участка дуги. Тогда $\Delta\vec{s}_k=(\Delta x_k;\ \Delta y_k)$ (рис. 186).

При сделанных допущениях работа на k-м участке выражается приближенной формулой

$$A_k\approx P(\xi_k,\ \eta_k)\,\Delta x_k+Q(\xi_k,\ \eta_k)\,\Delta y_k,$$

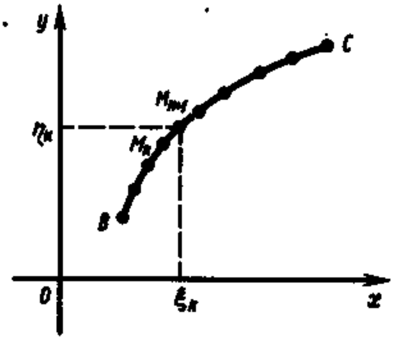


Рис. 185

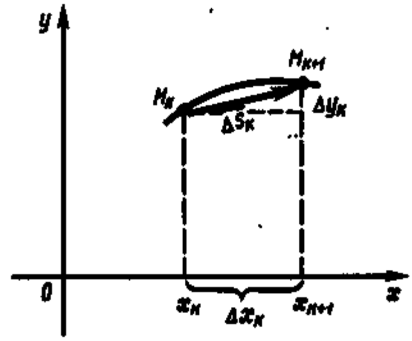


Рис. 186

а вся работа — приближенной формулой $A \approx S_T$, где

$$S_T = \sum_{k=0}^{n-1} P(\xi_k,\ \eta_k)\,\Delta x_k + Q(\xi_k,\ \eta_k)\,\Delta y_k.$$

Точное равенство получится переходом к пределу при условии $d_T \to 0$, где d_T — диаметр разбиения (наибольшая из длин дуг участков разбиения): $A = \lim\limits_{d_T \to 0} S_T$.

2. Определение и свойства криволинейного интеграла

Осмыслим ту математическую модель, которая была построена в п. 1. Пусть функции $P(x,\ y)$ и $Q(x,\ y)$ определены во всех точках дуги L. Производят разбиение T дуги L на ряд мелких участков, в каждом из участков выбирают по точке, вычисляют значения обеих функций в выбранных точках, составляют интегральную сумму $S_T = \sum\limits_{k=0}^{n-1} P(\xi_k,\ \eta_k)\,\Delta x_k + Q(\xi_k,\ \eta_k)\,\Delta y_k$ и вычисляют $\lim\limits_{d_T \to 0} S_T$.

Определение. Если $\lim\limits_{d_T \to 0} S_T$ существует и не зависит ни от способа разбиения дуги L на ряд мелких участков, ни от выбора точек внутри участков разбиения, то он называется *криволинейным интегралом* и обозначается $\int\limits_L P(x,\ y)\,dx + Q(x,\ y)\,dy$.

Опираясь на рассуждения, проведенные в п. 1, приходим к следующему выводу: *если в силовом поле действует сила $\vec{F} = (P(x,\ y);\ Q(x,\ y))$, то при движении в этом поле по дуге L совершается работа, вычисляемая по формуле*

$$A = \int\limits_L P(x,\ y)\,dx + Q(x,\ y)\,dy.$$

В этом состоит физический смысл криволинейного интеграла.

Из определения криволинейного интеграла вытекают следующие его свойства:

1^0. $\int\limits_L P(x,\ y)\,dx + Q(x,\ y)\,dy = \int\limits_L P(x,\ y)\,dx + \int\limits_L Q(x,\ y)\,dy.$

2^0. *Если* $L=L_1\cup L_2$, *причем* $L_1\cap L_2$ — *пустое или одноэлементное множество, то*

$$\int\limits_L P(x, y)\,dx+Q(x, y)\,dy=\int\limits_{L_1} P(x, y)\,dx+Q(x, y)\,dy+\int\limits_{L_2} P(x, y)\,dx+Q(x, y)\,dy$$

(*аддитивное свойство криволинейного интеграла*).

3^0. $\int\limits_{AB} P(x, y)\,dx+Q(x, y)\,dy=-\int\limits_{BA} P(x, y)\,dx+Q(x, y)\,dy.$

4^0. *Если* L — *отрезок прямой, параллельной оси* x, *то* $\int\limits_L Q(x, y)\,dy=0.$

5^0 *Если* L — *отрезок прямой, параллельной оси* y, *то* $\int\limits_L P(x, y)\,dx=0.$

Докажем, например, последнее свойство.

Если L — отрезок прямой, параллельной оси y, то при любом разбиении T отрезка L получим $\Delta x_k=0$, т. е. $\sum\limits_{k=0}^{n-1} P(\xi_k,\eta_k)\Delta x_k=0$, а, следовательно, и

$\lim\limits_{d_T\to 0}\sum\limits_{k=0}^{n-1} P(\xi_k, \eta_k)\Delta x_k=0$, т. е. $\int\limits_L P(x, y)\,dx=0$.

3. Вычисление криволинейного интеграла

Пусть дуга L задана условиями $y=y(x)$, $a\leqslant x\leqslant b$, причем $y'(x)$ непрерывна на $[a, b]$ (такая кривая называется гладкой, см. определение 3 из § 41), и пусть $P(x, y)$, $Q(x, y)$ — непрерывные функции на некотором множестве, содержащем кривую L внутри себя. Тогда для вычисления криволинейного интеграла $\int\limits_L P(x, y)dx+Q(x, y)dy$ достаточно заменить под знаком интеграла y на $y(x)$ и вычислить определенный интеграл по отрезку $[a, b]$. Аналогично обстоит дело в случае, если гладкая кривая L задана условиями $x=x(y)$, $c\leqslant y\leqslant d$.

Если гладкая дуга L задана параметрически уравнениями $x=x(t)$, $y=y(t)$, $\alpha\leqslant t\leqslant\beta$, то криволинейный интеграл сводится к определенному интегралу по отрезку $[\alpha, \beta]$ после замены под знаком интеграла переменных x и y их выражениями через t, т. е. $x(t)$, $y(t)$.

Примеры. 1. Вычислить $\int\limits_L 2xy\,dx+x^2dy$, где: а) L — дуга параболы $y=x^2$, $0\leqslant x\leqslant 2$;

б) L — дуга параболы $x=\frac{1}{8}y^2$, $0\leqslant y<4$.

○ а) $\int\limits_L 2xy\,dx+x^2dy=\int\limits_0^2 2x\cdot x^2\cdot dx+x^2\cdot d(x^2)=\int\limits_0^2 2x^3dx+2x^3dy=\int\limits_0^2 4x^3dx=x^4\Big|_0^2=16$;

б) $\int\limits_L 2xy\,dx+x^2dy=\int\limits_0^4 2\frac{1}{8}y^2\cdot y\,d\left(\frac{1}{8}y^2\right)+\left(\frac{1}{8}y^2\right)^2dy=\int\limits_0^4\left(\frac{y^4}{16}+\frac{y^4}{64}\right)dy=$

$=\frac{5}{64}\int\limits_0^4 y^4dy=\frac{1}{64}y^5\Big|_0^4=16.$

2. Вычислить $\int\limits_L y\,dx-x\,dy$, где L — единичная окружность с центром в начале координат.

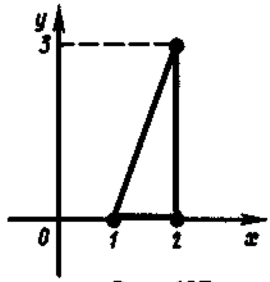


Рис. 187

○ Зададим L параметрически: $x=\cos t$, $y=\sin t$, $0\leqslant t\leqslant 2\pi$. Тогда

$$\int\limits_L y\,dx-x\,dy=\int\limits_0^{2\pi}\sin t\cdot d(\cos t)-\cos t\cdot d(\sin t)=$$

$$=\int\limits_0^{2\pi}\sin t(-\sin t\,dt)-\cos t(\cos t\,dt)=-\int\limits_0^{2\pi}(\sin^2 t+\cos^2 t)dt=-\int\limits_0^{2\pi}dt=-2\pi.●$$

3. Пусть в плоскости xy действует сила $\vec{F}(x,y)$ такая, что ее горизонтальная составляющая $P(x,y)=xy$, а вертикальная составляющая $Q(x,y)=x-y$. Материальная точка перемещается под действием этой силы из точки $(1;0)$ в точку $(2;3)$ по двум различным путям: а) по прямолинейному отрезку, соединяющему эти две точки (путь L_1); б) по ломаной, соединяющей точки $(1;0)$, $(2;0)$ и $(2;3)$ (путь L_2). Найти работу, совершенную силой $\vec{F}$ в каждом из этих случаев (рис. 187).

○ а) Имеем $A_1=\int\limits_L xy\,dx+(x-y)dy$. Составим уравнение прямой L_1, проходящей через точки $(1;0)$ и $(2;3)$: $\frac{x-1}{2-1}=\frac{y-0}{3-0}$, т. е. $y=3x-3$. Значит,

$$A_1=\int\limits_1^2 x(3x-3)dx+(x-3x+3)\cdot 3dx=\int\limits_1^2(3x^2-9x+9)dx=$$

$$=\left(x^3-\frac{9x^2}{2}+9x\right)\Big|_1^2=2{,}5.$$

б) Здесь $A_2=\int\limits_{L_2} xy\,dx+(x-y)dy$. Путь L_2 состоит из двух участков. Первый задается условиями $y=0$, $1\leqslant x\leqslant 2$; подставив $y=0$ под знак интеграла, получим, что интеграл по этому участку равен нулю. Второй участок задается условиями $x=2$, $0\leqslant y\leqslant 3$; интеграл по этому участку принимает вид $\int\limits_0^3(2-y)dy$ и равен 1,5. Воспользовавшись аддитивностью криволинейного интеграла, находим $A_2=0+1{,}5=1{,}5.$●

§ 49. Формула Грина и ее применения

1. Формула Грина

В настоящем пункте мы рассмотрим одну из основных формул интегрального исчисления функций двух переменных — формулу Грина*. Она связывает криволинейный интеграл по замкнутому контуру L, представляющему собой глад-

* Дж. Грин (1793—1841) — английский математик.

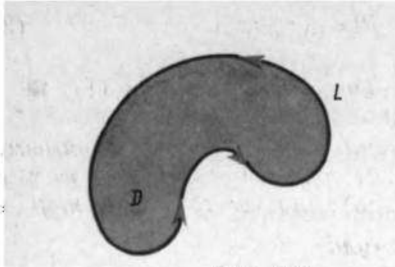


Рис. 188

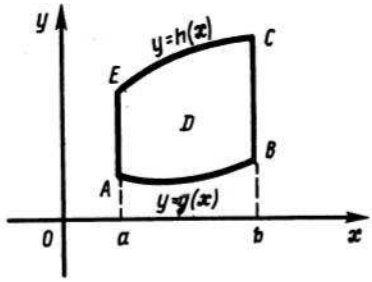


Рис. 189

кую или кусочно-гладкую кривую (т. е. объединение конечного числа гладких кривых) и двойной интеграл по области, ограниченной этим контуром.

В § 48 мы отмечали, что при изменении направления движения по дуге меняется и знак криволинейного интеграла; поэтому, связывая двойной интеграл по области D с криволинейным интегралом по контуру L этой области, нужно (во избежание недоразумений) условиться о направлении обхода контура L. Всюду в этом параграфе будем считать, что движение по замкнутому контуру L осуществляется так, что сама область D при этом движении все время остается слева (рис. 188).

Для вывода формулы Грина нам понадобятся два вспомогательных утверждения.

Лемма 1. *Пусть D — криволинейная трапеция, ограниченная прямыми $x=a$, $x=b$ $(a<b)$ и графиками гладких на $[a, b]$ функций $y=g(x)$, $y=h(x)$ таких, что $g(x)\leqslant h(x)$ на $[a, b]$* (рис. 189); *пусть $P(x, y)$ и $\frac{\partial P}{\partial y}$ непрерывны на D. Тогда справедлива формула*

$$\iint\limits_D \frac{\partial P}{\partial y}\,dx\,dy = -\int\limits_L P(x,y)\,dx. \qquad (1)$$

□ Имеем $\iint\limits_D \frac{\partial P}{\partial y}\,dx\,dy = \int\limits_a^b dx \int\limits_{g(x)}^{h(x)} \frac{\partial P}{\partial y}\,dy$. Так как первообразной от производной некоторой функции является сама эта функция, то $\int\limits_{g(x)}^{h(x)} \frac{\partial P}{dy}\,dy = P(x, y)\Big|_{g(x)}^{h(x)} = P(x, h(x)) - P(x, g(x))$. Следовательно,

$$\iint\limits_D \frac{\partial P}{\partial y}\,dx\,dy = \int\limits_a^b (P(x, h(x)) - P(x, g(x)))\,dx. \qquad (2)$$

Теперь рассмотрим правую часть доказываемого равенства (1) и воспользуемся аддитивным свойством криволинейного интеграла (см. § 48). Имеем

$$-\int\limits_L P(x, y)dx = -\Big(\int\limits_{AB} P(x, y)dx + \int\limits_{BC} P(x, y)dx + \int\limits_{CE} P(x, y)dx + \int\limits_{EA} P(x, y)dx\Big) =$$

$$= -\Big(\int\limits_a^b P(x, g(x))dx + 0 - \int\limits_a^b P(x, h(x)dx + 0\Big) = \int\limits_a^b P(x, h(x))dx -$$

$$-\int\limits_a^b P(x, g(x))dx = \int\limits_a^b P(x, h(x)) - P(x, g(x))dx.$$

Итак,

$$-\int_L P(x,y)dx = \int_a^b (P(x,h(x)) - P(x,g(x)))dx\,. \qquad (3)$$

Сопоставив равенства (2) и (3), получаем требуемое соотношение (1). ■

Аналогично доказывается следующая лемма.

Лемма 2. *Пусть D — криволинейная трапеция, ограниченная прямыми $y=c$, $y=d$ $(c<d)$ и графиками гладких на $[c,d]$ функций $x=\varphi(y)$, $x=\psi(y)$ таких, что $\varphi(y)\leqslant\psi(y)$ на $[c,d]$* (см. рис. 183); *пусть также $Q(x,y)$ и $\frac{\partial Q}{\partial x}$ непрерывны на D. Тогда справедлива формула*

$$\iint_D \frac{\partial Q}{\partial x}dxdy = \int_L Q(x,y)dy. \qquad (4)$$

Криволинейную трапецию, о которой шла речь в лемме 1, условимся ради краткости называть *криволинейной трапецией первого вида*, а в лемме 2 — *криволинейной трапецией второго вида.*

Следствие. *Формулы* (1) *и* (4) *справедливы и в случае, если D представляет собой объединение конечного числа криволинейных трапеций первого или второго вида без общих внутренних точек.*

□ Приведем в качестве примера схему доказательства формулы (1) для случая объединения двух трапеций первого вида (рис. 190). Имеем

$$\iint_D = \iint_{D_1} + \iint_{D_2} = -\int_{ACC_1A_1A} - \int_{CBB_1C_1C} =$$

$$= -\left(\int_{AC} + \int_{CC_1} + \int_{C_1A_1} + \int_{A_1A}\right) - \left(\int_{CB} + \int_{BB_1} + \int_{B_1C_1} + \int_{C_1C}\right) =$$

$$= -\left(\int_{AC} + 0 + \int_{C_1A_1} + 0 + \int_{CB} + 0 + \int_{B_1C_1} + 0\right) = -\int_{ABB_1A_1A} = -\int_L. \ ■$$

Теорема 1. *Пусть D — фигура, которую можно представить в виде объединения конечного числа криволинейных трапеций первого или второго вида, и пусть функции $P(x,y)$, $Q(x,y)$, $\frac{\partial P}{\partial y}$, $\frac{\partial Q}{\partial x}$ непрерывны на D. Тогда*

$$\int_L P(x,y)dx + Q(x,y)dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dxdy. \qquad (5)$$

Равенство (5) называется *формулой Грина.*

□ При указанных условиях справедливы как равенство (1), так и равенство (4). Сложив их, получим требуемое соотношение (5). ■

Пример. Используя формулу Грина, вычислить

$$\int_L \left(xy - x + \frac{1}{2}y^2 - y\right)dx + \left(xy + \frac{1}{2}x^2 - y\right)dy,$$

где L — контур параболического сегмента $y = x^2 - 3x - 2$, $y = x + 3$ (рис. 191).

○ Имеем $P = xy - x + \frac{1}{2}y^2 - y$, $Q = xy + \frac{1}{2}x^2 - y$,

$$\frac{\partial P}{\partial y} = x + y - 1,\ \frac{\partial Q}{\partial x} = y + x,\ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1.$$

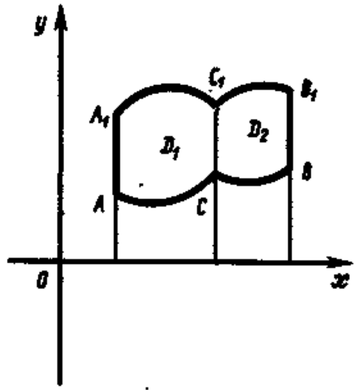


Рис. 190

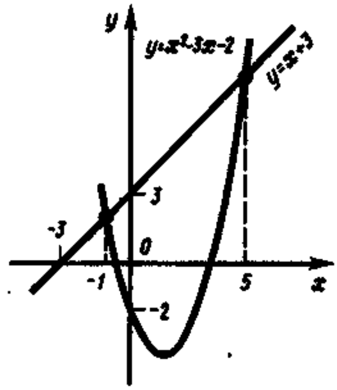


Рис. 191

Тогда

$$\int\limits_L \left(xy - x + \frac{1}{2}y^2 - y\right)dx + \left(xy + \frac{1}{2}x^2 - y\right)dy = \iint\limits_D dx\,dy.$$

Чтобы расставить пределы интегрирования в двойном интеграле, надо найти координаты точек пересечения параболы и прямой. Имеем $x^2 - 3x - 2 = x + 3$; $x^2 - 4x - 5 = 0$; $x_1 = -1$, $x_2 = 5$. Значит,

$$\iint\limits_D dx\,dy = \int\limits_{-1}^{5} dx \int\limits_{x^2-3x-2}^{x+3} dy = \int\limits_{-1}^{5} dx \cdot y\Big|_{x^2-3x-2}^{x+3} =$$

$$= \int\limits_{-1}^{5}(x + 3 - x^2 + 3x + 2)dx = \int\limits_{-1}^{5}(-x^2 + 4x + 5)dx =$$

$$= \left(-\frac{x^3}{3} + 2x^2 + 5x\right)\Big|_{-1}^{5} = \left(-\frac{125}{3} + 50 + 25\right) - \left(\frac{1}{3} + 2 - 5\right) = 36. \bullet$$

2. Условие независимости криволинейного интеграла от пути интегрирования

Выше (см. пример 1 п. 3 § 48) мы вычислили криволинейный интеграл $\int\limits_L 2xy\,dx + x^2dy$ по двум дугам: $y = x^2$, $0 \leqslant x \leqslant 2 (L_1)$ и $x = \frac{1}{8}y^2$, $0 \leqslant y \leqslant 4 (L_2)$. Это два различных пути интегрирования, но начало и конец пути в обоих случаях одинаковы: начало — точка $(0; 0)$, конец — точка $(2; 4)$. Оказалось, что $\int\limits_{L} 2xy\,dx + x^2dy = \int\limits_{L_2} 2xy\,dx + x^2dy = 16$. Это не случайно; более того, если взять любую гладкую дугу L, соединяющую точки $(0;0)$ и $(2;4)$, то $\int\limits_L 2xy\,dx + x^2dy = 16$, т. е. значение интеграла не зависит от выбора пути интегрирования, оно зависит только от начала и конца дуга. Так бывает далеко не всегда (см. пример 3 п. 3 § 48), а только при определенных условиях, о которых речь идет ниже.

Теорема 2. *Пусть функции $P(x, y)$, $Q(x, y)$, $\frac{\partial P}{\partial y}$, $\frac{\partial Q}{\partial x}$ непрерывны на связном множестве E. Для того чтобы интеграл $\int_L P(x, y)\,dx + Q(x, y)\,dy$, где L — произвольная гладкая или кусочно-гладкая замкнутая кривая, лежащая внутри E, был равен нулю, необходимо и достаточно, чтобы во всех внутренних точках множества E выполнялось равенство $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.*

□ Необходимость. Пусть для любого замкнутого контура L выполняется равенство $\int_L P(x, y)dx + Q(x, y)dy = 0$. Докажем, что тогда $\frac{\partial P}{dy} = \frac{\partial Q}{\partial x}$ всюду в E.

Предположим противное, т. е. что в некоторой внутренней точке $(x_0; y_0) \in E$ равенство не выполняется; пусть для определенности в этой точке $\frac{\partial P}{\partial y} < \frac{\partial Q}{\partial x}$. Поскольку функция $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ непрерывна в точке $(x_0; y_0)$ и положительна в этой точке, по теореме 7 § 44 существует такая окрестность точки $(x_0; y_0)$, в которой $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} > 0$. Эта окрестность представляет собой круг радиуса r с центром в точке $(x_0; y_0)$. Пусть D — этот круг, а L — окружность круга. Тогда по формуле Грина имеем

$$\int_L P(x, y)dx + Q(x, y)dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Так как $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} > 0$ в D, то $\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy > 0$ (см. свойство 4^0 из п. 5 § 47). Но по условию $\int_L P(x, y)dx + Q(x, y)dy = 0$, а значит, и $\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = 0$. Получили противоречие, которое показывает, что сделанное предположение неверно. Таким образом, $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ всюду в E.

Достаточность. Пусть $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ всюду в E и L — любой замкнутый контур, лежащий внутри E. Тогда интеграл $\int_L P(x, y)dx + Q(x, y)dy$ равен нулю, что непосредственно следует из формулы Грина. ■

Теорема 3. *Пусть функции $P(x, y)$, $Q(x, y)$, $\frac{\partial P}{\partial y}$, $\frac{\partial Q}{\partial x}$ непрерывны на связном множестве E и пусть A и B — любые две точки из E, а L_1 и L_2-различные гладкие или кусочно-гладкие дуги, соединяющие точки A и B и лежащие внутри E. Для того чтобы интегралы $\int_{L_1} P(x,y)dx + Q(x,y)dy$ и $\int_{L_2} P(x, y)dx + Q(x, y)dy$ были равны, необходимо и достаточно, чтобы выполнялось равенство $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.*

□ Необходимость. Пусть $L=ACBD$ — произвольный замкнутый контур, лежащий внутри E; A, B — две точки на этом контуре (рис. 192). По условию, $\int\limits_{L_1} P(x,y)dx+Q(x,y)dy=\int\limits_{L_2} P(x,y)dx+Q(x,y)dy$, т. е. $\int\limits_{ACB}=\int\limits_{ADB}$. Но тогда $\int\limits_{L}=\int\limits_{ACB}+\int\limits_{BDA}=\int\limits_{ACB}-\int\limits_{ADB}=0$. Значит, интеграл по любому замкнутому контуру равен нулю, откуда, согласно теореме 2, следует, что $\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$.

Достаточность. Пусть L_1 и L_2 — две дуги, соединяющие точки A и B и образующие в совокупности замкнутый контур L (рис. 193). Так как $\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$, то в силу теоремы 2 получаем $\int\limits_{L}=0$, откуда и следует, что $\int\limits_{L_1}=\int\limits_{L_2}$. ■

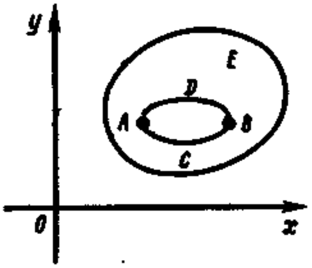


Рис. 192

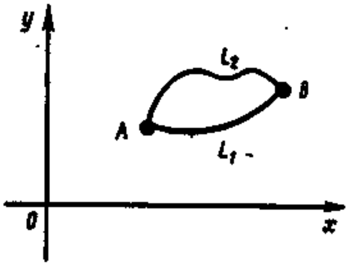


Рис. 193

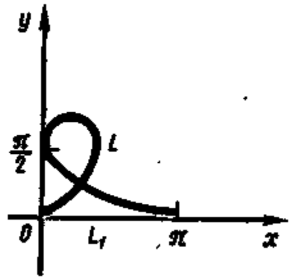


Рис. 194

Пример. Вычислить $\int\limits_{L} 2xydx+x^2dy$, где дуга L задана параметрически:

$$\begin{cases} x=t\cos^2 t \\ y=t(\cos t+1), \end{cases} \qquad 0\leqslant t\leqslant \pi.$$

○ Здесь $P(x,y)=2xy$, $Q(x,y)=x^2$, $\frac{\partial P}{\partial y}=2x$, $\frac{\partial Q}{\partial x}=2x$; следовательно, $\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}$ и вместо интеграла по пути L можно вычислить интеграл по любому пути L_1, более простому, чем заданный, и соединяющему начало и конец дуги L. Для дуги L имеем: если $t=0$, то $x=0$, $y=0$, а если $t=\pi$, то $x=\pi$, $y=0$. Значит, дуга L соединяет точки $(0;0)$ и $(\pi;0)$ (рис. 194). Самый простой путь L_1, соединяющий эти точки, такой: $y=0$, $0\leqslant x\leqslant\pi$. Подставив $y=0$ в интеграл $\int 2xydx+x^2dy$, получаем, что он равен нулю. Поэтому равен нулю и заданный интеграл. ●

3. Восстановление функции по ее полному дифференциалу

Если дана функция $z=f(x,y)$, то ее дифференциал нетрудно найти по формуле $dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy$, т. е. $dz=P(x,y)dx+Q(x,y)dy$, где $P(x,y)=\frac{\partial z}{\partial x}$, а $Q(x,y)=\frac{\partial z}{\partial y}$. Труднее решается обратная задача: зная, что $dz=P(x,y)dx+Q(x,y)dy$, найти первообразную $z=f(x,y)$, т. е. функцию, для которой $P(x,y)=$

$= \frac{\partial z}{\partial x}$, а $Q(x, y) = \frac{\partial z}{\partial y}$. Более того, далеко не всегда выражение $P(x, y)dx + Q(x, y)dy$ является дифференциалом некоторой функции (для сравнения заметим, что в случае функции одной переменной аналогичная ситуация проще: практически всегда выражение $P(x)dx$ является дифференциалом некоторой функции; чтобы найти ее, достаточно проинтегрировать выражение $P(x)dx$).

Определение. Выражение $P(x, y)dx + Q(x, y)dy$ называется *полным дифференциалом,* если существует функция $z = f(x, y)$ такая, что $dz = P(x, y)dx + Q(x, y)dy$; эта функция $z = f(x, y)$ называется *первообразной* для полного дифференциала.

Таким образом, нам нужно научиться отвечать на два вопроса:

1) как узнать, является ли выражение $P(x, y)dx + Q(x, y)dy$ полным дифференциалом?

2) если $P(x, y)dx + Q(x, y)dy$ полный дифференциал, то как найти первообразную?

Ответы на оба вопроса мы получим, доказав следующую теорему.

Теорема 4. *Пусть функции $P(x,y)$, $Q(x,y)$, $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ непрерывны на связном множестве E. Для того чтобы выражение $P(x,y)dx + Q(x,y)dy$ было полным дифференциалом, необходимо и достаточно, чтобы всюду в E выполнялось равенство $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.*

Необходимость. Пусть $P(x,y)dx + Q(x,y)dy$ — полный дифференциал. Тогда существует функция $z = f(x,y)$ такая, что $\frac{\partial z}{\partial x} = P(x,y)$, а $\frac{\partial z}{\partial y} = Q(x,y)$.

Так как $\frac{\partial P}{\partial y} = \frac{\partial^2 z}{\partial x \partial y}$, а $\frac{\partial Q}{\partial x} = \frac{\partial^2 z}{\partial y \partial x}$, то, используя теорему о равенстве смешанных производных (см. теорему 6 § 46), получим $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

Достаточность. Пусть $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ всюду в E. Докажем, что существует функция, дифференциал которой равен $P(x, y)dx + Q(x, y)dy$.

Возьмем внутри E две точки: фиксированную точку $A(x_0; y_0)$ и текущую точку $M(x,y)$. Так как $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, то по теореме 3 интеграл $\int\limits_{AM} P(x,y)dx + Q(x, y)\,dy$ не зависит от выбора дуги, соединяющей точки A и M, он зависит только от точки M (напомним, что точка A фиксирована). Значит, этот интеграл есть функция точки M, т. е. зависит от ее координат $(x; y)$. Обозначим $v(x, y) = \int\limits_{AM} P(x, y)\,dx + Q(x, y)\,dy$.

Докажем, что $v(x, y)$ — искомая первообразная, т. е. что $\frac{\partial v}{\partial x} = P(x, y)$, а $\frac{\partial v}{\partial y} = Q(x, y)$.

Найдем $\frac{\partial v}{\partial x}$, воспользовавшись обычным алгоритмом нахождения производной (см. п. 3 § 28).

1°. Для точки $M(x, y)$ имеем $v(x, y) = \int\limits_{AM} P(x, y)dx + Q(x, y)dy$.

2°. Дадим аргументу x приращение Δx, получим новую точку $N(x + \Delta x; y)$ и найдем $v(x + \Delta x; y) = \int\limits_{AMN} = \int\limits_{AM} + \int\limits_{MN}$ (рис. 195).

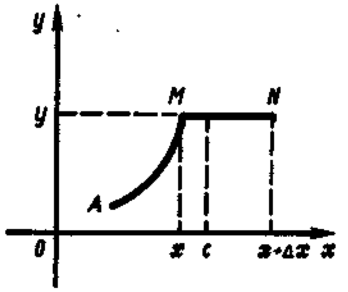


Рис. 195

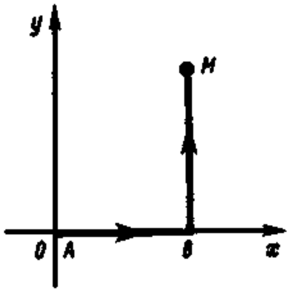


Рис. 196

3^0. $\Delta_x v = v(x+\Delta x, y) - v(x, y) = \left(\int\limits_{AM} + \int\limits_{MN}\right) - \int\limits_{AM} = \int\limits_{MN} P(x,y)dx + Q(x,y)dy.$

Но на горизонтальном отрезке MN имеем $\int\limits_{MN} Q(x, y)dy = 0$ (см. свойство 4^0 из § 48). Значит, $\Delta_x v = \int\limits_{MN} P(x, y)dx = \int\limits_x^{x+\Delta x} P(x, y)dx$ (здесь $y = \text{const}$).

Воспользовавшись для определенного интеграла $\int\limits_x^{x+\Delta x} P(x, y)dx$ теоремой о среднем (см. теорему 4 § 39), получим $\int\limits_x^{x+\Delta x} P(x,y)dx = P(c,y)\Delta x$, где $c \in [x, x+\Delta x]$.

4^0. $\frac{\Delta_x v}{\Delta x} = \frac{P(c,y)\Delta x}{x} = P(c, y).$

5^0. $\lim\limits_{\Delta x \to 0} \frac{\Delta_x v}{\Delta x} = \lim\limits_{\Delta x \to 0} P(c,y) = \lim\limits_{c \to x} P(c,y) = P(x,y)$ (так как $P(x,y)$ непрерывная функция).

Итак, $\lim\limits_{\Delta x \to 0} \frac{\Delta_x v}{\Delta x} = P(x, y)$, а это и означает, что $\frac{\partial v}{\partial x} = P(x, y)$.

Аналогично доказывается, что $\frac{\partial v}{\partial y} = Q(x, y)$. ■

Каковы же ответы на поставленные выше вопросы?

1. Чтобы узнать, является ли выражение $P(x, y)dx + Q(x, y)dy$ полным дифференциалом, достаточно проверить выполнение тождества $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ $\left(\text{при условии непрерывности функций } P(x, y),\ Q(x, y),\ \frac{\partial P}{\partial x},\ \frac{\partial Q}{\partial y}\right)$.

2. Чтобы найти первообразную для полного дифференциала $P(x, y)dx + Q(x, y)dy$, достаточно вычислить $\int\limits_{AM} P(x, y)dx + Q(x, y)dy$, где AM — любая гладкая или кусочно-гладкая дуга, соединяющая выбранную фиксированную точку A с текущей точкой $M(x; y)$. В качестве фиксированной точки A можно взять любую точку, в которой функции $P(x, y)$ и $Q(x, y)$ непрерывны (обычно берут одну из точек $(0; 0)$, $(1; 0)$, $(0; 1)$, $(1; 1)$).

Пример. Доказать, что $(2x\cos y-y^2\sin x)dx+(2y\cos x-x^2\sin y)dy$ является полным дифференциалом, и найти первообразную.

○ Имеем $P(x, y)=2x\cos y-y^2\sin x$, $Q(x, y)=2y\cos x-x^2\sin y$, $\dfrac{\partial P}{\partial y}=-2x\sin y-$ $-2y\sin x$, $\dfrac{\partial Q}{\partial x}=-2y\sin x-2x\sin y$. Так как $\dfrac{\partial P}{\partial y}=\dfrac{\partial Q}{\partial x}$, то заданное выражение есть полный дифференциал.

В качестве фиксированной точки A здесь можно взять точку $A\ (0;0)$, а в качестве пути интегрирования AM — путь, изображенный на рис. 196; тогда $\int\limits_{AM}=\int\limits_{AB}+\int\limits_{BM}$.

Отрезок AB описывается условиями $y=0$, $0\leqslant x\leqslant x$ (эта запись означает, что переменная x пробегает все значения от 0 до некоторого фиксированного значения x). Тогда получим

$$\int\limits_{AB}(2x\cos y-y^2\sin x)dx+(2y\cos x-x^2\sin y)dy=\int\limits_0^x 2x\,dx$$

(мы воспользовались тем, что $y=0$). Далее, находим $\int\limits_0^x 2x\,dx=x^2\Big|_0^x=x^2$.

Отрезок BM описывается условиями $x=x$ (эта запись означает, что $x=\text{const}$), $0\leqslant y\leqslant y$ (эта запись означает, что переменная y пробегает все значения от 0 до некоторого фиксированного значения y). Тогда получим

$$\int\limits_{BM}(2x\cos y-y^2\sin x)dx+(2y\cos x-x^2\sin y)dy=\int\limits_0^y(2y\cos x-x^2\sin y)dy$$

(мы воспользовались тем, что $x=\text{const}$ и, следовательно, $dx=0$). Далее, находим

$$\int\limits_0^y(2y\cos x-x^2\sin y)dy=(y^2\cos x+x^2\cos y)\Big|_0^y=(y^2\cos x+x^2\cos y)-$$

$$-(\cos x\cdot 0+x^2\cdot 1)=y^2\cos x+x^2\cos y-x^2.$$

В результате получаем

$$v(x, y)=\int\limits_{AM}=\int\limits_{AB}+\int\limits_{BM}=x^2+(y^2\cos x+x^2\cos y-x^2)=y^2\cos x+x^2\cos y.$$

Это и есть искомая первообразная. ●

8

Числовые и функциональные ряды

При решении многих задач математического анализа приходится рассматривать суммы, составленные из бесконечного количества слагаемых. Из теории действительных чисел известно, что значит сложить любые два числа, а также три, четыре числа и т. д. — вообще, что означает сумма любого конечного множества чисел. Однако что следует понимать под суммой бесконечного множества чисел (а в более сложных ситуациях — уже не чисел, а скажем, функций, векторов, матриц и т. п.)? Этого мы пока не знаем. Между тем задача суммирования бесконечного множества каких-то однотипных объектов (чисел, функций и т. п.) постоянно встречается в математике. Эта задача и решается в теории рядов, составляющей одну из важных глав курса математического анализа.

§ 50. Числовые ряды. Сходимость ряда

1. Числовой ряд. Общий член ряда

Пусть задана бесконечная последовательность чисел a_1, a_2, a_3, ... Рассмотрим выражение

$$a_1 + a_2 + a_3 + \dots, \tag{1}$$

представляющее собой «сумму бесконечного множества слагаемых». Оно называется *числовым рядом*, а сами числа a_1, a_2, a_3, ... — *членами* ряда.

Член ряда a_n с произвольным (неопределенным) номером n называется *общим членом*. Ясно, что a_n есть некоторая функция от n. Эту функцию необходимо указать, так как иначе нельзя считать заданной саму последовательность a_1, a_2, a_3, ... Однако чаще всего выражение для общего члена ряда (как функции от n) не указывают отдельно, а вносят в ряд и вместо (1) пишут

$$a_1 + a_2 + a_3 + \dots + a_n + \dots . \tag{2}$$

Например,

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$$

есть ряд с общим членом $a_n = \frac{1}{n}$, а

$$\frac{1}{1\cdot 2} - \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} - \ldots + (-1)^{n-1}\frac{1}{n(n+1)} + \ldots$$

— ряд с общим членом $a_n = \frac{(-1)^{n-1}}{n(n+1)}$.

2. Сходящиеся и расходящиеся ряды. Сумма ряда

Пусть дан ряд (2). Введем понятие *частичных сумм* ряда. Так называются числа

$$A_1 = a_1,$$
$$A_2 = a_1 + a_2,$$
$$A_3 = a_1 + a_2 + a_3$$

и т. д. Вообще, n-я частичная сумма A_n есть сумма первых n членов ряда:

$$A_n = a_1 + a_2 + \ldots + a_n.$$

Определение 1. Если последовательность A_n частичных сумм имеет предел, т. е. существует число $A = \lim\limits_{n\to\infty} A_n$, то ряд называется *сходящимся*, а число A называется *суммой ряда*. В этом случае также говорят, что *ряд сходится к сумме* A и пишут

$$a_1 + a_2 + \ldots + a_n + \ldots = A.$$

Если же $\lim\limits_{n\to\infty} A_n$ не существует, то говорят, что ряд *расходится* или что он *не имеет суммы*.

Фактически с понятием суммы ряда мы уже встречались при выводе формулы суммы «всех» членов бесконечной убывающей геометрической прогрессии. Напомним, что бесконечная геометрическая прогрессия, т. е. ряд

$$b + bq + bq^2 + \ldots + bq^{n-1} + \ldots$$

сходится лишь при $|q| < 1$; в этом случае его сумма равна $\frac{b}{1-q}$.

Приведем пример сходящегося ряда, не являющегося прогрессией. Рассмотрим ряд

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n(n+1)} + \cdots$$

Так как $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, то для n-й частичной суммы ряда получаем выражение

$$A_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

После раскрытия скобок все слагаемые, кроме первого и последнего, взаимно уничтожаются и в результате получаем $A_n = 1 - \frac{1}{n+1}$. Отсюда следует, что $\lim\limits_{n\to\infty} A_n = 1$. Итак, ряд сходится и его сумма равна 1.

3. Общие свойства сходящихся рядов

1°. *Если ряд $a_1 + a_2 + \ldots + a_n + \ldots$ сходится, то сходится и любой ряд, полученный из него отбрасыванием конечного числа членов.*

□ Ограничимся случаем, когда отбрасываемые члены имеют номера 1, 2, ..., p (в принципе можно отбрасывать члены с какими угодно номерами, лишь бы только их было конечное число). Будем называть оставшийся ряд $a_{p+1} + a_{p+2} + \ldots$ «укороченным». При $n>p$, очевидно, имеем

$$A_n = (a_1 + \ldots + a_p) + (a_{p+1} + \ldots + a_n) = A_p + A^*_{n-p}, \tag{3}$$

где A^*_{n-p} — частичная сумма укороченного ряда. Из этого равенства следует, что если существует $\lim\limits_{n\to\infty} A_n$, то существует и $\lim\limits_{n\to\infty} A^*_{n-p}$, и обратно. Это доказывает справедливость свойства 1°. ■

Замечание. Из равенства (3) попутно вытекает равенство $A = A_p + A^*$, где A — сумма исходного ряда, A^* — сумма укороченного ряда и A_p — сумма отбрасываемых членов.

Так как сумма ряда определяется как предел последовательности его частичных сумм, то многие свойства суммы ряда автоматически следуют из соответствующих свойств пределов. Укажем некоторые из них, используемые в дальнейшем.

Пусть даны два ряда

$$a_1 + a_2 + \ldots + a_n + \ldots, \tag{4}$$

$$b_1 + b_2 + \ldots + b_n + \ldots. \tag{5}$$

Составим из них новый ряд

$$(a_1 + b_1) + (a_2 + b_2) + \ldots + (a_n + b_n) + \ldots. \tag{6}$$

Говорят, что ряд (6) получен *почленным сложением* рядов (4) и (5). При этом имеет место следующее свойство сходимости.

2°. *Если оба ряда* (4) и (5) *сходятся, а их суммы соответственно равны A и B, то сходится и ряд* (6), *причем его сумма равна $A + B$.*

□ Очевидно, что при любом n имеем $C_n = A_n + B_n$, где A_n, B_n, C_n — соответственно частичные суммы рядов (4), (5), (6). Переходя к пределу при $n\to\infty$, находим, что существует число $C = \lim\limits_{n\to\infty} C_n$ и что $C = A + B$. ■

3°. *Если ряд $a_1 + a_2 + \ldots + a_n + \ldots$ сходится и его сумма равна A, то сходится и ряд $\lambda a_1 + \lambda a_2 + \ldots + \lambda a_n + \ldots$,* где λ — *любое число, причем его сумма равна λA.*

Рекомендуем доказать это свойство самостоятельно.

4°. *Если ряд $a_1 + a_2 + \ldots + a_n + \ldots$ сходится, то сходится и любой ряд, полученный из него группировкой слагаемых, причем суммы обоих рядов одинаковы.*

Например, сходится ряд

$$(a_1 + a_2) + (a_3 + a_4) + \ldots + (a_{2n-1} + a_{2n}) + \ldots,$$

полученный из данного ряда группировкой членов по два. Впрочем, необязательно составлять группы из одного и того же числа членов; количество членов, объединяемых в n-ю группу, может быть произвольной функцией от n.

□ Назовем ряд, полученный из данного группировкой членов, «группированным» рядом. Очевидно, любая частичная сумма A^*_k группированного ряда совпадает с какой-то частичной суммой A_{n_k} исходного ряда:

$$A^*_k = A_{n_k}. \tag{7}$$

При этом если $k \to \infty$, то и $n_k \to \infty$ (поскольку $n_k \geqslant k$). Поэтому из равенства (7) в пределе при $k \to \infty$ получаем $A^* = A$, т. е. сумма исходного ряда совпадает с суммой группированного ряда. ■

Определение 2. Пусть дан ряд $a_1 + a_2 + \ldots + a_n + \ldots$. Ряд

$$a_{k+1} + a_{k+2} + \ldots + a_{k+n} + \ldots , \qquad (8)$$

где k — произвольное натуральное число, называется *остатком данного ряда после k-го члена* или просто *k-м остатком*.

Сумму ряда (8) (если она существует) будем, как правило, обозначать R_k. Из свойства 1^0 сходящихся рядов и замечания к нему вытекает следующее свойство.

5^0. *Если сходится ряд $a_1 + a_2 + \ldots + a_n + \ldots$, то сходится и его остаток* (8), *причем их суммы связаны соотношением $A = a_1 + \ldots + a_k + R_k$.*

4. Необходимый признак сходимости ряда

В приложениях обычно применяются лишь сходящиеся ряды. Поэтому важно знать признаки, по которым можно было бы судить, сходится данный ряд или нет. Установим сначала необходимый признак сходимости.

Теорема (необходимый признак сходимости). *Если ряд сходится, то его n-й член стремится к нулю при $n \to \infty$.*

В краткой форме: *если ряд сходится, то его общий член стремится к нулю.*

Пусть ряд сходится и его сумма равна A. Для любого номера n имеем $A_n = A_{n-1} + a_n$ или

$$a_n = A_n - A_{n-1} . \qquad (9)$$

Так как при $n \to \infty$ обе частичные суммы A_n и A_{n-1} стремятся к пределу A, то из равенства (9) следует $\lim\limits_{n\to\infty} a_n = A - A = 0$.

Подчеркнем, еще раз, что мы установили лишь необходимый признак сходимости, т. е. такой, при нарушении которого ряд не может сходиться. С помощью этого признака можно доказывать лишь расходимость ряда. Так, например, для ряда

$$\frac{1}{2} + \frac{2}{3} + \ldots + \frac{n}{n+1} + \ldots$$

указанный признак не выполняется (здесь $\lim\limits_{n\to\infty} a_n = 1$), и, следовательно, ряд расходится.

Приведем пример, из которого видно, что условие $\lim\limits_{n\to\infty} a_n = 0$ не является достаточным для сходимости ряда. Рассмотрим ряд

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \ldots$$

(выражение для n-го члена не указываем, оно выглядит громоздко). Его первый член равен 1, следующие два члена равны $\frac{1}{2}$, следующие три члена равны $\frac{1}{3}$ и т. д. Очевидно, n-й член ряда стремится к нулю при $n \to \infty$, т. е. для ряда выполнен необходимый признак сходимости. Тем не менее ряд расходится. Действительно, если бы ряд сходился, то сходился бы и ряд

$$1 + \left(\frac{1}{2} + \frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{3}\right) + \ldots ,$$

полученный из данного группировкой одинаковых членов. Но последний ряд есть $1+1+1+\ldots$, а значит, он расходится.

§ 51. Ряды с положительными членами. Признаки сходимости

В настоящем параграфе рассматриваются только ряды $a_1+a_2+\ldots+a_n+$ $+\ldots$, все члены которых положительны. Для таких рядов вопросы сходимости и расходимости решаются особенно просто.

1. Критерий сходимости

Так как все члены ряда положительны, т. е. $a_n>0$ для $n=1, 2, \ldots$, то последовательность частичных сумм монотонно возрастает: $A_1<A_2<\ldots<A_n<\ldots$. Как известно (см. § 13), для существования предела монотонно возрастающей последовательности необходимо и достаточно, чтобы эта последовательность была ограничена сверху, т. е. чтобы существовало число M такое, что $A_n<M$ для всех $n=1, 2, \ldots$.

Отсюда получаем следующий критерий (т. е. необходимый и достаточный признак) сходимости.

Теорема 1 (критерий сходимости ряда с положительными членами). *Для сходимости ряда с положительными членами необходимо и достаточно, чтобы все частичные суммы ряда были ограничены сверху одним и тем же числом.*

Указанный критерий имеет скорее теоретическое, чем практическое значение. В последующих пунктах мы приведем ряд других признаков сходимости; в отличие от указанного выше необходимого и достаточного признака все они являются только достаточными.

2. Достаточные признаки сходимости

Теорема 2 (первый признак сравнения). *Пусть даны два ряда с положительными членами:*

$$a_1+a_2+\ldots+a_n+\ldots, \tag{1}$$

$$b_1+b_2+\ldots+b_n+\ldots, \tag{2}$$

причем члены первого ряда не превосходят соответствующих членов второго: $a_n \leqslant b_n$ ($n=1, 2, \ldots$). Тогда из сходимости ряда (2) *(«большего») вытекает сходимость ряда* (1) *(«меньшего»).*

□ Очевидно, что для любого $n=1, 2, \ldots$ имеем

$$A_n \leqslant B_n, \tag{3}$$

где A_n и B_n — соответственно частичные суммы рядов (1) и (2). Если ряд (2) сходится, то существует число $B=\lim\limits_{n\to\infty} B_n$. Поскольку при этом последовательность $\{B_n\}$ — возрастающая, ее предел больше любого из ее членов, т. е. $B_n<B$ для любого n. Отсюда и из неравенства (3) следует $A_n<B$. Таким образом, все частичные суммы ряда (1) ограничены сверху числом B. Согласно теореме 1 этот ряд сходится. ■

Следствие. *В условиях первого признака сравнения из расходимости ряда* (1) *следует расходимость ряда* (2).

Действительно, если бы ряд (2) сходился, то по признаку сравнения сходился бы и ряд (1).

Примеры. 1. Рассмотрим ряд

$$\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{(n+1)^2}+\ldots. \tag{4}$$

Его можно сравнить с рядом

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n(n+1)}+\ldots,$$

сходимость которого была доказана раньше (см. § 50, п. 2). Так как $\frac{1}{(n+1)^2}<$ $<\frac{1}{n(n+1)}$, то сходится и ряд (4).

2. Доказать сходимость ряда с общим членом $a_n=\frac{1}{n^\alpha}$, где $\alpha\geqslant 2$.

○ Случай $\alpha=2$ рассмотрен в примере 1. При $\alpha>2$ имеем $\frac{1}{n^\alpha}<\frac{1}{n^2}$ $(n=1, 2, \ldots)$; следовательно, по признаку сравнения ряд $1+\frac{1}{2^\alpha}+\cdots+\frac{1}{n^\alpha}+\cdots$ сходится.●

Теорема 3 (второй признак сравнения). *Если для рядов* (1) *и* (2) *с положительными членами существует отличный от нуля предел*

$$\lim_{n\to\infty}\frac{a_n}{b_n}=u, \tag{5}$$

то ряды сходятся или расходятся одновременно.

□ Пусть ряд (2) сходится; докажем, что тогда сходится и ряд (1). Выберем какое-нибудь число v, большее, чем u. Из условия (5) вытекает существование такого номера n_0, что для всех $n>n_0$ справедливо неравенство $a_n/b_n<v$, или, что то же,

$$a_n<vb_n. \tag{6}$$

Отбросив в рядах (1) и (2) первые n_0 членов (что не влияет на сходимость), можно считать, что неравенство (6) справедливо для всех $n=1, 2, \ldots$. Но ряд с общим членом vb_n сходится в силу сходимости ряда (2). Согласно первому признаку сравнения, из неравенства (6) следует сходимость ряда (1).

Пусть теперь сходится ряд (1); докажем сходимость ряда (2). Для этого следует просто поменять ролями заданные ряды. Так как

$$\lim_{n\to\infty}\frac{b_n}{a_n}=\frac{1}{u}\neq 0,$$

то, по доказанному выше, из сходимости ряда (1) должна следовать сходимость ряда (2). ■

Примеры. 3. Ряд

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\cdots \tag{7}$$

называется *гармоническим*. Доказать, что он расходится, используя для сравнения ряд

$$\ln(1+1)+\ln\left(1+\frac{1}{2}\right)+\ln\left(1+\frac{1}{3}\right)+\cdots+\ln\left(1+\frac{1}{n}\right)+\cdots. \tag{8}$$

○ Покажем сначала, что ряд (8) расходится. Для этого заметим, что его n-й член равен $\ln\frac{n+1}{n}=\ln(n+1)-\ln n$. Исходя из этого, можно найти частичные суммы ряда (8):

$$A_n=(\ln 2-\ln 1)+(\ln 3-\ln 2)+\dots+(\ln(n+1)-\ln n)=\ln(n+1)-\ln 1=\ln(n+1).$$

Отсюда следует, что $A_n\to\infty$ при $n\to\infty$, т. е. ряд (8) расходится.

Применим теперь к рядам (7) и (8) второй признак сравнения. Используя известный предел $\lim\limits_{n\to\infty}\frac{\ln\left(1+\frac{1}{n}\right)}{\frac{1}{n}}=1$ (см. теорему 3 § 27), заключаем, что из расходимости ряда (8) вытекает расходимость ряда (7). ●

4. Доказать сходимость ряда с общим членом $a_n=\frac{3n+1}{n^3+5n-1}$.

○ Если в числителе и знаменателе выражения для a_n оставить только старшие (относительно n) члены, то получим число $b_n=\frac{3n}{n^3}=\frac{3}{n^2}$. Ряд с общим членом b_n сходится (см. пример 1). Так как

$$\lim_{n\to\infty}\frac{a_n}{b_n}=\lim_{n\to\infty}\frac{n^2(3n+1)}{3(n^3+5n-1)}=1,$$

то из сходимости ряда $b_1+b_2+\dots+b_n+\dots$ вытекает сходимость ряда $a_1+a_2+\dots+a_n+\dots$. ●

Для выяснения вопроса о сходимости того или иного ряда обычно пытаются сравнить его с одним из «стандартных» рядов. Таким «стандартным» рядом чаще всего служит геометрическая прогрессия. Сравнение с геометрической прогрессией лежит в основе следующего часто применяемого признака сходимости.

Теорема 4 (признак сходимости Даламбера*). *Если для ряда $a_1+a_2+\dots+a_n+\dots$ с положительными членами существует такое число $q<1$, что при всех n (или хотя бы начиная с некоторого n) выполняется неравенство*

$$\frac{a_{n+1}}{a_n}\leqslant q \tag{9}$$

(число q одно и то же для всех n), то ряд сходится. Если же $\frac{a_{n+1}}{a_n}\geqslant 1$ (для всех n или начиная с некоторого n), то ряд расходится.

□ Отбросив, если нужно, несколько первых членов ряда (что не влияет на сходимость), можно считать, что неравенство (9) выполняется для всех $n=1$, 2, ... Перепишем это неравенство в виде $a_{n+1}\leqslant qa_n$. Отсюда имеем: $a_2\leqslant qa_1$, $a_3\leqslant qa_2\leqslant q^2a_1$, $a_4\leqslant qa_3\leqslant q^3a_1$ и т. д.; вообще, при любом n справедливо неравенство

$$a_n\leqslant q^{n-1}a_1. \tag{10}$$

Это показывает, что члены ряда $a_1+a_2+\dots+a_n+\dots$ не превосходят соответствующих членов бесконечной геометрической прогрессии $a_1+qa_1+q^2a_1+\dots+q^{n-1}a_1+\dots$. Поскольку $0\leqslant q<1$, эта прогрессия является сходящейся. Отсюда в силу первого признака сравнения вытекает сходимость ряда $a_1+a_2+\dots+a_n+\dots$.

* Ж. Даламбер (1717—1783) — французский математик.

В том случае, если $\frac{a_{n+1}}{a_n} \geqslant 1$, имеем $a_{n+1} \geqslant a_n$, т. е. члены данного ряда возрастают. Тем самым нарушен необходимый признак сходимости ($\lim a_n \neq 0$) и, следовательно, ряд расходится. ■

Признак Даламбера часто применяется в следующей предельной форме.

Теорема 5. *Если существует предел*

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = l, \tag{11}$$

то ряд сходится в случае $l < 1$ *и расходится в случае* $l > 1$.

□ Пусть $l < 1$. Возьмем какое-либо число q, заключенное между l и 1: $l < q < 1$. Из условия (11) следует, что начиная с некоторого номера n выполняется неравенство $\frac{a_{n+1}}{a_n} \leqslant q$; на основании теоремы 4 отсюда следует сходимость ряда. Случай $l > 1$ предоставляем рассмотреть самостоятельно. ■

В формулировке теоремы 5 ничего не говорится о случае $l = 1$. При $l = 1$ возможна как сходимость, так и расходимость ряда. Например, ряд $1 + 1 + \cdots + 1 + \cdots$, для которого $l = 1$, расходится, в то время как ряд $1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \cdots$ (для него также $l = 1$) сходится.

Полезно заметить, что признак Даламбера дает и оценку остатка ряда: из неравенства (10) следует, что

$$R_k = a_{k+1} + a_{k+2} + a_{k+3} + \cdots \leqslant a_{k+1} + a_{k+1}q + a_{k+1}q^2 + \cdots,$$

т. е.

$$R_k \leqslant \frac{a_{k+1}}{1-q}. \tag{12}$$

Примеры. 5. Доказать сходимость ряда

$$1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots.$$

○ Находим $\frac{a_{n+1}}{a_n} = \frac{1}{(n+1)!} : \frac{1}{n!} = \frac{1}{n+1}$. Так как предел этого отношения равен 0 и, следовательно, меньше 1, то согласно признаку Даламбера ряд сходится. ●

6. Доказать сходимость ряда

$$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \cdots + \frac{n}{2^n} + \cdots$$

и оценить остаток R_9.

○ Имеем

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{2^{n+1}} : \frac{n}{2^n} = \frac{1}{2} \frac{n+1}{n} = \frac{1}{2}\left(1 + \frac{1}{n}\right).$$

Предел этого отношения равен 1/2, т. е. ряд сходится. Для оценки остатка $R_9 = a_{10} + a_{11} + \cdots$ заметим, что если $n \geqslant 10$, то

$$\frac{a_{n+1}}{a_n} = \frac{1}{2}\left(1 + \frac{1}{n}\right) \leqslant \frac{1}{2} \cdot 1{,}1 = 0{,}55,$$

поэтому

$$R_9 \leqslant \frac{a_{10}}{1-q} = \frac{\frac{10}{2^{10}}}{1 - 0{,}55} = \frac{10}{2^{10} \cdot 0{,}45} = \frac{1}{2^9 \cdot 0{,}09} < \frac{1}{46}. \text{ ●}$$

Ранее было отмечено, что n-й член ряда есть некоторая функция от n: $a_n = f(n)$. Следовательно, всякий ряд может быть представлен в виде

$$f(1)+f(2)+\dots+f(n)+\dots. \tag{13}$$

В тех случаях, когда члены ряда не только положительны, но и монотонно стремятся к нулю, может оказаться полезной следующая теорема.

Теорема 6 (интегральный признак сходимости). *Пусть функция $f(x)$ определена для $x>0$, положительна, монотонно убывает и стремится к нулю при $x\to\infty$. Тогда для сходимости ряда* (13) *необходимо и достаточно, чтобы сходился несобственный интеграл*

$$\int_1^\infty f(x)\,dx. \tag{14}$$

□ Сходимость интеграла (14) означает существование предела при $n\to\infty$ интеграла

$$\int_1^n f(x)\,dx, \tag{15}$$

а, значит, сходимость ряда

$$\int_1^2 f(x)\,dx+\int_2^3 f(x)\,dx+\dots+\int_n^{n+1} f(x)\,dx+\dots, \tag{16}$$

для которого интеграл (15) является частичной суммой (если сложить первые $n-1$ членов ряда, то получим интеграл (15)). Поэтому задача состоит в том, чтобы доказать одновременную сходимость или расходимость рядов (13) и (16).

В силу монотонности $f(x)$ на любом отрезке $[n, n+1]$ имеем $f(n+1)\leqslant f(x)\leqslant f(n)$ $(n=1, 2, \dots)$. Интегрируя по отрезку $[n, n+1]$, получим $f(n+1)\times \int_n^{n+1} dx\leqslant \int_n^{n+1} f(x)\,dx\leqslant f(n)\int_n^{n+1} dx$, т. е.

$$f(n+1)\leqslant \int_n^{n+1} f(x)\,dx\leqslant f(n). \tag{17}$$

Если сходится ряд (13), то из второго неравенства в (17) согласно признаку сравнения вытекает сходимость ряда (16). Обратно, если сходится ряд (16), то из первого неравенства в (17) согласно признаку сравнения вытекает сходимость ряда $f(2)+f(3)+\dots+f(n+1)+\dots$ а, значит, и ряда (13). ■

Пример 7. Выяснить, при каких $\alpha>0$ сходится ряд

$$1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\dots+\frac{1}{n^\alpha}+\dots. \tag{18}$$

○ Положим $f(x)=\frac{1}{x^\alpha}$ $(x>0)$. Функция $f(x)$ положительна, монотонно убывает и стремится к нулю при $x\to\infty$. Поэтому сходимость ряда эквивалентна сходимости интеграла $\int_1^\infty \frac{dx}{x^\alpha}$. Как известно (см. пример 4 § 42), этот интеграл сходится при $\alpha>1$ и расходится при $\alpha\leqslant 1$. Значит, и ряд (18) сходится при $\alpha>1$ и расходится при $\alpha\leqslant 1$. В частности, расходящимся является ряд $1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}+\dots$, названный ранее гармоническим (для него $\alpha=1$). Это уже известный нам факт (см. пример 3). ●

3. Оценка остатка ряда

При выводе интегрального признака сходимости мы использовали неравенства (17). Первое из них

$$f(n+1) \leqslant \int_{n}^{n+1} f(x)\,dx \quad (n=1,\ 2,\ \ldots) \tag{19}$$

позволяет оценить сверху k-й остаток ряда (13), т. е. число $R_k = f(k+1) + f(k+2) + \cdots$. Чтобы получить такую оценку, достаточно сложить почленно неравенства (19), записанные для $n=k,\ k+1,\ k+2,\ \ldots$. Тогда получим

$$R_k \leqslant \int_{k}^{\infty} f(x)\,dx.$$

Аналогично получается оценка остатка снизу. Для этого следует воспользоваться вторым неравенством в (17), а именно,

$$\int_{n}^{n+1} f(x)\,dx \leqslant f(n) \quad (n=1,\ 2,\ \ldots). \tag{20}$$

Если сложить неравенства (20), записанные для $n=k+1,\ k+2,\ k+3,\ \ldots$, то получим $\int\limits_{k+1}^{\infty} f(x)\,dx \leqslant R_k$.

В результате находим оценки остатка с обеих сторон (сверху и снизу):

$$\int_{k+1}^{\infty} f(x)\,dx \leqslant R_k \leqslant \int_{k}^{\infty} f(x)\,dx. \tag{21}$$

Пример. Сколько членов ряда $1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots$ следует сложить, чтобы вычислить его сумму с точностью до 0,001?

○ В данном случае $f(x)=\frac{1}{x^2}$; поэтому оценки (21) дают

$$\int_{k+1}^{\infty} \frac{dx}{x^2} \leqslant R_k \leqslant \int_{k}^{\infty} \frac{dx}{x^2}, \text{ или } \frac{1}{k+1} \leqslant R_k \leqslant \frac{1}{k}.$$

Второе из этих неравенств показывает, что при $k \geqslant 1000$ остаток R_k не превосходит 0,001. Таким образом, для вычисления суммы ряда с требуемой точностью следует сложить 1000 членов ряда! Как видим, сходимость ряда весьма «медленная». ●

§ 52. Свойства рядов с положительными членами

Свойства рядов с положительными членами во многом аналогичны свойствам конечных сумм. Например, сумма любого из таких рядов не меняется при перестановке его членов, умножение таких рядов сводится (в определенном смысле) к умножению их членов и т. д.

1. Перестановка членов ряда

Теорема 1. *Пусть ряд*

$$a_1 + a_2 + \cdots + a_n + \cdots, \tag{1}$$

члены которого положительны, сходится и имеет сумму A. Тогда ряд, полученный из него любой перестановкой членов, также сходится и имеет ту же сумму A.

□ Пусть

$$a_{i_1}+a_{i_2}+\cdots+a_{i_n}+\cdots \tag{2}$$

есть ряд (1) с переставленными членами. Так как любая частичная сумма A_n^* ряда (2) целиком содержится в некоторой частичной сумме A_N (где $N\geqslant n$) ряда (1), то $A_n^*\leqslant A_N$. Следовательно,

$$A_n^*\leqslant A, \tag{3}$$

где A сумма ряда (1). Это показывает, что частичные суммы ряда (2) ограничены сверху числом A; значит, ряд (2) сходится. Переходя в неравенстве (3) к пределу при $n\to\infty$, получим $A^*\leqslant A$. Но и ряд (1) получается из ряда (2) перестановкой членов; следовательно, $A\leqslant A^*$. Из двух последних неравенств вытекает, что $A=A^*$. ■

2. Умножение рядов

Как известно, для умножения двух конечных сумм $a_1+\cdots+a_n$ и $b_1+\cdots+b_m$ достаточно умножить каждое слагаемое первой суммы на каждое слагаемое второй суммы и затем сложить все полученные произведения. Оказывается, что такой способ умножения конечных сумм переносится и на умножение рядов с положительными членами.

Теорема 2. *Пусть ряды $a_1+a_2+\cdots+a_n+\cdots$ и $b_1+b_2+\cdots+b_n+\cdots$ с положительными членами сходятся соответственно к суммам A и B. Тогда ряд, составленный из всех произведений вида a_ib_j ($i=1, 2, \ldots, j=1, 2, \ldots$), расположенных в каком угодно порядке, также сходится, и его сумма равна AB.*

Другими словами, *сходящиеся ряды с положительными членами можно перемножать почленно.*

□ Пусть

$$a_{i_1}b_{j_1}+a_{i_2}b_{j_2}+\ldots+a_{i_n}b_{j_n}+\ldots \tag{4}$$

— какой-нибудь ряд, полученный почленным умножением данных рядов.

Докажем, что ряд (4) сходится, а его сумма не превосходит AB. Пусть C_n — частичная сумма ряда (4):

$$C_n=a_{i_1}b_{j_1}+a_{i_2}b_{j_2}+\cdots+a_{i_n}b_{j_n}.$$

Среди номеров $i_1, \ldots, i_n, j_1, \ldots, j_n$ существует наибольший; обозначим его N. Очевидно, $N=N(n)$. Имеем

$$C_n\leqslant(a_1+a_2+\cdots+a_N)(b_1+b_2+\cdots+b_N),$$

поскольку при перемножении двух сумм, заключенных в скобки, получатся все произведения $a_{i_1}b_{j_1}, \ldots, a_{i_n}b_{j_n}$ плюс, быть может, какие-то еще. Отсюда вытекает неравенство $C_n\leqslant AB$, означающее, что все частичные суммы ряда (4) ограничены сверху числом AB. Следовательно, ряд (4) сходится, а его сумма C не превосходит AB, т. е. $C\leqslant AB$.

Докажем, что $C=AB$. Пусть k — любое натуральное число. Все члены произведения A_kB_k, равного $\sum a_ib_j$ ($i=1, \ldots, k$; $j=1, \ldots, k$), содержатся в некоторой частичной сумме ряда (4) (ведь в этом ряде собраны все произведения a_ib_j, где i, j — два любых натуральных числа). Следовательно, существует такое $K=K(k)$, что $A_kB_k\leqslant C_K$. Поскольку $C_K\leqslant C$, имеем $A_kB_k\leqslant C$; переходя затем к пределу при $k\to\infty$, получаем $AB\leqslant C$. Итак, справедливы оба неравенства $C\leqslant AB$ и $AB\leqslant C$. Отсюда вытекает, что $C=AB$. ■

§ 53. Знакопеременные ряды

В § 51 и 52 рассматривались ряды с положительными членами. Теперь переходим к изучению рядов с членами произвольного знака, или, как принято говорить, *знакопеременных* рядов.

1. Ряды с чередующимися знаками. Теорема Лейбница

Среди знакопеременных рядов наиболее интересный класс образует ряды с чередующимися знаками, т. е. такие, в которых соседние члены имеют противоположные знаки. Для краткости будем называть такие ряды *знакочередующимися*.

Условимся считать первый член ряда положительным; тогда знакочередующийся ряд в общем виде запишется так:

$$a_1-a_2+a_3-a_4+\cdots+(-1)^{n-1}a_n+\cdots, \tag{1}$$

где все a_n положительны ($n=1, 2, \ldots$).

Теорема 1 (теорема Лейбница). *Если члены знакочередующегося ряда монотонно убывают по абсолютной величине и стремятся к нулю, то:* 1) *ряд сходится;* 2) *любой остаток ряда не превосходит по абсолютной величине первого из своих членов и имеет одинаковый с ним знак.*

Второе утверждение теоремы означает, что при любом k число

$$R_k=(-1)^k a_{k+1}+(-1)^{k+1}a_{k+2}+\cdots$$

(k-й остаток ряда) имеет тот же знак, что и $(-1)^k$, и удовлетворяет неравенству $|R_k|\leqslant a_{k+1}$.

□ Докажем первое утверждение. Рассмотрим частичную сумму ряда (1) с любым четным номером $2n$. Представим ее в виде

$$A_{2n}=(a_1-a_2)+(a_3-a_4)+\cdots+(a_{2n-1}-a_{2n}).$$

Все разности в скобках положительны, так как по условию $a_i>a_{i+1}$ при любом i. При переходе от A_{2n} к A_{2n+2} добавляется еще одна разность $a_{2n+1}-a_{2n+2}$; следовательно, $A_{2n}<A_{2n+2}$. Итак, частичные суммы с четными номерами образуют возрастающую последовательность. Покажем, что эта последовательность ограничена сверху. Для этого перепишем частичную сумму A_{2n} в виде

$$A_{2n}=a_1-(a_2-a_3)-(a_4-a_5)-\cdots-(a_{2n-2}-a_{2n-1})-a_{2n}.$$

Значит, $A_{2n}<a_1$ при любом n.

Как известно, монотонно возрастающая и ограниченная последовательность имеет предел. Поэтому существует число $A=\lim\limits_{n\to\infty}A_{2n}$, причем из сказанного выше следует, что $0<A\leqslant a_1$.

Так как $A_{2n+1}=A_{2n}+a_{2n+1}$ и по условию $\lim a_{2n+1}=0$, то $\lim A_{2n+1}=\lim A_{2n}=A$. Мы доказали, что ряд (1) сходится и его сумма A удовлетворяет условиям

$$0<A\leqslant a_1. \tag{2}$$

Докажем теперь второе утверждение. Рассмотрим остаток ряда (1) с четным номером $2k$: $R_{2k}=a_{2k+1}-a_{2k+2}+\cdots$. Этот ряд, подобно исходному ряду

(1), удовлетворяет условиям теоремы; отсюда по доказанному следует, что его сумма лежит в границах

$$0 < R_{2k} \leqslant a_{2k+1} \tag{3}$$

Что касается остатков ряда с нечетными номерами, то для любого из них можно записать $R_{2k+1} = -a_{2k+2} + a_{2k+3} - \cdots = -(a_{2k+2} - a_{2k+3} + \cdots)$. Ряд в скобках снова удовлетворяет условиям теоремы, поэтому имеем $0 < -R_{2k+1} \leqslant a_{2k+2}$ или

$$-a_{2k+2} \leqslant R_{2k+1} < 0. \tag{4}$$

Сходимость ряда (1) вместе с неравенствами (2); (3), (4) полностью доказывают теорему. ■

Пример. Доказать сходимость ряда

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots + \frac{(-1)^{n-1}}{n^2} + \cdots. \tag{5}$$

Сколько членов ряда нужно взять, чтобы вычислить сумму ряда с точностью до 0,001?

○ Сходимость ряда непосредственно следует из теоремы 1. Для вычисления с нужной точностью суммы ряда воспользуемся оценкой остатка, о которой говорится в теореме: $|R_k| \leqslant a_{k+1}$. В данном случае получаем неравенство $|R_k| \leqslant 1/(k+1)^2$. В частности, при $k = 31$ находим $|R_{31}| \leqslant 1/32^2 < 0{,}001$. Таким образом, достаточно сложить первые 31 членов ряда. При этом R_{31} имеет тот же знак, что и 32-й член ряда, т. е. $R_{31} < 0$. ●

Поучительно сопоставить последний пример с примером из п. 3 § 51. Для ряда

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \cdots,$$

состоящего из тех же членов, что и ряд (5), но взятых со знаком плюс, вычисление суммы с той же точностью требовало сложения 1000 членов! Эти примеры показывают, что знакочередующиеся ряды сходятся «быстрее», чем соответствующие ряды с положительными членами. Поэтому они удобнее для вычислений.

2. Ряды с членами произвольного знака. Плюс- и минус-ряды для данного ряда

Рассмотрим теперь ряды, члены которых имеют произвольные знаки. С самого начала условимся считать все члены ряда отличными от нуля (нулевые члены всегда можно отбросить — это не отразится ни на сходимости ряда, ни на его сумме). С каждым знакопеременным рядом

$$a_1 + a_2 + \cdots + a_n + \cdots \tag{6}$$

можно связать два ряда с положительными членами; один из них будем называть *плюс-рядом*, а другой *минус-рядом* для данного ряда (6). Чтобы получить плюс-ряд, следует оставить в ряде (6) только положительные члены (а отрицательные отбросить); для получения минус-ряда нужно оставить только отрицательные члены ряда (6), сменив в них знаки — на +. Например, для знакочередующегося ряда

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots$$

плюс- и минус-ряды соответственно имеют вид

$$1+\frac{1}{3}+\frac{1}{5}+\cdots \quad \text{и} \quad \frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\cdots .$$

Любая частичная сумма A_n ряда (6) включает некоторую часть плюс-ряда и некоторую часть минус-ряда; поэтому

$$A_n = A^+_{n_1} - A^-_{n_2}, \tag{7}$$

где $A^+_{n_1}$ — некоторая частичная сумма плюс-ряда и $A^-_{n_2}$ — частичная сумма минус-ряда.

3. Абсолютно сходящиеся ряды и их свойства

Пусть дан знакопеременный ряд (6). Рассмотрим ряд

$$|a_1|+|a_2|+\cdots+|a_n|+\cdots, \tag{8}$$

членами которого являются модули членов ряда (6). Очевидно, что (8) есть ряд с положительными членами.

Определение 1. Ряд (6) называется *абсолютно сходящимся*, если сходится ряд (8), составленный из модулей его членов.

Справедливо следующее утверждение.

Теорема 2. *Всякий абсолютно сходящийся ряд сходится. Сумма такого ряда равна разности между суммой его плюс-ряда и суммой минус-ряда.*

□ Пусть ряд (6) абсолютно сходится, т. е. сходится ряд (8). Обозначим частичные суммы ряда (8) через A^*_n. Аналогично равенству (7) имеем

$$A^*_n = A^+_{n_1} + A^-_{n_2}. \tag{9}$$

Ввиду сходимости ряда (8) его частичные суммы A^*_n ограничены сверху некоторым числом C. Тогда из равенства (9) следует $A^+_{n_1} \leqslant C$ и $A^-_{n_2} \leqslant C$, т. е. частичные суммы плюс- и минус-ряда также ограничены сверху числом C. Согласно критерию сходимости рядов с положительными членами отсюда вытекает сходимость плюс- и минус-рядов, т. е. существование чисел $A^+ = \lim\limits_{k\to\infty} A^+_k$ и $A^- = \lim\limits_{l\to\infty} A^-_l$. Если теперь в равенстве (7) перейти к пределу при $n\to\infty$, то получим $\lim\limits_{n\to\infty} A_n = A^+ - A^-$. ■

Принимая во внимание теорему 2, можно сказать, что абсолютно сходящийся ряд — это ряд, который не только сходится сам, но и обладает сходящимся рядом из модулей своих членов.

Пример. Доказать абсолютную сходимость ряда

$$\sum_{n=1}^{\infty} (-1)^n \frac{\sin\frac{\pi}{n}}{n}. \tag{10}$$

○ Составим ряд из модулей членов данного ряда:

$$\sum_{n=1}^{\infty} \frac{\sin\frac{\pi}{n}}{n}. \tag{11}$$

Так как $\sin\frac{\pi}{n} < \frac{\pi}{n}$, то члены ряда (11) не превосходят соответствующих членов сходящего ряда $\sum\limits_{n=1}^{\infty} \frac{\pi}{n^2}$. Следовательно, ряд (11) сходится, а ряд (10) — сходится абсолютно. ●

Поскольку абсолютная сходимость ряда (6) означает сходимость ряда (8), члены которого положительны, для установления абсолютной сходимости можно пользоваться любым из признаков сходимости, рассмотренных в § 51.

Обратим попутно внимание на следующее обстоятельство: в некоторых случаях из свойств ряда (8) можно делать заключение о расходимости ряда (6). Например, справедливо следующее утверждение.

Теорема 3 (признак расходимости знакопеременного ряда). *Если для всех n (или начиная с некоторого номера n) выполняется неравенство* $\frac{|a_{n+1}|}{|a_n|} \geqslant 1$, *то ряд* $a_1 + a_2 + \dots + a_n + \dots$ *расходится.*

□ Имеем $|a_{n+1}| \geqslant |a_n|$, что делает невозможным равенство $\lim a_n = 0$, являющееся необходимым условием сходимости. ■

Свойства абсолютно сходящихся рядов аналогичны доказанным в § 52 свойствам рядов с положительными членами. Имеют место следующие теоремы.

Теорема 4. *Если ряд сходится абсолютно и имеет сумму A, то ряд, полученный из него перестановкой (произвольной) членов, также сходится абсолютно и имеет ту же сумму A.*

Теорема 5. *Если* $a_1 + a_2 + \dots + a_n + \dots$ *и* $b_1 + b_2 + \dots + b_n + \dots$ — *два абсолютно сходящихся ряда, суммы которых равны соответственно A и B, то ряд, составленный из всевозможных произведений* $a_i b_j$ $(i = 1, 2, \dots;\ j = 1, 2, \dots)$, *также сходится абсолютно и имеет сумму AB.*

Теорема 4 следует из того, что при перестановке членов ряда: 1) переставляются члены ряда, составленного из модулей, причем, как известно, последний ряд сходится (см. теорему 1 § 52); 2) переставляются члены плюс- и минус-рядов, в результате чего сходимость этих рядов, а также их суммы A^+ и A^- сохраняются; 3) сохраняется и сумма исходного ряда, поскольку она равна $A^+ - A^-$.

Для доказательства теоремы 5 достаточно заметить, что: 1) из сходимости рядов $\sum\limits_{i=1}^{\infty} |a_i|$ и $\sum\limits_{j=1}^{\infty} |b_j|$ следует сходимость ряда $\sum\limits_{i,j=1}^{\infty} |a_i b_j|$ (см. теорему 2 § 52) и, значит, ряд $\sum\limits_{i,j=1}^{\infty} a_i b_j$ сходится абсолютно; 2) ввиду абсолютной сходимости последнего ряда его сумма не меняется при любой перестановке членов, в частности, если расположить его члены так, чтобы для частичной суммы с номером n^2 $(n = 1, 2, \dots)$ было выполнено равенство $S_{n^2} = \sum\limits_{i=1}^{n} a_i \sum\limits_{j=1}^{n} b_j$; 3) из последнего равенства в пределе при $n \to \infty$ вытекает, что сумма ряда $\sum\limits_{i,j=1}^{\infty} a_i b_j$ равна AB.

4. Условно сходящиеся ряды и их свойства

Определение 2. Ряд $a_1 + a_2 + \dots + a_n + \dots$ называется *условно сходящимся*, если он сходится, но в то же время ряд из модулей $|a_1| + |a_2| + \dots + |a_n| + \dots$ расходится.

Таким образом, все сходящиеся ряды можно разделить на два класса: абсолютно сходящиеся, т. е. такие, для которых ряд из модулей их членов сходится, и условно сходящиеся, для которых ряд из модулей членов расходится.

Примером условно сходящегося ряда может служить ряд

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+\frac{(-1)^{n-1}}{n}+\cdots. \tag{12}$$

Действительно, по теореме Лейбница этот ряд сходится; в то же время ряд из модулей его членов

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{n}+\cdots$$

(гармонический ряд), как известно, расходится.

Свойства условно сходящихся рядов резко отличаются от свойств абсолютно сходящихся рядов. В частности, для таких рядов неверна теорема о возможности перестановки членов ряда.

Пример. В условно сходящемся ряде (12) переставим и сгруппируем члены следующим образом:

$$\left(1-\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{6}-\frac{1}{8}\right)+\cdots+\left(\frac{1}{2n-1}-\frac{1}{4n-2}-\frac{1}{4n}\right)+\cdots.$$

Если затем внутри каждых скобок сложить первые два (из трех) слагаемых, то получим ряд

$$\left(\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{6}-\frac{1}{8}\right)+\cdots+\left(\frac{1}{4n-2}-\frac{1}{4n}\right)+\cdots,$$

сумма которого равна

$$\frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\cdots=\frac{1}{2}\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots\right).$$

Таким образом, при перестановке членов исходного ряда (12) его сумма уменьшилась вдвое!

Справедлива следующая теорема, которую мы приведем без доказательства.

Теорема 6 (теорема Римана*). *Если ряд сходится условно, то в результате перестановки его членов можно получить ряд, имеющий любую наперед заданную сумму, а также расходящийся ряд.*

§ 54. Функциональные ряды

1. Функциональный ряд и его область сходимости

До сих пор члены ряда мы считали постоянными числами. Перейдем теперь к рассмотрению рядов, членами которых являются функции некоторой переменной x. Такие ряды называют *функциональными рядами.*

Будем записывать функциональный ряд в виде

$$f_1(x)+f_2(x)+\cdots+f_n(x)+\cdots. \tag{1}$$

* Б. Риман (1826—1866) — немецкий математик.

Если взять какое-либо конкретное значение x_0 аргумента x и подставить в ряд (1), то получим числовой ряд

$$f_1(x_0)+f_2(x_0)+\cdots+f_n(x_0)+\cdots.$$

Так как x может принимать, вообще говоря, бесчисленное множество значений, то с данным функциональным рядом оказывается связанным бесчисленное множество числовых рядов. Некоторые из них являются сходящимися, некоторые — расходящимися.

Определение 1. **Множество тех значений x, для которых функциональный ряд (1) сходится, называется *областью сходимости* этого ряда.**

Примеры. 1. Найти область сходимости функционального ряда

$$\frac{x}{1}+\frac{x^2}{2}+\frac{x^3}{3}+\cdots+\frac{x^n}{n}+\cdots. \tag{2}$$

○ Исследуем данный ряд сразу на абсолютную сходимость. Ряд, составленный из модулей его членов, имеет вид

$$|x|+\frac{|x|^2}{2}+\frac{|x|^3}{3}+\cdots+\frac{|x|^n}{n}+\cdots. \tag{3}$$

Найдем отношение последующего члена к предыдущему (т. е. $(n+1)$-го члена к n-му):

$$\frac{|x|^{n+1}}{n+1}:\frac{|x|^n}{n}=|x|\frac{n}{n+1}.$$

Предел этого отношения при $n\to\infty$ равен $|x|$. Согласно признаку Даламбера ряд (3) сходится при $|x|<1$ и расходится при $|x|>1$; в последнем случае члены ряда начиная с некоторого номера возрастают. Отсюда следует, что исходный ряд (2) сходится при $|x|<1$, а при $|x|>1$ расходится (см. теорему 3 § 53). Остается выяснить поведение ряда при $x=1$ и $x=-1$. При $x=1$ получаем гармонический ряд $1+\frac{1}{2}+\frac{1}{3}+\cdots$, который расходится; при $x=-1$ имеем ряд $-1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\cdots$, который сходится по теореме Лейбница. Итак, область сходимости ряда (2) есть промежуток $(-1,1)$. В дальнейшем мы установим, что суммой ряда (2) является функция $-\ln(1-x)$. ●

2. Найти область сходимости функционального ряда

$$\frac{x}{1+x^2}+\frac{x^2}{1+x^4}+\cdots+\frac{x^n}{1+x^{2n}}+\cdots \tag{4}$$

○ Исследуем на сходимость ряд из модулей

$$\frac{|x|}{1+x^2}+\frac{|x|^2}{1+x^4}+\cdots+\frac{|x|^n}{1+x^{2n}}+\cdots. \tag{5}$$

При любом x имеем $\frac{|x|^n}{1+x^{2n}}\leqslant|x|^n$. Если $|x|<1$, то ряд $|x|+|x|^2+\cdots+|x|^n+\cdots$ сходится (убывающая геометрическая прогрессия); следовательно, сходится и ряд (5). Если же $|x|>1$, то записав неравенство

$$\frac{|x|^n}{1+x^{2n}}\leqslant\frac{|x|^n}{x^{2n}}=\frac{1}{|x|^n},$$

снова установим сходимость ряда (5) с помощью сравнения его с геометрической прогрессией $\frac{1}{|x|}+\frac{1}{|x|^2}+\cdots+\frac{1}{|x|^n}+\cdots$. Наконец, при $|x|=1$ получаем из (5) ряд, все члены которого равны $\frac{1}{2}$; этот ряд (а вслед за ним ряд (4)) расходится. Итак, область сходимости ряда (4) есть вся числовая прямая, за исключением точек $x=1$ и $x=-1$. ●

Для функционального ряда (1) понятия частичной суммы и остатка ряда определяются точно так же, как и для числового ряда: *частичная сумма* $F_n(x)$ есть функция

$$F_n(x) = f_1(x) + f_2(x) + \dots + f_n(x) \quad (n = 1, 2, \dots),$$

а *n-й остаток* ряда есть ряд

$$f_{n+1}(x) + f_{n+2}(x) + \cdots . \tag{6}$$

Впрочем, остатком называют не только самый ряд (6), но и его сумму $R_n(x)$. Для всех значений x из области сходимости имеем

$$F(x) = F_n(x) + R_n(x), \tag{7}$$

где $F(x)$ — сумма ряда (1), т. е. $F(x) = \lim\limits_{n\to\infty} F_n(x)$. Из соотношения (7) в пределе при $n \to \infty$ находим $\lim\limits_{n\to\infty} R_n(x) = 0$. Еще раз подчеркнем, что это равенство имеет место для всех значений x из области сходимости ряда (1).

2. Правильно сходящиеся функциональные ряды

Оперируя с функциональными рядами (как, впрочем, и с числовыми), нужно иметь в виду, что свойства бесконечных сумм функций могут отличаться от свойств конечных сумм. В частности, над конечными суммами функций можно выполнять почленно целый ряд операций: переход к пределу (предел суммы равен сумме пределов), интегрирование, дифференцирование. В случае же бесконечных сумм функций почленное выполнение указанных операций может привести к неверным результатам. Поэтому одним из первых вопросов теории функциональных рядов должен быть следующий: когда можно ту или иную операцию (переход к пределу, интегрирование, дифференцирование) выполнять над рядом почленно? Чтобы дать ответ на этот вопрос, введем понятие правильно сходящегося ряда.

Определение 2. Пусть дан функциональный ряд (1), причем все функции $f_n(x)$ $(n = 1, 2, \dots)$ определены на одном и том же множестве M. Говорят, что ряд (1) *правильно сходится* на множестве M, если существует такой сходящийся числовой ряд с положительными членами

$$a_1 + a_2 + \cdots + a_n + \cdots, \tag{8}$$

что для всех $x \in M$ выполняются неравенства

$$|f_n(x)| \leqslant a_n \quad (n = 1, 2, \dots). \tag{9}$$

При этом ряд (8) называют *мажорантой* (от лат. major — «больший») функционального ряда (1); говорят также, что ряд (1) *мажорируется* числовым рядом (8).

Если ряд (1) правильно сходится на множестве M, то он обязательно сходится, и притом абсолютно, при каждом значении x из M: это следует из неравенств (9), если воспользоваться признаком сравнения рядов с положительными членами.

Пример. Функциональный ряд

$$\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \cdots + \frac{\sin nx}{n^2} + \cdots$$

сходится правильно на всей числовой прямой, так как при любом x справедливы неравенства $\left|\frac{\sin nx}{n^2}\right| \leqslant \frac{1}{n^2}$ ($n=1, 2, \ldots$). Здесь мажорантой является сходящийся числовой ряд

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots.$$

3. Свойства правильно сходящихся рядов

Для доказательства основных теорем о правильно сходящихся рядах нам потребуется одно важное свойство таких рядов. Сформулируем его в виде леммы.

Лемма. *Пусть функциональный ряд* (1) *правильно сходится на некотором множестве* M. *Тогда для всех* $x \in M$ *и всех номеров* n *выполняются неравенства*

$$|R_n(x)| \leqslant R_n, \tag{10}$$

где R_n — *остаток мажорирующего числового ряда* (8)

□ Пусть $x \in M$. В силу неравенств (9) имеем

$$\begin{gathered}|f_{n+1}(x)+f_{n+2}(x)+\ldots+f_{n+m}(x)| \leqslant |f_{n+1}(x)|+|f_{n+2}(x)|+\ldots+|f_{n+m}(x)| \leqslant \\ \leqslant a_{n+1}+a_{n+2}+\ldots+a_{n+m},\end{gathered}$$

каковы бы ни были n и m. В пределе при $m \to \infty$ получаем неравенство (10). ■

Значение правильно сходящихся рядов заключается прежде всего в том, что с ними можно оперировать как с конечными суммами. Справедливы следующие утверждения.

Теорема 1. *Пусть функциональный ряд* (1) *правильно сходится на отрезке* $[a, b]$. *Если все члены ряда непрерывны на* $[a, b]$, *то и сумма* $F(x)$ *этого ряда также непрерывна на* $[a, b]$.

□ Пусть точки x_0 и $x_0+\Delta x$ принадлежат $[a, b]$. Имеем $F(x_0)=F_n(x_0)+R_n(x_0)$, $F(x_0+\Delta x)=F_n(x_0+\Delta x)+R_n(x_0+\Delta x)$. Вычитая из второго равенства первое, получаем $\Delta F=\Delta F_n+R_n(x_0+\Delta x)-R_n(x_0)$, откуда

$$|\Delta F| \leqslant |\Delta F_n|+|R_n(x_0+\Delta x)|+|R_n(x_0)|.$$

Из этого неравенства на основании леммы следует

$$|\Delta F| \leqslant |\Delta F_n|+R_n+R_n, \tag{11}$$

где R_n — остаток межорирующего ряда (8).

Зададим произвольное $\varepsilon>0$. Так как ряд (8) сходится, то $R_n \to 0$ при $n \to \infty$. Это означает, что для данного ε найдется такой номер $N=N(\varepsilon)$, что для всех $n>N$ справедливо неравенство $R_n<\varepsilon$. Зафиксируем один из таких номеров n и обозначим его n_0. Поскольку функция $F_{n_0}(x)$ непрерывна на отрезке $[a, b]$ и, в частности, в точке x_0, найдется положительное число δ такое, что при $|\Delta x|<\delta$ будет выполнено неравенство $|\Delta F_{n_0}|<\varepsilon$. Тогда из соотношения (11) при $|\Delta x|<\delta$ следует $|\Delta F|<\varepsilon+\varepsilon+\varepsilon=3\varepsilon$, что и доказывает непрерывность функции $F(x)$ в точке x_0. ■

Теорема 2. *Пусть функциональный ряд* (11), *члены которого являются непрерывными функциями на отрезке* $[a, b]$, *сходится правильно на* $[a, b]$ *и пусть* $F(x)$ — *сумма этого ряда. Тогда*

$$\int_a^b F(x)\,dx=\int_a^b f_1(x)\,dx+\int_a^b f_2(x)\,dx+\cdots+\int_a^b f_n(x)\,dx+\cdots \tag{12}$$

т. е. интеграл от суммы ряда равен сумме интегралов от членов ряда.

□ Снова воспользуемся равенством $F(x)=F_n(x)+R_n(x)$ или, подробнее, $F(x)=f_1(x)+\cdots+f_n(x)+R_n(x)$. Интегрируя обе части равенства, находим

$$\int_a^b F(x)dx=\int_a^b f_1(x)dx+\ldots+\int_a^b f_n(x)dx+\int_a^b R_n(x)dx.$$

Сумма первых n слагаемых правой части есть не что иное, как n-я частичная сумма ряда (12). Поэтому, чтобы доказать справедливость равенства (12), остается проверить, что

$$\int_a^b R_n(x)dx\to 0 \text{ при } n\to\infty. \tag{13}$$

Имеем $\left|\int_a^b R_n(x)dx\right|\leqslant\int_a^b |R_n(x)|dx$, откуда на основании леммы следует $\left|\int_a^b R_n(x)dx\right|\leqslant R_n(b-a)$. Так как $R_n\to 0$ при $n\to\infty$, то справедливо соотношение (13). ■

Теорема 3. *Пусть функциональный ряд* (12) *сходится на отрезке* $[a, b]$ *к некоторой функции* $F(x)$. *Если члены ряда имеют непрерывные производные на отрезке* $[a, b]$ *и если ряд*

$$f_1'(x)+f_2'(x)+\ldots+f_n'(x)+\ldots, \tag{14}$$

составленный из производных, сходится правильно, то для любого x *из отрезка* $[a, b]$ *справедливо равенство*

$$F'(x)=f_1'(x)+f_2'(x)+\ldots+f_n'(x)+\ldots, \tag{15}$$

т. е. производная суммы ряда существует и равна сумме производных от членов ряда.

Отметим, что правильная сходимость требуется не от самого ряда (1), а от ряда, составленного из производных.

□ Сумму ряда (14) обозначим $\Phi(x)$. Мы должны доказать, что $F'(x)=\Phi(x)$. Пусть $x\in[a, b]$. На основании теоремы 2 можем записать равенство

$$\int_a^x \Phi(t)dt=\int_a^x f_1'(t)dt+\int_a^x f_2'(t)dt+\ldots+\int_a^x f_n'(t)dt+\ldots$$

или

$$\int_a^x \Phi(t)dt=[f_1(x)-f_1(a)]+[f_2(x)-f_2(a)]+\ldots+[f_n(x)-f_n(a)]+\ldots.$$

Стоящий справа ряд есть результат почленного сложения двух рядов: $f_1(x)+$ $+f_2(x)+\ldots+f_n(x)+\ldots$ и $f_1(a)-f_2(a)-\ldots-f_n(a)-\ldots$; значит, его сумма равна $F(x)-F(a)$. Итак,

$$\int_a^x \Phi(x)dx=F(x)-F(a).$$

Дифференцируя по x обе части равенства, получаем $F'(x)=\Phi(x)$. ■

Еще раз подчеркнем то обстоятельство, что правильно сходящимся должен быть ряд (14), а не исходный ряд (1). Одно условие нельзя заменить другим, что показывает следующий пример.

Пример. Ряд

$$\sin x+\frac{\sin 4x}{4}+\frac{\sin 9x}{9}+\cdots+\frac{\sin n^2x}{n^2}+\cdots$$

правильно сходится на всей числовой прямой и, значит, его сумма $F(x)$ существует для всех x. Однако заключение теоремы 3 здесь неверно, поскольку ряд из производных

$$\cos x+\cos 4x+\cos 9x+\cdots+\cos n^2x+\cdots$$

сходится не при всех x; например, при $x=0$ он расходится (можно показать, что этот ряд расходится при «почти всех» значениях x).

§ 55. Степенные ряды

1. Степенной ряд. Теорема Абеля

Степенные ряды — важнейший вид функциональных рядов.

Определение 1. *Степенным рядом* называется функциональный ряд вида

$$a_0+a_1(x-x_0)+a_2(x-x_0)^2+\cdots+a_n(x-x_0)^n+\cdots,$$

где a_0, a_1, a_2, ..., а также x_0 — постоянные числа. Точку x_0 часто называют *центром* степенного ряда.

Так, например,

$$1+x+x^2+\cdots+x^n+\cdots$$

есть степенной ряд с центром $x_0=0$, а

$$5+(x-1)+\frac{(x-1)^2}{2}+\frac{(x-1)^3}{3}+\cdots+\frac{(x-1)^n}{n}+\cdots$$

есть степенной ряд с центром $x_0=1$.

Замечание. При рассмотрении степенного ряда условимся считать, что член ряда, записанный на первом месте, имеет нулевой номер, затем следуют первый, второй и т. д. члены. Такое соглашение удобно тем, что n-й член ряда содержит x в степени n.

Сначала рассмотрим степенные ряды с центром 0, т. е. ряды вида

$$a_0+a_1x+a_2x^2+\cdots+a_nx^n+\cdots. \tag{1}$$

Такой ряд всегда сходится при $x=0$ и, значит, его область сходимости есть непустое множество.

Важную роль при изучении степенных рядов играет следующее утверждение.

Теорема 1 (теорема Абеля*). *Если ряд* (1) *сходится при $x=\alpha$, где α — число, не равное нулю, то он сходится, и притом абсолютно, при любом значении x таком, что $|x|<|\alpha|$.*

□ По условию числовой ряд $a_0+a_1\alpha+a_2\alpha^2+\cdots+a_n\alpha^n+\cdots$ является сходящимся; следовательно, его общий член $a_n\alpha^n$ стремится к нулю. Но любая последовательность, имеющая предел, ограничена; значит, существует такое число M, что $|a_n\alpha^n|<M$ для всех $n=0, 1, 2,...$

Рассмотрим теперь ряд

$$|a_0|+|a_1x|+|a_2x^2|+\cdots+|a_nx^n|+\cdots; \tag{2}$$

* Н. Х. Абель (1802—1829) — норвежский математик.

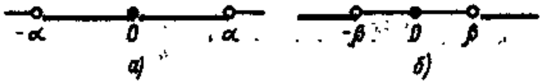


Рис. 197

предполагая, что $|x|<|\alpha|$. Так как

$$|a_n x^n| = |a_n \alpha^n| \cdot \left|\frac{x}{\alpha}\right|^n < M\left|\frac{x}{\alpha}\right|^n \tag{3}$$

и при этом $\left|\frac{x}{\alpha}\right|<1$, то члены ряда (2) не превосходят соответствующих членов сходящегося ряда

$$M + M\left|\frac{x}{\alpha}\right| + M\left|\frac{x}{\alpha}\right|^2 + \dots + M\left|\frac{x}{\alpha}\right|^n + \dots$$

(геометрической прогрессии). Следовательно, ряд (2) сходится, а ряд (1) — абсолютно сходится. ■

Следствие. *Если ряд* (1) *расходится в точке* β, *то он расходится и во всякой точке* x *такой, что* $|x|>|\beta|$.

Действительно, если бы в точке x ряд сходился, то по теореме Абеля он сходился бы и в точке β, что противоречит условию.

Геометрическое содержание теоремы 1 (и ее следствия) весьма просто: *если степенной ряд* (1) *сходится в точке* α, *то он сходится и во всех точках, расположенных ближе к началу, чем* α (рис. 197, *а*), *если же ряд расходится в точке* β, *то он расходится и во всех более удаленных от начала точках* (рис. 197, *б*).

2. Область сходимости степенного ряда

Теорема Абеля позволяет в принципе дать описание области сходимости степенного ряда.

Теорема 2. *Для степенного ряда* (1) *возможны только три случая:*

1) *ряд сходится в единственной точке* $x=0$;
2) *ряд сходится при всех значениях* x;
3) *существует такое* $R>0$, *что ряд сходится при всех значениях* x, *для которых* $|x|<R$, *и расходится при всех* x, *для которых* $|x|>R$.

В случае 3 сходимость ряда при $x=R$ и $x=-R$ зависит конкретно от вида этого ряда.

□ Обозначим через A множество всех положительных значений x, для которых ряд (1) сходится.

Если A пусто, то, как следует из теоремы Абеля, ряд (1) сходится только при $x=0$. Это дает случай 1.

Пусть A непусто. Если A не ограничено сверху, т. е. содержит как угодно большие числа, то из теоремы Абеля следует, что ряд (1) сходится при всех x. Это дает случай 2.

Остается рассмотреть случай, когда A непусто и ограничено сверху. Согласно теореме 1 § 10 существует точная верхняя граница для множества A; обозначим ее R. Число R обладает тем свойством, что в любой его окрестности имеются точки, как принадлежащие, так и не принадлежащие A; другими словами, в любой окрестности точки R имеются точки, в которых ряд (1) сходится, и точки, в которых он расходится. Используя теорему Абеля и следствие из нее, заключаем, что при всех x таких, что $|x|<R$, ряд (1) сходится, а при всех x, для которых $|x|>R$, — расходится. Это дает случай 3. ■

Итак, *областью сходимости степенного ряда* (1) *является или точка* $x=0$, *или вся числовая прямая, или конечный интервал* $(-R, R)$, *к которому могут присоединяться один или оба конца.*

Определение 2. Интервал $(-R,R)$, где число R определено в теореме 2, называется ***интервалом сходимости*** ряда (1), а число R — ***радиусом сходимости*** этого ряда.

Понятие радиуса сходимости будем распространять на все три случая в теореме 2: для этого в случае 1 условимся считать $R=0$, а в случае 2 $R=\infty$.

3. Отыскание радиуса сходимости степенного ряда

На практике радиус сходимости степенного ряда чаще всего определяют с помощью признака сходимости Даламбера. Предположим сначала, что все коэффициенты ряда

$$a_0+a_1x+a_2x^2+\cdots+a_nx^n+\cdots$$

отличны от нуля и что существует предел $p=\lim\limits_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}$ отношения модулей коэффициентов последующего и предыдущего членов ряда. Тогда радиус сходимости ряда находится по формуле

$$R=1/p.$$

Действительно, в силу признака Даламбера ряд

$$|a_0|+|a_1x|+|a_2x^2|+\cdots+|a_nx^n|+\cdots$$

сходится, если число

$$l(x)=\lim_{n\to\infty}\frac{|a_{n+1}x^{n+1}|}{|a_nx^n|}=p|x|$$

меньше 1, и расходится, если $l(x)>1$. Иначе говоря, ряд сходится для всех x таких, что $|x|<1/p$, и расходится для всех x таких, что $|x|>1/p$. Это и означает, что число $1/p$ является радиусом сходимости ряда.

Встречаются степенные ряды, в которых часть коэффициентов равна нулю. Как правило, и в этих случаях можно воспользоваться признаком Даламбера, если, конечно, удалить из ряда нулевые члены. С примерами подобного рода мы познакомимся ниже (см. пример 2).

Примеры. 1. Найти область сходимости ряда

$$1+\frac{x}{1^2}+\frac{x^2}{2^2}+\cdots+\frac{x^n}{n^2}+\cdots.$$

○ Здесь все коэффициенты отличны от нуля и

$$p=\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}=\lim_{n\to\infty}\left[\frac{1}{(n+1)^2}:\frac{1}{n^2}\right]=\lim\frac{n^2}{(n+1)^2}=1.$$

Следовательно, радиус сходимости есть $1/p=1$. Ряд сходится при $-1<x<1$ и расходится при $|x|>1$. В точках $x=1$ и $x=-1$ ряд также сходится (объясните, почему). Итак, область сходимости ряда — отрезок $[-1, 1]$. ●

2. Найти область сходимости ряда

$$x+\frac{x^3}{3!}+\frac{x^5}{5!}+\dots+\frac{x^{2n+1}}{(2n+1)!}+\dots.$$

○ Полагая $f_n(x)=\frac{x^{2n+1}}{(2n+1)!}$, находим

$$\frac{|f_{n+1}(x)|}{|f_n(x)|}=\frac{|x|^{2n+3}}{(2n+3)!}:\frac{|x|^{2n+1}}{(2n+1)!}=|x|^2\frac{1}{(2n+2)(2n+3)}.$$

Предел (при $n\to\infty$) этого отношения при любом x равен 0. Ряд сходится при всех x; его радиус сходимости $R=\infty$. ●

3. Найти область сходимости ряда

$$x+(2x)^2+(3x)^3+\dots+(nx)^n+\dots.$$

○ Здесь $f_n(x)=n^nx^n$. Каково бы ни было значение x, отличное от нуля, при $n>1/|x|$ имеем $n^n|x|^n>1$, т. е. ряд при данном x расходится. Область сходимости состоит из единственной точки $x=0$ (радиус сходимости $R=0$). ●

4. Непрерывность суммы степенного ряда. Почленное интегрирование и дифференцирование ряда

В предыдущем параграфе мы отмечали, какое значение при установлении тех или иных свойств функциональных рядов имеет правильная сходимость ряда. Здесь мы сформулируем и докажем теоремы, выражающие важнейшие свойства правильно сходящихся степенных рядов.

Теорема 3. *Степенной ряд сходится правильно на любом отрезке, целиком содержащемся внутри интервала сходимости.*

□ Пусть степенной ряд (1) имеет интервал сходимости $(-R, R)$. Рассмотрим какой-нибудь отрезок $[\alpha, \beta]$, целиком содержащийся в $(-R, R)$. Очевидно, что всегда можно найти отрезок вида $[-r, r]$, содержащий $[\alpha, \beta]$ и

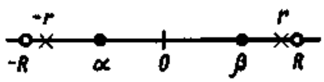


Рис. 198

целиком лежащий в $(-R, R)$ (рис. 198). Если $x\in[\alpha, \beta]$, то $|x|\leqslant r$ и, следовательно, члены ряда (1) не превосходят по модулю членов ряда

$$|a_0|+|a_1r|+|a_2r^2|+\dots+|a_nr^n|+\dots.$$

Но последний ряд сходится, так как $r<R$. Таким образом, для всех $x\in[\alpha, \beta]$ члены ряда (1) не превосходят по модулю членов сходящегося числового ряда. Это доказывает правильную сходимость ряда (1) на отрезке $[\alpha, \beta]$. ■

Теорема 4. *На любом отрезке $[\alpha, \beta]$, целиком содержащемся внутри интервала сходимости степенного ряда, сумма ряда есть непрерывная функция.*

□ Так как члены ряда являются непрерывными функциями и ряд сходится правильно на отрезке $[\alpha, \beta]$, то согласно теореме 1 § 54 сумма ряда есть непрерывная функция на $[\alpha, \beta]$. ■

Теорема 5. *Если отрезок $[\alpha, \beta]$ целиком содержится внутри интервала сходимости степенного ряда, то интеграл по $[\alpha, \beta]$ от суммы ряда равен сумме интегралов от членов ряда.*

Доказательство вытекает из теоремы 2 § 54.

Перейдем теперь к вопросу о возможности почленного дифференцирования степенного ряда. Основополагающее значение для всей теории степенных рядов имеет следующее утверждение.

Теорема 6. *Пусть степенной ряд*

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots, \tag{4}$$

радиус сходимости которого равен R, имеет сумму $F(x)$. Рассмотрим ряд

$$a_1 + 2a_2x + \dots + na_nx^{n-1} + \dots, \tag{5}$$

полученный почленным дифференцированием ряда (4); *пусть $\Phi(x)$ есть сумма этого ряда. Тогда:*

1) *ряд* (5) *имеет тот же радиус сходимости R, что и ряд* (4);

2) *во всем интервале $(-R, R)$ существует производная $F'(x)$, причем $F'(x) = \Phi(x)$, т. е. производная от суммы ряда* (4) *равна сумме производных от членов ряда.*

□ Докажем сначала, что ряд (5) правильно сходится на любом отрезке $[-r, r]$, где $r < R$. Для этого рассмотрим любую точку α между r и R. В этой точке ряд (4) сходится; следовательно, существует такое число M, что $|a_n\alpha^n| < M$ $(n = 0, 1, 2, \dots)$. Но при $|x| \leqslant r$ имеем

$$|na_nx^{n-1}| \leqslant |na_nr^{n-1}| = n|a_n\alpha^n| \cdot \left|\frac{r}{\alpha}\right|^{n-1} \frac{1}{\alpha} < n\frac{M}{\alpha}q^{n-1},$$

где $q = r/\alpha < 1$. Таким образом, для указанных x члены ряда (5) не превосходят членов числового ряда

$$\frac{M}{\alpha}(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots),$$

который является сходящимся (по признаку Даламбера). Следовательно, ряд (5) на $[-r, r]$ сходится правильно.

Из правильной сходимости ряда (5) на основании теоремы 3 § 54 приходим к заключению, что на всем отрезке $[-r, r]$ функция $F(x)$ имеет производную, причем $F'(x) = \Phi(x)$.

Так как любую точку из $(-R, R)$ можно охватить отрезком вида $[-r, r]$, содержащимся в $(-R, R)$, то во всем интервале $(-R, R)$ ряд (5) сходится и справедливо равенство $F'(x) = \Phi(x)$.

Пусть R_1 — радиус сходимости ряда (5). По доказанному, $R_1 \geqslant R$. Покажем, что не может быть $R_1 > R$.

Рассуждая от противного, допустим, что $R_1 > R$. Выберем число β, заключенное между R_1 и R: $R < \beta < R_1$. Так как на отрезке $[0, \beta]$ ряд (5) сходится правильно (по теореме 3), то его можно почленно интегрировать по этому отрезку. Это означает, в частности, что сходится ряд из интегралов

$$\int_0^\beta a_1dx + \int_0^\beta 2a_2x\,dx + \dots + \int_0^\beta na_nx^{n-1}dx + \dots,$$

т. е. ряд $a_1\beta + a_2\beta^2 + \dots + a_n\beta^n + \dots$. Однако это невозможно, поскольку ряд (4) в точке β расходится. Итак, $R_1 = R$, т. е. радиус сходимости ряда (5) есть R. ■

Применяя теорему 6 повторно, заключаем, что вторая производная $F''(x)$ также существует и равна сумме ряда, полученного двукратным почленным дифференцированием ряда (4). Аналогичное заключение справедливо для $F'''(x)$ и т. д. В итоге приходим к следующей теореме.

Теорема 7. *Степенной ряд в пределах его интервала сходимости можно дифференцировать почленно любое число раз. При этом радиусы сходимости всех рядов, полученных почленным дифференцированием данного ряда, совпадают с радиусом сходимости этого ряда.*

§ 56. Разложение функций в степенные ряды

1. Ряд Тейлора для заданной функции

Одна из центральных задач, решаемых в теории степенных рядов, заключается в следующем.

Пусть дана некоторая функция $f(x)$. Требуется установить:

1) может ли эта функция на данном отрезке быть представлена в виде суммы некоторого степенного ряда, или, как говорят, может ли функция $f(x)$ быть «разложенной в степенной ряд»?

2) если да, то как найти этот ряд?

В этом пункте мы дадим ответ на второй из поставленных вопросов; ответ на первый вопрос будет дан в следующем параграфе.

Итак, пусть известно, что функция $f(x)$ на некотором отрезке $[-r, r]$ разлагается в степенной ряд:

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n+\cdots. \qquad (1)$$

Найдем коэффициенты a_0, a_1, a_2, ... этого ряда.

Решение основано на теореме 7 из предыдущего параграфа. Согласно этой теореме, при любом x из интервала $(-r, r)$ имеем

$$f'(x)=a_1+2a_2x+3a_3x^2+4a_4x^3+\cdots,$$
$$f''(x)=2\cdot1a_2+3\cdot2a_3x+4\cdot3a_4x^2+\cdots,$$
$$f'''(x)=3\cdot2\cdot1a_3+4\cdot3\cdot2a_4x+\cdots$$

и т. д.; вообще,

$$f^{(n)}(x)=n(n-1)\cdots3\cdot2\cdot1a_n+\cdots$$

(в разложении для $f^{(n)}(x)$ нам важен только свободный член). Полагая в этих равенствах, а также в исходном разложении (1) $x=0$, находим $f(0)=a_0$, $f'(0)=$ $=a_1$, $f''(0)=2a_2$ и т. д.; вообще, $f^{(n)}(0)=n!a_n$. Отсюда получим формулу

$$a_n=\frac{f^{(n)}(0)}{n!} \quad (n=0,\ 1,\ 2,\ \ldots).$$

Определение 1. Пусть функция $f(x)$ определена в некоторой окрестности точки $x=0$ и имеет в этой точке производные всех порядков. Степенной ряд

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

называется *рядом Тейлора** для функции $f(x)$.

* Иногда указанный ряд называют *рядом Маклорена;* К. Маклорен (1698—1746) — английский ученый.

Ответ на поставленный выше вопрос можно теперь оформить в виде следующей теоремы.

Теорема 1. *Если функция $f(x)$ разлагается в некоторой окрестности точки $x=0$ в ряд по степеням x, то он является ее рядом Тейлора.*

Теорема 1 имеет принципиальное значение. Если нас интересует возможность разложения функции $f(x)$ в окрестности точки $x=0$ в ряд по степеням x, то мы можем с самого начала ограничить свои поиски одним единственным рядом — рядом Тейлора для $f(x)$. При этом вопрос сводится к следующему: имеет ли ряд Тейлора, составленный для функции $f(x)$, своей суммой именно эту функцию, или же сходится к какой-то другой функции? Ответ (для наиболее важных случаев) мы дадим в следующем пункте.

2. Достаточное условие разложимости функции в ряд Тейлора

Для решения вопроса о разложимости функции $f(x)$ в ряд Тейлора в окрестности точки $x=0$ необходимо установить, какими свойствами должна обладать функция $f(x)$, определенная в окрестности точки $x=0$, чтобы в этой окрестности имело место равенство между функцией $f(x)$ и ее рядом Тейлора:

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

Во многих случаях справедливость равенства (2) можно установить, используя следующее утверждение.

Теорема 2 (достаточное условие разложимости функции в ряд Тейлора). *Пусть функция $f(x)$ определена и дифференцируема любое число раз в интервале $(-r, r)$. Если существует такое число M, что во всех точках указанного интервала выполняются неравенства*

$$|f^{(n)}(x)|<M \quad (n=0, 1, 2, \dots), \qquad (3)$$

то в этом интервале справедливо равенство (2).

□ Правая часть равенства (2) есть сумма ряда, т. е. предел последовательности его частичных сумм. Частичная сумма с произвольным номером n имеет вид

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

Таким образом, нужно доказать, что при всех $x\in(-r, r)$ выполняется соотношение $\lim\limits_{x\to\infty}F_n(x)=f(x)$ или, что то же самое, $\lim\limits_{n\to\infty}[f(x)-F_n(x)]=0$.

Выражение $F_n(x)$ представляет собой многочлен Тейлора n-й степени для функции $f(x)$. Многочлены Тейлора были введены в § 33 с целью получить приближенное представление функции $f(x)$ в виде многочлена той или иной степени. Там же рассматривалась и разность $f(x)-F_n(x)$, которую мы называли остаточным членом формулы Тейлора; для этой разности была получена формула

$$f(x)-F_n(x)=\frac{x^{n+1}}{(n+1)!}f^{(n+1)}(\xi), \qquad (4)$$

где ξ — некоторое число, заключенное между 0 и x. Из формулы (4) и неравенства (3) следует, что при всех x из интервала $(-r, r)$ справедливо неравенство

$$|f(x)-F_n(x)|\leqslant\frac{Mr^{n+1}}{(n+1)!}.$$

Последовательность чисел $a_n=\frac{Mr^n}{n!}$ стремится к нулю, поскольку отношение последующего числа к предыдущему равно

$$\frac{a_{n+1}}{a_n}=\frac{r^{n+1}}{(n+1)!}:\frac{r^n}{n!}=\frac{r}{n+1}$$

и при достаточно большом n это отношение меньше, чем, скажем, 1/2. Таким образом, при любом $x\in(-r, r)$ разность $f(x)-F_n(x)$ стремится к нулю при $n\to\infty$, что и доказывает равенство (2). ■

3. Разложения функций $\sin x$ и $\cos x$

Каждая из функций $\sin x$, $\cos x$ удовлетворяет условиям теоремы 2 (при любом r): действительно, производная любого порядка от $\sin x$ или $\cos x$ есть одна из функций $\pm\sin x$, $\pm\cos x$, а значит, $|(\sin x)^{(n)}|\leqslant 1$, $|(\cos x)^{(n)}|\leqslant 1$ $(n=0, 1, 2, \ldots)$. Таким образом, в любом интервале $(-r, r)$ каждая из функций $\sin x$ и $\cos x$ равна сумме своего ряда Тейлора.

Имеем

$$(\sin x)^{(0)}=\sin x,\ (\sin x)'=\cos x,\ (\sin x)''=-\sin x,\ (\sin x)'''=-\cos x,$$
$$(\sin x)^{(4)}=\sin x;$$
$$(\cos x)^{(0)}=\cos x,\ (\cos x)'=-\sin x,\ (\cos x)''=-\cos x,\ (\cos x)'''=\sin x,$$
$$(\cos x)^{(4)}=\cos x.$$

Видно, что каждая из последовательностей $\{(\sin x)^{(n)}\}$ и $\{(\cos x)^{(n)}\}$ «периодична» с «периодом» 4. При $x=0$ получаем

$$\sin 0=0,\ \sin'(0)=1,\ \sin''(0)=0,\ \sin'''(0)=-1;\ \cos 0=1,\ \cos'(0)=0,$$
$$\cos''(0)=-1,\ \cos'''(0)=0;$$

вообще,

$$\sin^{(2n)}(0)=0,\ \sin^{(2n+1)}(0)=(-1)^n,\ \cos^{(2n+1)}(0)=0,$$
$$\cos^{(2n)}(0)=(-1)^n\quad(n=0, 1, 2, \ldots).$$

Отсюда находим ряды Тейлора для $\sin x$ и $\cos x$:

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\cdots,\tag{5}$$

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots+(-1)^n\frac{x^{2n}}{(2n)!}+\cdots.\tag{6}$$

Как следует из предыдущего, эти разложения справедливы при любом значении x. В частности, это означает, что каждый из написанных рядов имеет радиус сходимости $R=\infty$; впрочем, это можно проверить и с помощью признака Даламбера.

Пользуясь разложениями (5) и (6), можно находить значения $\sin x$ и $\cos x$ с любой степенью точности.

4. Разложение функции e^x

Пусть $f(x)=e^x$. В любом интервале $(-r, r)$ имеем $|f^{(n)}(x)|=e^x<e^r$. В силу теоремы 2 отсюда следует, что функция e^x равна сумме своего ряда Тейлора при всех $x\in(-r, r)$, а значит (ввиду произвольности r), при всех x вообще. Учитывая, что $f^{(n)}(0)=e^0=1$ при любом n, получаем разложение

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+\cdots, \tag{7}$$

справедливое для всех x.

5. Разложение функции $\ln(1+x)$

Будем исходить из разложения

$$\frac{1}{1+t}=1-t+t^2-\cdots+(-1)^n t^n+\cdots, \tag{8}$$

справедливого при $|t|<1$ (мы записали формулу для суммы членов бесконечной геометрической прогрессии). Ряд (8) можно интегрировать почленно по любому отрезку $[0, x]$, где $|x|<1$. Получим

$$\int_0^x \frac{dt}{1+t}=\int_0^x dt-\int_0^x t\,dt+\int_0^x t^2dt+\cdots+(-1)^n\int_0^x t^n dt+\cdots$$

или

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots+(-1)^{n-1}\frac{x^n}{n}+\cdots. \tag{9}$$

Это равенство справедливо для всех x из интервала $(-1, 1)$; радиус сходимости полученного ряда равен 1.

Можно показать, что равенство (9) сохраняется и при $x=1$, т. е. что

$$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots.$$

Это вытекает из следующей теоремы, которую мы приводим без доказательства.

Теорема 3. *Если степенной ряд сходится во всех точках некоторого промежутка (интервала, полуинтервала, отрезка), то его сумма непрерывна в этом промежутке.*

6. Разложение функции $(1+x)^\alpha$

При натуральном α разложение функции $(1+x)^\alpha$ в ряд Тейлора совпадает с формулой бинома Ньютона:

$$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{1\cdot 2}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot 2\cdot 3}x^3+\cdots+\frac{\alpha(\alpha-1)(\alpha-2)\cdots 2\cdot 1}{1\cdot 2\ldots\alpha}x^\alpha.$$

Можно считать, что правая часть этого равенства представляет собой бесконечный степенной ряд

$$1+\alpha x+\frac{\alpha(\alpha-1)}{1\cdot 2}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot 2\cdot 3}x^2+\cdots.$$

Действительно, при натуральном α все члены такого ряда с номерами $\alpha+2$, $\alpha+3$ и т. д. равны нулю, поскольку каждый из них содержит (в числителе) произведение $\alpha(\alpha-1)(\alpha-2)\cdots2\cdot1\cdot0$.

Напрашивается предположение, что при любом α (не обязательно натуральном) справедливо разложение

$$(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{1\cdot2}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot2\cdot3}x^3+\cdots. \quad (10)$$

Можно показать, что такое разложение действительно имеет место, однако со следующей оговоркой: если число α не является натуральным, то равенство (10) справедливо лишь при $-1<x<1$.

§ 57. Степенные ряды с произвольным центром

1. Интервал сходимости

В § 55 мы дали определение степенного ряда как функционального ряда вида

$$a_0+a_1(x-x_0)+a_2(x-x_0)^2+\cdots+a_n(x-x_0)^n+\cdots, \quad (1)$$

однако в дальнейшем рассматривали только частный случай $x_0=0$. Общий случай легко сводится к указанному частному случаю введением новой переменной $X=x-x_0$. После такой замены ряд (1) принимает вид

$$a_0+a_1X+a_2X^2+\cdots+a_nX^n+\cdots, \quad (2)$$

т. е. превращается в степенной ряд с центром $x_0=0$. По доказанному, ряд (2) имеет интервал сходимости вида $(-R, R)$, т. е. сходится при $|X|<R$ и расходится при $|X|>R$. По отношению к ряду (1) это означает, что он сходится

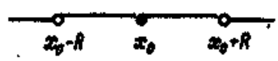


Рис. 199

при $|x-x_0|<R$ и расходится при $|x-x_0|>R$. Следовательно, *интервал сходимости ряда* (1) *есть* $(x_0-R,\ x_0+R)$ (рис. 199).

Все свойства, которыми обладает ряд (2) в интервале $(-R, R)$, переносятся на ряд (1) в интервале $(x_0-R,\ x_0+R)$. К ним относятся: непрерывность суммы ряда; возможность почленного интегрирования по любому отрезку, содержащемуся в интервале сходимости $(x_0-R,\ x_0+R)$; возможность почленного дифференцирования в указанном интервале.

2. Ряд Тейлора с произвольным центром

Аналогично задаче о разложении функции в ряд по степеням x ставится и задача о разложении по степеням $x-x_0$.

Определение. Пусть функция $f(x)$ определена в некоторой окрестности точки x_0 и имеет в этой точке производные всех порядков. Степенной ряд

$$f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\cdots$$

называется *рядом Тейлора* с центром x_0 для функции $f(x)$.

Справедливо следующее утверждение, обобщающее теорему 1 § 56.

Теорема. *Если функция $f(x)$ разлагается в некоторой окрестности точки x_0 в ряд по степеням $x-x_0$, то он является ее рядом Тейлора с центром x_0.*

□ Положим $X=x-x_0$ и $F(X)=f(x)$. Если функция $f(x)$ в окрестности точки x_0 разлагается в ряд (1), то функция $F(X)$ в окрестности точки 0 разлагается в ряд (2). Тогда согласно теореме 1 § 56 последний ряд должен быть рядом Тейлора для $F(X)$, т. е. должно иметь место разложение

$$F(X)=F(0)+\frac{F'(0)}{1!}X+\frac{F''(0)}{2!}X^2+\dots+\frac{F^{(n)}(0)}{n!}X^n+\cdots.$$

Возвращаясь к переменной x, отсюда получаем

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\dots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\cdots. \ ■$$

§ 58. Приложения степенных рядов к приближенным вычислениям

Возможность разложения функций в степенные ряды, сравнительная простота выражений для коэффициентов ряда Тейлора делают степенные ряды незаменимым средством приближенных вычислений. Ниже это положение иллюстрируется на нескольких примерах. Подчеркнем, что в каждом случае желательна оценка ошибки, получаемой при замене ряда той или иной частичной суммой.

1. Приближенное нахождение значений функций

Покажем, как решается эта задача, на примере логарифмической функции.

Согласно предыдущему (см. формулу (9) § 56), при $|x|<1$ имеет место разложение

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots.$$

Заменяя x на $-x$, получаем

$$\ln(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-\cdots.$$

Вычитая из первого равенства второе, находим

$$\ln\frac{1+x}{1-x}=2\left(x+\frac{x^3}{3}+\frac{x^5}{5}+\cdots\right), \qquad (1)$$

где $|x|<1$.

Формула (1) очень удобна для нахождения логарифмов. При этом следует иметь в виду, что остаток ряда (1) оценивается легко. Так как отношение последующего члена ряда к предыдущему равно

$$\frac{x^{2n+1}}{2n+1}:\frac{x^{2n-1}}{2n-1}=\frac{2n-1}{2n+1}x^2,$$

а значит, меньше x^2, то при $x>0$ остаток ряда после члена $\frac{x^{2n-1}}{2n-1}$ меньше, чем сумма членов геометрической прогрессии с первым членом $2\frac{x^{2n+1}}{2n+1}$ и знаменателем $q=x^2$, т. е. меньше, чем

$$\Delta_n=\frac{2x^{2n+1}}{(2n+1)(1-x^2)}. \qquad (2)$$

Пусть, например, требуется найти $\ln 3$. Полагая $\frac{1+x}{1-x}=3$, находим $x=\frac{1}{2}$. Так как $\frac{1}{2}<1$, то можем воспользоваться разложением (1) при $x=\frac{1}{2}$ и записать

$$\ln 3=2\left(\frac{1}{2}+\frac{1}{3}\left(\frac{1}{2}\right)^3+\frac{1}{5}\left(\frac{1}{2}\right)^5+\frac{1}{7}\left(\frac{1}{2}\right)^7+\ldots\right).$$

Если ограничиться только первыми четырьмя членами ряда, т. е. записать

$$\ln 3=2\left(\frac{1}{2}+\frac{1}{24}+\frac{1}{160}+\frac{1}{896}\right),$$

то согласно оценке (2) ошибка будет меньше, чем

$$\Delta_4=2\frac{1}{2^9}\cdot\frac{1}{9}\cdot\frac{1}{1-\frac{1}{4}}=\frac{1}{2^6}\cdot\frac{1}{27}=\frac{1}{1728}.$$

Заметим, кстати, что и само число e — основание натуральных логарифмов — можно находить с любой точностью при помощи разложения в ряд. В качестве примера вычислим e с точностью до 0,001. Имеем

$$e=1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots.$$

Остаток ряда после члена $\frac{1}{(n-1)!}$ оценивается так:

$$\frac{1}{n!}+\frac{1}{(n+1)!}+\cdots<\frac{1}{n!}\left(1+\frac{1}{n+1}+\frac{1}{(n+1)^2}+\cdots\right)=$$
$$=\frac{1}{n!}\frac{1}{1-\frac{1}{n+1}}=\frac{n+1}{n!n}.$$

Полагая $n=7$, находим, что остаток меньше $\frac{8}{7!\,7}$ и, следовательно, меньше 0,001. Итак, с точностью 0,001 получим

$$e\approx 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\frac{1}{6!}=2{,}718\ldots$$

2. Приближенное вычисление корней

Пусть, например, требуется вычислить $\sqrt[4]{650}$ с точностью до 0,001. Легко найти целую часть числа $\sqrt[4]{650}$; так как $5^4=625<650$, а $6^4=1296>650$, то целая часть равна 5. Запишем теперь равенство

$$\sqrt[4]{650}=\sqrt[4]{625+25}=\sqrt[4]{5^4\left(1+\frac{25}{625}\right)}=5\left(1+\frac{1}{25}\right)^{1/4}.$$

Применив разложение для $(1+x)^\alpha$ (см. формулу (10) § 56), получим

$$\sqrt[4]{650}=5\left[1+\frac{1}{4}\cdot\frac{1}{25}+\frac{\frac{1}{4}\left(\frac{1}{4}-1\right)}{1\cdot 2}\cdot\frac{1}{25^2}+\frac{\frac{1}{4}\left(\frac{1}{4}-1\right)\left(\frac{1}{4}-2\right)}{1\cdot 2\cdot 3}\cdot\frac{1}{25^3}+\cdots\right].$$

Члены ряда, записанного в квадратных скобках, имеют чередующиеся знаки, а их модули монотонно убывают (проверьте это, составив отношение последующего члена к предыдущему). Согласно теореме Лейбница, остаток ряда после n-го члена не превосходит по модулю $(n+1)$-го члена ряда. В частности, полагая

$$\sqrt[4]{650} \approx 5\left(1+\frac{1}{4\cdot 25}+\frac{\frac{1}{4}\left(\frac{1}{4}-1\right)}{1\cdot 2}\cdot\frac{1}{25^2}\right),$$

мы допустим ошибку, не превышающую по модулю числа

$$\frac{\frac{1}{4}\left(\frac{1}{4}-1\right)\left(\frac{1}{4}-2\right)}{1\cdot 2\cdot 3}\cdot\frac{1}{25^3}=\frac{1\cdot 3\cdot 7}{1\cdot 2\cdot 3}\cdot\frac{1}{100^3}=0{,}0000035.$$

3. Приближенное вычисление интегралов

Пусть требуется вычислить интеграл $I=\int_0^1 e^{-x^2}dx$. Неопределенный интеграл от функции e^{-x^2} не выражается в элементарных функциях, поэтому мы вынуждены ограничиться приближенным вычислением интеграла I. Найдем I с точностью до 0,01.

Используя разложение

$$e^x = 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+\cdots,$$

справедливое для всех x (см. формулу (7) § 56), находим

$$e^{-t^2} = 1-t^2+\frac{t^4}{2!}-\frac{t^6}{3!}+\cdots+(-1)^n\frac{t^{2n}}{n!}+\cdots.$$

Интегрируя этот ряд почленно в пределах от 0 до 1, имеем

$$I = 1-\frac{1}{3}+\frac{1}{2!}\cdot\frac{1}{5}-\frac{1}{3!}\cdot\frac{1}{7}+\cdots+(-1)^n\frac{1}{n!}\frac{1}{2n+1}+\cdots.$$

Для нахождения I с требуемой точностью достаточно взять 4 первых члена ряда. Тогда получим $I=1-\frac{1}{3}+\frac{1}{10}-\frac{1}{42}$, причем остаток по модулю не превосходит пятого члена ряда, т. е. числа $\frac{1}{4!}\cdot\frac{1}{9}=\frac{1}{24\cdot 9}=\frac{1}{216}<0{,}01$.

§ 59. Тригонометрические ряды (ряды Фурье)

Так называемые тригонометрические ряды (или ряды Фурье*) занимают после степенных рядов важнейшее место как в теоретическом отношении, так и по богатству приложений.

Тригонометрические ряды используются прежде всего при изучении периодических процессов. Любая величина $f(t)$, связанная с периодическим процессом, по истечении периода T возвращается к своему первоначальному значению, т. е. является периодической функцией с периодом T. Рассмотрим сначала случай $T=2\pi$. Наиболее известные периодические функции с периодом 2π — это (если не считать постоянных функций) $\cos x$ и $\sin x$. Разумеется,

---

* Ж. Фурье (1768—1830) — французский ученый.

косинус и синус кратного аргумента, т. е. $\cos 2x$; $\sin 2x$, $\cos 3x$, $\sin 3x$, ..., также имеют период 2π (хотя и не наименьший). Условимся называть периодические функции

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cos 3x,\ \sin 3x,\ \ldots$$

простейшими функциями с периодом 2π.

Любая линейная комбинация простейших функций, т. е. любая функция вида

$$a_0+a_1\cos x+b_1\sin x+a_2\cos 2x+b_2\sin 2x+\cdots \tag{1}$$

также является периодической с периодом 2π; это справедливо как в отношении суммы (1), состоящей из конечного числа слагаемых, так и в отношении ряда (1). Естественно поставить вопрос: можно ли данную периодическую функцию с периодом 2π представить в виде суммы конечного или бесконечного множества простейших, иначе говоря, в виде суммы ряда (1)? Этот вопрос является основополагающим для всей теории тригонометрических рядов.

1. Ортогональность системы простейших функций

Определение 1. Функции $f_1(x)$ и $f_2(x)$, определенные на отрезке $[a, b]$, называются *ортогональными* друг другу на этом отрезке, если

$$\int_a^b f_1(x)f_2(x)\,dx=0. \tag{2}$$

Отметим, что условие (2) есть обобщение равенства $a_1a_2+b_1b_2+c_1c_2=0$, выражающего ортогональность векторов $\{a_1, b_1, c_1\}$ и $\{a_2, b_2, c_2\}$ в прямоугольной декартовой системе координат.

Теорема 1. *Любые две функции из совокупности* $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$ *ортогональны друг другу на отрезке* $[0, 2\pi]$.

□ Заметим сначала, что $1=\cos 0x$. Таким образом, необходимо проверить справедливость следующих равенств:

$$\int_0^{2\pi}\cos mx\cos nx\,dx=0\quad (m,\ n=0,\ 1,\ 2,\ \ldots;\ m\neq n); \tag{3}$$

$$\int_0^{2\pi}\sin mx\sin nx\,dx=0\quad (m,\ n=1,\ 2,\ \ldots;\ m\neq n); \tag{4}$$

$$\int_0^{2\pi}\cos mx\sin nx\,dx=0\quad (m=0,\ 1,\ 2,\ \ldots;\ n=1,\ 2,\ \ldots). \tag{5}$$

На основании известных тригонометрических формул имеем

$$\cos mx\cos nx=\frac{1}{2}[\cos(m+n)x+\cos(m-n)x],$$

$$\sin mx\sin nx=\frac{1}{2}[\cos(m-n)x-\cos(m+n)x],$$

$$\cos mx\sin nx=\frac{1}{2}[\sin(n+m)x+\sin(n-m)x].$$

Поэтому для проверки равенств (3) — (5) достаточно убедиться, что каждый из интегралов $\int_0^{2\pi}\cos kx\,dx$, $\int_0^{2\pi}\sin lx\,dx$, где k и l — любые целые числа, причем $k\neq 0$, равен нулю. Рекомендуем сделать это самостоятельно. ■

2. Тригонометрические ряды. Коэффициенты Фурье и ряд Фурье данной функции

Определение 2. Функциональный ряд вида

$$\frac{a_0}{2}+(a_1\cos x+b_1\sin x)+(a_2\cos 2x+b_2\sin 2x)+\cdots, \tag{6}$$

или короче, ряд

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

называется *тригонометрическим рядом*, а постоянные числа a_0, a_1, b_1, ... называются *коэффициентами* этого ряда.

Из соображений удобства начальный член ряда обозначен $\frac{a_0}{2}$ (а не просто a_0).

Обратимся теперь к вопросу, поставленному в начале этого параграфа: о разложении заданной функции $f(x)$ в ряд (6). Предположим, что $f(x)$ — периодическая функция с периодом 2π и что для всех x справедливо разложение

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx). \tag{7}$$

Как найти коэффициенты a_0, a_1, b_1, a_2, b_2, ...?

Теорема 2. *Если для всех x имеет место равенство (7), причем ряд в правой части этого равенства сходится правильно на всей числовой прямой, то справедливы формулы*

$$a_n=\frac{1}{\pi}\int_0^{2\pi}f(x)\cos nx\,dx \quad (n=0,\ 1,\ 2,\ \ldots); \tag{8}$$

$$b_n=\frac{1}{\pi}\int_0^{2\pi}f(x)\sin nx\,dx \quad (n=1,\ 2,\ \ldots). \tag{9}$$

Заметим, что именно для того чтобы формула (8) была справедлива для всех n, включая $n=0$, начальный член ряда (7) и обозначен $a_0/2$, а не a_0.

☐ Как известно (см. теорему 2 § 54) правильно сходящийся функциональный ряд можно интегрировать почленно. Отсюда следует, что

$$\int_0^{2\pi}f(x)\,dx=\frac{a_0}{2}\int_0^{2\pi}dx+\sum_{n=1}^{\infty}\left(a_n\int_0^{2\pi}\cos nx\,dx+b_n\int_0^{2\pi}\sin nx\,dx\right).$$

Первый член ряда, стоящего в правой части, равен πa_0, остальные же члены равны нулю. Поэтому

$$a_0=\frac{1}{\pi}\int_0^{2\pi}f(x)\,dx,$$

что совпадает с формулой (8) при $n=0$.

Умножив теперь обе части равенства (7) на $\cos kx$, где k — целое положительное число, получим

$$f(x)\cos kx=\frac{a_0}{2}\cos kx+\sum_{n=1}^{\infty}(a_n\cos nx\cos kx+b_n\sin nx\cos kx).$$

Ряд в правой части последнего равенства сходится правильно, так как каждый его член отличается от соответствующего члена ряда (7) множителем $\cos kx$, модуль которого не превосходит 1. Поэтому возможно почленное интегрирование этого ряда:

$$\int_0^{2\pi} f(x)\cos kx\,dx = \frac{a_0}{2}\int_0^{2\pi}\cos kx\,dx + \sum_{n=1}^{\infty}\left(a_n\int_0^{2\pi}\cos nx\cos kx\,dx + b_n\int_0^{2\pi}\sin nx\cos kx\,dx\right).$$

Первое слагаемое в правой части равно нулю, поскольку $\int_0^{2\pi}\cos kx\,dx=0$.

Что касается остальных слагаемых, то в силу соотношений (3) и (5) отличным от нуля среди них может быть только одно, а именно

$$a_k\int_0^{2\pi}\cos kx\cos kx\,dx = a_k\int_0^{2\pi}\cos^2 kx\,dx = \pi a_k.$$

Итак, $\pi a_k = \int_0^{2\pi} f(x)\cos kx\,dx$, что и завершает проверку равенства (8).

Чтобы установить справедливость равенства (9), достаточно умножить обе части разложения (7) на $\sin kx$ ($k=1, 2, \ldots$) и взять интеграл от обеих частей полученного равенства по отрезку $[0, 2\pi]$. ■

Определение 3. Пусть $f(x)$ — функция, заданная на отрезке $[0, 2\pi]$. Числа a_n ($n=0, 1, 2, \ldots$) и b_n ($n=1, 2, \ldots$), определенные формулами (8) и (9), называются *коэффициентами Фурье* функции $f(x)$, а тригонометрический ряд $\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$, коэффициентами которого служат эти числа, — *рядом Фурье* функции $f(x)$.

Учитывая данное определение, теорему 2 можно сформулировать следующим образом.

Теорема 3. *Если функция $f(x)$ разлагается в правильно сходящийся тригонометрический ряд, то этот ряд является ее рядом Фурье.*

Аналогичное положение имеет место в случае степенных рядов: согласно теореме из § 57, если в некоторой окрестности точки x_0 функция $f(x)$ разлагается в ряд по степеням $x-x_0$, то этот ряд не может быть ничем иным, как ее рядом Тейлора с центром x_0.

3. Разложение функций в тригонометрические ряды

С учетом теоремы 3 вопрос о возможности разложения заданной функции $f(x)$ в тригонометрический ряд сводится к следующему: какими свойствами должна обладать функция $f(x)$, чтобы построенный для нее ряд Фурье сходился и его сумма совпадала с $f(x)$? При выяснении ответа на этот вопрос обнаруживается резкое различие между тригонометрическими и степенными рядами. В то время как в степенные ряды могут разлагаться лишь функции, имеющие производные всех порядков, в тригонометрический ряд разлагается, грубо говоря, почти любая функция.

Условимся называть функцию $f(x)$ *кусочно монотонной* на отрезке $[a, b]$, если этот отрезок с помощью конечного числа точек $x_1, x_2, \ldots, x_{n-1}$ можно разбить на отрезки $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$ (где $x_0=a$, $x_n=b$), внутри каж-

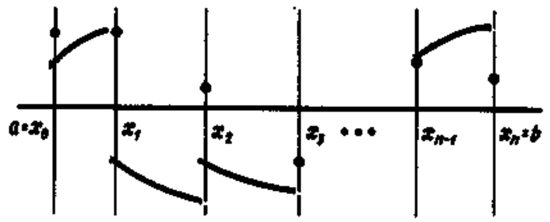


Рис. 200

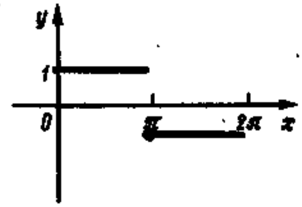


Рис. 201

дого из которых (т. е. в соответствующем интервале (x_{i-1}, x_i)) функция $f(x)$ непрерывна и монотонна.

На рис. 200 изображен график некоторой кусочно монотонной функции $f(x)$. Обратим внимание на точки, соответствующие значениям функции при $x=x_0, x_1, \ldots, x_n$, т. е. числам $f(x_0), f(x_1), \ldots, f(x_n)$. В принципе эти числа могут быть какими угодно; они никак не должны быть согласованы со значениями $f(x)$ вблизи точек $x_0, x_1, \ldots, x_n$.

Из данного определения следует, что если функция $f(x)$ кусочно монотонна, то она может иметь разрывы лишь в точках $x_0, x_1, \ldots, x_n$, причем каждый из них есть разрыв первого рода. Действительно, если в точке x_i имеет место разрыв, то ввиду монотонности $f(x)$ слева от x_i существует предел $\lim\limits_{x\to x_i-0} f(x)=$ $=f(x_i-0)$, а ввиду монотонности справа — предел $\lim\limits_{x\to x_i+0} f(x)=f(x_i+0)$.

Достаточные условия разложимости функции в ряд Фурье дает следующая теорема, которую мы примем без доказательства.

Теорема 4 (теорема Дирихле). *Если периодическая функция $f(x)$ с периодом 2π ограничена и кусочно монотонна на отрезке $[0, 2\pi]$, то ряд Фурье, построенный для $f(x)$, сходится во всех точках. Сумма $F(x)$ этого ряда равна $f(x)$ в точках непрерывности функции $f(x)$. Если же c есть точка разрыва $f(x)$, то*

$$F(c)=\frac{f(c+0)+f(c-0)}{2}. \tag{10}$$

Заметим, что если c — точка непрерывности $f(x)$, то равенство (10) также выполняется. В этом случае $f(c-0)=f(c+0)=f(c)$ и равенство (10) принимает вид $F(c)=f(c)$.

Теорема Дирихле показывает, что класс функций, представимых рядами Фурье, очень широк. Поэтому ряды Фурье находят применение в самых различных областях науки и практики, особенно — в математической физике.

4. Примеры разложения функций в тригонометрические ряды

Предварительно докажем следующее утверждение: *если функция $f(x)$ имеет период 2π, то интеграл от нее по любому отрезку длины 2π имеет одно и то же значение*, т. е.

$$\int_a^{a+2\pi} f(x)\,dx=\int_0^{2\pi} f(x)\,dx.$$

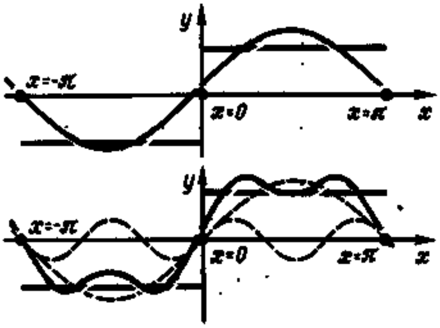


Рис. 202

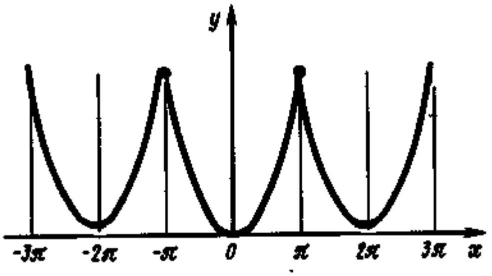


Рис. 203

□ Имеем

$$\int_a^{a+2\pi} f(x)dx=\int_a^0 f(x)dx+\int_0^{2\pi} f(x)dx+\int_{2\pi}^{2\pi+a} f(x)dx. \tag{11}$$

Если в последнем из интегралов правой части сделать замену переменной $x=2\pi+t$, то интеграл примет вид $\int_0^a f(2\pi+t)dt=\int_0^a f(t)dt$, откуда следует, что сумма первого и третьего интегралов в правой части равенства (11) равна нулю. ■

Примеры. 1. Периодическая функция с периодом 2π в полуинтервале $[0, 2\pi)$ задана формулами

$$f(x)=\begin{cases} 1 \text{ при } 0\leqslant x\leqslant \pi, \\ -1 \text{ при } \pi<x<2\pi \end{cases}$$

(рис. 201). Разложить $f(x)$ в тригонометрический ряд.

○ Вычислим коэффициенты Фурье. С учетом сделанного выше замечания интегралы по $[0, 2\pi]$ можно заменить соответствующими интегралами по $[-\pi, \pi]$. Имеем

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx.$$

Подынтегральная функция, если отвлечься от ее значения при $x=0$, — нечетная; поэтому $a_0=0$. Далее, находим

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx \quad (n=1, 2, \ldots);$$

снова используя нечетность подынтегральной функции, получим $a_n=0$. Наконец,

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx=\frac{1}{\pi}\left[\int_{-\pi}^{0}(-1)\sin nx\,dx+\int_0^{\pi}\sin nx\,dx\right]=$$

$$=\frac{1}{\pi}\left[\left.\frac{\cos nx}{n}\right|_{-\pi}^{0}-\left.\frac{\cos nx}{n}\right|_0^{\pi}\right]=\frac{2}{\pi n}(\cos 0-\cos n\pi).$$

Отсюда следует

$$b_n=\begin{cases} 0, & \text{если } n \text{ — четное;} \\ \dfrac{4}{\pi n}, & \text{если } n \text{ — нечетное.} \end{cases}$$

Итак, разложение $f(x)$ в ряд Фурье имеет вид

$$f(x)=\frac{4}{\pi}\left[\frac{\sin x}{1}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots\right]. \tag{12}$$

В частности, при $x=\frac{\pi}{2}$ получаем $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$.

На рис. 202 изображены графики нескольких частичных сумм ряда (12); видно, что с увеличением номера частичная сумма все точнее представляет $f(x)$. ●

2. Функция $f(x)$ на отрезке $[-\pi, \pi]$ определена формулой $f(x)=x^2$, а на остальные значения x продолжена периодически (рис. 203). Разложить $f(x)$ в ряд Фурье.

○ Имеем

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} x^2\sin nx\,dx=0 \quad (n=1, 2, \ldots),$$

поскольку подынтегральная функция — нечетная. Далее, находим

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi} x^2dx=\frac{1}{\pi}\left.\frac{x^3}{3}\right|_{-\pi}^{\pi}=\frac{2\pi^2}{3},$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos nx\,dx=\frac{1}{\pi}\left[\left.\frac{x^2\sin nx}{n}\right|_{-\pi}^{\pi}-\frac{2}{n}\int_{-\pi}^{\pi} x\sin nx\,dx\right]=$$

$$=-\frac{2}{n\pi}\left[\left.-\frac{x\cos nx}{n}\right|_{-\pi}^{\pi}+\frac{1}{n}\int_{-\pi}^{\pi}\cos nx\,dx\right]=\frac{4}{n^2}\cos n\pi \quad (n=1, 2, \ldots).$$

Таким образом,

$$a_n=\begin{cases}\frac{4}{n^2}, & \text{если } n \text{ — четное;}\\ -\frac{4}{n^2}, & \text{если } n \text{ — нечетное.}\end{cases}$$

Разложение в ряд Фурье имеет вид

$$f(x)=\frac{\pi^2}{3}-4\left[\frac{\cos x}{1}-\frac{\cos 2x}{2^2}+\frac{\cos 3x}{3^2}-\frac{\cos 4x}{4^2}+\cdots\right].$$

В частности, при $x=\pi$ получаем $\frac{\pi^2}{6}=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$. ●

5. Разложение в ряд Фурье четной и нечетной функций. Разложение функции, заданной на отрезке $[0,\pi]$

В рассмотренных выше примерах обращает на себя внимание следующее обстоятельство: в разложение функции из примера 1 входят только слагаемые вида $b_n\sin nx$, а в разложение функции из примера 2 — только слагаемые вида $a_n\cos nx$. Чем обусловлены эти особенности разложений? Ответ заключается в том, что функция из примера 1 является нечетной, а из примера 2 — четной.

Справедливо следующее утверждение: *в случае нечетной функции $f(x)$ все коэффициенты a_n ее ряда Фурье равны нулю, в случае четной функции $f(x)$ равны нулю все коэффициенты b_n.*

□ Если $f(x)$ — нечетная функция, то функция $f(x)\cos nx$ — также нечетная; следовательно, $a_n=0$. Заметим, что в этом случае функция $f(x)\sin nx$ — четная, поэтому интеграл от $-\pi$ до π можно заменить удвоенным интегралом от 0 до π:

$$b_n=\frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx. \tag{13}$$

Если же $f(x)$ — четная функция, то функция $f(x)\sin nx$ — нечетная; следовательно, $b_n=0$. В этом случае, подобно предыдущему,

$$a_n=\frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\,dx. \quad \blacksquare \tag{14}$$

Согласно теореме 2, для любой периодической функции $f(x)$ с периодом 2π имеется не более чем один (правильно) сходящийся к ней тригонометрический ряд: им может быть только ряд Фурье для $f(x)$. Однако если $f(x)$ первоначально задана не на всей числовой прямой, а лишь на некотором интервале, длина которого меньше, чем 2π, то для нее существует не одно, а бесчисленное множество разложений в тригонометрический ряд. Действительно, пусть $f(x)$ задана на интервале $(0, a)$, где $a<2\pi$. Продолжим $f(x)$ произвольным образом на промежуток $[0, 2\pi)$, а затем с периодом 2π — на всю числовую прямую (рис. 204). Полученное указанным образом продолжение функции $f(x)$ обозначим $\tilde{f}(x)$. Функцию $\tilde{f}(x)$ как периодическую с периодом 2π можно разложить в ряд Фурье; последний на интервале $(0, a)$ сходится к $f(x)$ (во всех точ-

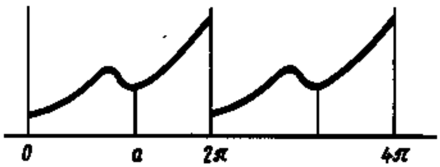


Рис. 204

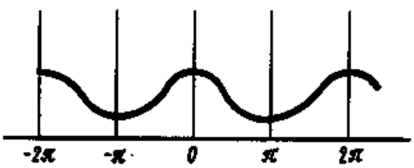


Рис. 205

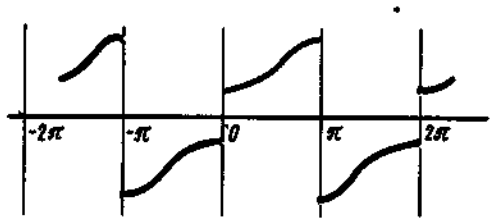


Рис. 206

ках, где $f(x)$ непрерывна). Так как продолжение $\tilde{f}(x)$ первоначально заданной функции $f(x)$ может быть выбрано бесчисленным множеством способов, то существует бесчисленное множество тригонометрических рядов, сходящихся на интервале $(0, a)$ к $f(x)$.

Возможность варьировать разложение $f(x)$ в тригонометрический ряд приобретает особое значение для функции $f(x)$, заданной на интервале $(0, \pi)$ (вообще, на интервале длины π). В этом случае среди различных продолжений функций $f(x)$ имеются два особых продолжения $f_ч(x)$ и $f_н(x)$, где $f_ч(x)$ — четная функция, а $f_н(x)$ — нечетная. Чтобы получить $f_ч(x)$, сначала продолжим функцию $f(x)$ четным образом на интервал $(-\pi, 0)$, а затем полученную функцию — периодически (с периодом 2π) на всю числовую прямую (рис. 205). Аналогично получается $f_н(x)$: сначала продолжим $f(x)$ нечетным образом на интервал $(-\pi, 0)$, а затем полученную функцию — периодически на всю числовую прямую (рис. 206).

Поскольку $f_ч(x)$ — четная функция, ее разложение в ряд Фурье состоит только из слагаемых $a_n\cos nx$; это дает разложение $f(x)$ (на интервале $(0, \pi)$) в тригонометрический ряд «по косинусам». Так как $f_н(x)$ — нечетная функция, то ее разложение в ряд Фурье состоит только из слагаемых $b_n \sin nx$; это приводит к разложению $f(x)$ в тригонометрический ряд «по синусам». Итак, функцию $f(x)$, заданную на интервале длины π, можно разложить (по желанию) в тригонометрический ряд по одним косинусам или по одним синусам.

Пример. Разложить функцию $f(x)=x$, заданную в интервале $(0, \pi)$, в ряд Фурье по синусам.

○ Продолжим функцию $f(x)$ на интервал $(-\pi, 0)$ нечетным образом, а затем полученную функцию — периодически с периодом 2π на всю числовую прямую; получим функцию $f_н(x)$ (значение $f_н(\pi)$ можно выбрать произвольно, например взять $f_н(\pi)=0$). В ряд Фурье для $f_н(x)$ войдут только слагаемые $b_n \sin nx$, причем согласно формуле (13) коэффициенты b_n определяются так:

$$b_n=\frac{2}{\pi}\int_0^\pi x\sin nx\,dx \quad (n=1, 2, \ldots).$$

Интегрируя по частям, находим

$$\int_0^\pi x\sin nx\,dx=-x\frac{\cos nx}{n}\Big|_0^\pi+\frac{1}{n}\int_0^\pi \cos nx\,dx=-\pi\frac{\cos n\pi}{n}+\frac{1}{n^2}\sin nx\Big|_0^\pi=-\pi\frac{\cos n\pi}{n}.$$

Следовательно,

$$b_n=-\frac{2}{n}\cos n\pi=\begin{cases}-\dfrac{2}{n}, & \text{если } n \text{ — четное;}\\ \dfrac{2}{n}, & \text{если } n \text{ — нечетное.}\end{cases}$$

Таким образом, при $0<x<\pi$ имеем

$$x=2\left(\frac{\sin x}{1}-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\frac{\sin 4x}{4}+\cdots\right). \ ●$$

6. Тригонометрический ряд для функции с произвольным периодом $2l$

До сих пор мы рассматривали вопрос о разложении в тригонометрический ряд периодической функции с периодом $T=2\pi$. Рассмотрим теперь тот же вопрос для функций, имеющих произвольный период $T=2l$. Для этого произведем замену переменной:

$$x=\frac{l}{\pi}t, \qquad t=\frac{\pi}{l}x. \tag{15}$$

Функция $f(x)=f\left(\frac{l}{\pi}t\right)=\varphi(t)$ по аргументу t имеет период 2π, поскольку

$$\varphi(t+2\pi)=f\left(\frac{l}{\pi}(t+2\pi)\right)=f\left(\frac{l}{\pi}t+2l\right)=f\left(\frac{l}{\pi}t\right)=\varphi(t).$$

Следовательно, $\varphi(t)$ можно разложить в ряд Фурье с периодом 2π:

$$\varphi(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nt+b_n\sin nt), \tag{16}$$

где коэффициенты a_n, b_n находятся по формулам

$$a_n=\frac{1}{\pi}\int_0^{2\pi}\varphi(t)\cos nt\,dt \quad (n=0, 1, 2, \ldots); \tag{17}$$

$$b_n=\frac{1}{\pi}\int_0^{2\pi}\varphi(t)\sin nt\,dt \quad (n=1, 2, \ldots). \tag{18}$$

Рис. 207

Возвращаясь к «старой» переменной x по формулам (15), получаем вместо (16) разложение

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos n\frac{\pi}{l}x+b_n\sin n\frac{\pi}{l}x\right), \tag{19}$$

а вместо (17), (18) — формулы

$$a_n=\frac{1}{l}\int_0^{2l}f(x)\cos n\frac{\pi}{l}x\,dx \quad (n=0,1,2,\ldots), \tag{20}$$

$$b_n=\frac{1}{l}\int_0^{2l}f(x)\sin n\frac{\pi}{l}x\,dx \quad (n=1,2,\ldots). \tag{21}$$

Формулы (19)—(21) дают решение задачи о разложении функции $f(x)$, имеющей период $2l$. «Простейшими» функциями периода $2l$, с помощью которых осуществляется это разложение, являются 1, $\cos\omega x$, $\sin\omega x$, $\cos 2\omega x$, $\sin 2\omega x$, ..., где $\omega=\pi/l$.

Пример. Разложить в ряд Фурье периодическую функцию $f(x)$ с периодом $2l$, которая на отрезке $[-l, l]$ задается формулой $f(x)=|x|$ (рис. 207).

○ Так как функция $f(x)$ — четная, то все коэффициенты b_n равны нулю, а

$$a_n=\frac{2}{l}\int_0^{l}f(x)\cos n\frac{\pi}{l}x\,dx \quad (n=0,1,2,\ldots).$$

В частности,

$$a_0=\frac{2}{l}\int_0^{l}x\,dx=l,$$

а при $n\neq 0$ находим

$$a_n=\frac{2}{l}\int_0^{l}x\cos n\frac{\pi}{l}x\,dx=\frac{2}{l}\left[\frac{l}{n\pi}x\sin n\frac{\pi}{l}x\Big|_0^l-\frac{l}{n\pi}\int_0^{l}\sin n\frac{\pi}{l}x\,dx\right]=$$

$$=\frac{2}{l}\frac{l^2}{(n\pi)^2}\cos n\frac{\pi}{l}x\Big|_0^l=\frac{2l}{n^2\pi^2}(\cos n\pi-\cos 0)=\begin{cases}0, & \text{если } n \text{ — четное;}\\ -\dfrac{4l}{n^2\pi^2}, & \text{если } n \text{ — нечетное.}\end{cases}$$

Итак, искомое разложение имеет вид

$$|x|=\frac{l}{2}-\frac{4l}{\pi^2}\left(\cos\frac{\pi}{l}x+\frac{1}{3^2}\cos 3\frac{\pi}{l}x+\frac{1}{5^2}\cos 5\frac{\pi}{l}x+\cdots\right),$$

где $-l\leqslant x\leqslant l$. ●

9

Элементы теории функций комплексной переменной

§ 60. Построение системы комплексных чисел

1. Необходимость введения комплексных чисел

Комплексные числа возникают в связи с задачей решения квадратных уравнений. Известно, что не всякое квадратное уравнение можно решить, оставаясь в области **R** действительных чисел. Простейшее неразрешимое в **R** квадратное уравнение имеет вид

$$x^2+1=0. \tag{1}$$

Попытаемся встать на новую для себя точку зрения: будем считать, что уравнение (1) на самом деле разрешимо, но его корень не является действительным числом, а представляет собой какое-то новое число. Обозначим это число символом i. Итак, мы вводим новое число i, которому приписываем необычное свойство: $i^2=-1$. Теперь, помимо действительных чисел, обозначаемых, как обычно, a, b, c и т. д., у нас имеется новое число i.

Операция умножения, примененная к действительному числу b и числу i, приводит к числам вида $bi(b\in \mathbf{R})$, а операция сложения — к числам $a+bi$, где $a\in\mathbf{R}$, $b\in\mathbf{R}$. Итак, введение нового числа i влечет за собой необходимость рассматривать числа вида $a+bi$. Эти числа называются *комплексными;* при этом слагаемое a называется *действительной частью,* а слагаемое bi — *мнимой частью* комплексного числа. Множество всех комплексных чисел обозначается **C**.

Примечательно, что основные арифметические операции — сложение и умножение — уже не выводят нас за пределы множества **C**, т. е. не заставляют вводить каких-то новых чисел. В самом деле,

$$(a_1+b_1i)+(a_2+b_2i)=a_1+a_2+b_1i+b_2i=(a_1+a_2)+(b_1+b_2)i; \tag{2}$$

$$(a_1+b_1i)(a_2+b_2i)=a_1a_2+b_1b_2i^2+a_1b_2i+a_2b_1i=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i \tag{3}$$

(здесь учтено, что $i^2=-1$).

Именно таким путем и вводились первоначально комплексные числа. Разумеется, этот способ построения системы комплексных чисел порождает много вопросов: что же все-таки представляет собой новое число i? можно ли распространять на него обычные законы арифметики? законно ли рассматривать выражения, содержащие вместе действительные числа и число i? и т. д. Без ответа на эти вопросы вся теория комплексных чисел остается как бы плодом чистого воображения. Иначе говоря, возникает задача строгого и полного построения теории комплексных чисел. Как решается эта задача, мы укажем в следующих пунктах, при этом вопрос о законности комплексных чисел как бы снимется сам собой.

Строгое обоснование теории комплексных чисел важно еще и потому, что эти числа крайне необходимы в целом ряде приложений математики. В настоящее время теория функций комплексной переменной представляет собой действенный инструмент применения математических методов в физике (механике, электро- и радиотехнике, аэро- и гидродинамике и т. д.).

2. Основные определения

Упомянутая в п. 1 запись комплексного числа в виде $a+bi$ наводит на мысль о возможности задания комплексного числа упорядоченной парой a, b действительных чисел. В соответствии с этим введем такое определение.

Определение 1. *Комплексным числом* называется любая упорядоченная пара $(a;\ b)$ действительных чисел.

Множество всех комплексных чисел обозначается **C**.

Определение 2. Два комплексных числа $(a;\ b)$ и $(a';\ b')$ считают *равными* и пишут $(a;\ b)=(a';\ b')$, если $a=a'$ и $b=b'$.

Условимся в дальнейшем обозначать комплексные числа греческими буквами α, β, γ, ...

Определение 3. Пусть $\alpha=(a_1;\ b_1)$ и $\beta=(a_2;\ b_2)$ — два комплексных числа. *Сумма* $\alpha+\beta$ определяется равенством

$$\alpha+\beta=(a_1+a_2;\ b_1+b_2), \tag{4}$$

а *произведение* $\alpha\beta$ — равенством

$$\alpha\beta=(a_1a_2-b_1b_2;\ a_1b_2+a_2b_1). \tag{5}$$

Такое определение действий над комплексными числами подсказано равенствами (2) и (3), играющими роль «наводящих соображений».

Например:

$$(-3;\ 2)+(3;\ 4)=(0;\ 6);$$
$$(-3;\ 2)(3;\ 4)=((-3)3-2\cdot4;\ (-3)\cdot4+2\cdot3)=(-17;\ -6);$$
$$(0;\ 1)(0;\ 1)=(0\cdot0-1\cdot1,\ 0\cdot1+1\cdot0)=(-1;\ 0).$$

Операции сложения и умножения комплексных чисел обладают следующими свойствами:

1^0. $\alpha+\beta=\beta+\alpha$ *(коммутативность сложения)*.

2^0. $\alpha\beta=\beta\alpha$ *(коммутативность умножения)*.

3^0. $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ *(ассоциативность сложения)*.

4^0. $(\alpha\beta)\gamma=\alpha(\beta\gamma)$ *(ассоциативность умножения)*.

5^0 $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$ *(дистрибутивность умножения относительно сложения)*.

Свойства 1^0 и 3^0 очевидны, а 2^0, 4^0, 5^0 нуждаются в проверке.

Покажем, например, справедливость свойства 2^0. Имеем

$$\alpha\beta=(a_1a_2-b_1b_2;\ a_1b_2+a_2b_1),\quad \beta\alpha=(a_2a_1-b_2b_1;\ b_1a_2+b_2a_1),$$

откуда непосредственно следует $\alpha\beta=\beta\alpha$.

Так же просто проверяется свойство 5^0. Пусть $\gamma=(a_3;b_3)$; тогда

$$\alpha(\beta+\gamma)=(a_1;\ b_1)(a_2+a_3;\ b_2+b_3)=$$

$$=(a_1(a_2+a_3)-b_1(b_2+b_3);\ a_1(b_2+b_3)+b_1(a_2+a_3)),$$
$$\alpha\beta+\alpha\gamma=(a_1a_2-b_1b_2;\ a_1b_2+a_2b_1)+(a_1a_3-b_1b_3;\ a_1b_3+a_3b_1)=$$
$$=(a_1a_2-b_1b_2+a_1a_3-b_1b_3;\ a_1b_2+a_2b_1+a_1b_3+a_3b_1).$$

Сравнивая результаты двух вычислений, убеждаемся в справедливости равенства $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$.

Проверку свойства 4^0 рекомендуем провести самостоятельно.

3. Множество комплексных чисел как расширение множества действительных чисел

Рассмотрим комплексные числа вида $(a;0)$. Множество, состоящее из всех таких чисел, обозначим $\mathbf{C}^*$. Очевидно, $\mathbf{C}^*\subset\mathbf{C}$.

Если каждому действительному числу a сопоставим комплексное число $(a;0)$, т. е. $a\to(a;0)$, то получим некоторое соответствие между множеством $\mathbf{R}$ и множеством $\mathbf{C}^*$. Это соответствие, очевидно, является взаимно однозначным.

Для любых двух действительных чисел a и b имеем

$$a+b\to(a+b;0). \qquad (6)$$

Если учесть, что $(a+b;0)=(a;0)+(b;0)$, то вместо (6) можем записать $a+b\to(a;0)+(b;0)$. Это означает, что сумме действительных чисел a и b отвечает сумма соответствующих им комплексных чисел.

То же самое относится к произведению: так как $ab\to(ab;0)$ и $(ab;0)=$ $=(a;0)\,(b;0)$ (см. формулу (5)), то $ab\to(a;0)\,(b;0)$. Итак, произведению действительных чисел a и b отвечает произведение соответствующих им комплексных чисел.

Из сказанного следует, что если отождествить каждое действительное число a с комплексным числом $(a;0)$, то тем самым множество $\mathbf{R}$ действительных чисел с его обычной арифметикой окажется как бы вложенным в множество $\mathbf{C}$ комплексных чисел. В этом смысле говорят, что *множество комплексных чисел является расширением множества действительных чисел.*

4. Представление комплексных чисел в виде $a+bi$

Особую роль среди комплексных чисел играет, как мы увидим дальше, число $(0;1)$. Обозначим его буквой i. Согласно закону умножения комплексных чисел, имеем

$$(0;b)=(b;0)\,(0;1),$$

где b — любое действительное число; отсюда

$$(a;b)=(a;0)+(0;b)=(a;0)+(b;0)i.$$

Условимся в дальнейшем не делать различия между комплексным числом вида $(a;0)$ и действительным числом a; основанием для такого соглашения являются, как уже отмечалось, одинаковые «арифметики» в множествах $\mathbf{R}$ и $\mathbf{C}^*$. Тогда можем записать

$$(a;b)=a+bi.$$

Таким образом, мы пришли к представлению комплексного числа в виде $a+bi$. Именно такое представление послужило в п. 1 отправной точкой к (нестрогому) построению множества комплексных чисел.

Теперь можно забыть о первоначальном способе задания комплексного числа как пары $(a;b)$ и записывать комплексное число в виде $a+bi$.

Если, используя новую запись $a+bi$, переписать правила сложения и умножения комплексных чисел (4) и (5), то получим

$$(a_1+b_1i)+(a_2+b_2i)=(a_1+a_2)+(b_1+b_2)i, \quad (7)$$

$$(a_1+b_1i)(a_2+b_2i)=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i, \quad (8)$$

что совпадает с правилами действий над комплексными числами, указанными в п. 1.

В дальнейшем комплексное число $a+0i$ будем для краткости записывать просто как a, а комплексное число $0+bi$ — как bi. В частности, $i=0+1i$. Имеем

$$i^2=i\cdot i=(0+1i)(0+1i)=-1+0i=-1.$$

Таким образом, в системе комплексных чисел уравнение $x^2+1=0$ разрешимо: одно из решений этого уравнения есть $x=i$.

Определение 4. Если задано комплексное число $\alpha=a+bi$, то действительное число a называется *действительной частью* числа α, а комплексное число bi — *мнимой частью* числа α.

Действительную часть числа α обозначают $\mathrm{Re}(\alpha)$ (от франц. reele — «действительный»), а коэффициент при мнимой части обозначают $\mathrm{Im}(\alpha)$ (от франц. imaginaire — «мнимый»). Например, $\mathrm{Re}(1-3i)=1$, $\mathrm{Im}(1-3i)=-3$.

Если $\mathrm{Im}(\alpha)=0$, то число α — действительное; если $\mathrm{Re}(\alpha)=0$, то число α имеет вид bi и называется *чисто мнимым*.

Заметим, что частным случаем формулы (8) является следующее правило умножения действительного числа на комплексное:

$$a(a_2+b_2i)=(a+0i)(a_2+b_2i)=aa_2+ab_2i, \quad (9)$$

т. е. *при умножении комплексного числа α на действительное число a оба числа* $\mathrm{Re}(\alpha)$ *и* $\mathrm{Im}(\alpha)$ *умножаются на a*. Например, $-3(2-7i)=-6+21i$, $2(-1+3i)=-2+6i$.

5. Вычитание и деление комплексных чисел

Пусть α и β — два комплексных числа. *Разностью* $\beta-\alpha$ называется комплексное число ξ, удовлетворяющее уравнению

$$\alpha+\xi=\beta. \quad (10)$$

Если $\alpha=a+bi$, $\beta=c+di$, то, полагая $\xi=x+yi$, вместо равенства (10) можем записать

$$(a+x)+(b+y)i=c+di \text{ или } a+x=c,\ b+y=d.$$

Следовательно, $x=c-a$, $y=d-b$. Итак, разность комплексных чисел задается формулой

$$(c+di)-(a+bi)=(c-a)+(d-b)i.$$

Прежде чем обращаться к делению комплексных чисел, заметим, что роль нуля в системе комплексных чисел играет число $0=0+0i$. Его основное свойство состоит в том, что $\alpha+0=\alpha$, каково бы то ни было комплексное число α.

Пусть α и β — два комплексных числа, причем $\alpha\neq0$. *Частным* β/α называется комплексное число μ, удовлетворяющее уравнению

$$\alpha\mu=\beta. \quad (11)$$

Полагая $\alpha=a+bi$, $\beta=c+di$, $\mu=x+yi$, запишем уравнение для μ в виде

$$(ax-by)+(ay+bx)i=c+di,$$

что равносильно двум равенствам

$$\begin{cases} ax-by=c, \\ bx+ay=d. \end{cases} \tag{12}$$

Получили систему из двух уравнений для неизвестных x и y. Определитель этой системы

$$\begin{vmatrix} a & -b \\ b & a \end{vmatrix} = a^2+b^2$$

отличен от нуля, поскольку условие $\alpha \neq 0$ означает, что хотя бы одно из чисел a, b не равно нулю. Значит, система (12) имеет единственное решение $(x_0; y_0)$, а уравнение (11) — единственное решение $\mu = x_0 + y_0 i$.

В следующем пункте будет указан более простой способ для нахождения частного двух комплексных чисел.

6. Сопряженные комплексные числа

Определение 5. Пусть $\alpha = a+bi$. Число $a-bi$, отличающееся от α лишь знаком коэффициента при мнимой части, называется *сопряженным* числу α и обозначается $\bar{\alpha}$.

Итак, по определению, $\bar{\alpha} = a-bi$.

Например, $\overline{2-3i} = 2+3i$, $\bar{i} = -i$.

Если α — действительное число, т. е. $\alpha = a+0i$, то $\bar{\alpha} = a-0i = \alpha$. Таким образом, любое действительное число равно своему сопряженному.

Для любого комплексного числа $\alpha = a+bi$ имеем

$$\alpha + \bar{\alpha} = (a+bi)+(a-bi) = 2a, \quad \alpha\bar{\alpha} = (a+bi)(a-bi) = a^2-(bi)^2 = a^2+b^2,$$

т. е. сумма и произведение сопряженных комплексных чисел являются действительными числами.

Важное свойство операции сопряжения выражается следующей леммой.

Лемма. *Пусть α и β — два комплексных числа. Тогда $\overline{\alpha * \beta} = \bar{\alpha} * \bar{\beta}$, где знак $*$ означает любую из операций: сложение, вычитание, умножение и деление.*

Например, если $*$ есть умножение, то $\bar{\alpha}\bar{\beta} = \overline{\alpha\beta}$ или, как говорят, *число сопряженное произведению, равно произведению сопряженных чисел;* аналогичное правило можно сформулировать для операций сложения, вычитания и деления.

□ Разберем случай, когда $*$ есть умножение; остальные операции рассматриваются аналогично.

Итак, необходимо доказать справедливость равенства $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$.

Пусть $\alpha = a+bi$, $\beta = c+di$. Тогда

$$\alpha\beta = (ac-bd)+(ad+bc)\,i$$

и, следовательно,

$$\overline{\alpha\beta} = (ac-bd)-(ad+bc)\,i.$$

С другой стороны,

$$\bar{\alpha}\bar{\beta}=(a-bi)(c-di)=(ac-(-b)(-d))+(a(-d)+(-b)c)i=$$
$$=(ac-bd)-(ad+bc)i.$$

Мы видим, что $\overline{\alpha\beta}=\bar{\alpha}\bar{\beta}$. ■

Операция сопряжения очень удобна при выполнении деления комплексных чисел. Для нахождения частного β/α ($\alpha\neq 0$) необходимо решить уравнение $\alpha\mu=\beta$ (с неизвестным μ). Умножив обе части этого уравнения на число $\bar{\alpha}$, имеем

$$(\bar{\alpha}\alpha)\mu=\bar{\alpha}\beta.$$

Число $\bar{\alpha}\alpha$, как мы знаем, является действительным. Теперь умножаем обе части последнего уравнения на число, обратное $\bar{\alpha}\alpha$ (правило умножения комплексного числа на действительное определяется формулой (9)) и в результате находим

$$\mu=\frac{1}{\bar{\alpha}\alpha}\bar{\alpha}\beta=\frac{\bar{\alpha}\beta}{\bar{\alpha}\alpha}.$$

Полученное правило обычно формулируют следующим образом: *для нахождения частного β/α следует умножить числитель и знаменатель на число, сопряженное знаменателю.*

Пример. Вычислить: а) $\frac{1-i}{1+i}$; б) $\frac{1}{i}$; в) $\frac{1-i}{3+4i}$.

○ а) $\frac{1-i}{1+i}=\frac{(1-i)(1-i)}{(1-i)(1+i)}=\frac{1}{2}(0-2i)=-i$;

б) $\frac{1}{i}=\frac{-i}{(-i)i}=\frac{-i}{1}=-i$;

в) $\frac{1-i}{3+4i}=\frac{(3-4i)(1-i)}{(3-4i)(3+4i)}=\frac{(3-4i)(1-i)}{3^2+4^2}=\frac{1}{25}(-1-7i)=-\frac{1}{25}-\frac{7}{25}i$. ●

7. Геометрическое изображение комплексных чисел. Комплексная плоскость

Если для изображения действительных чисел используются точки числовой прямой, то изображениями комплексных чисел служат точки координатной плоскости.

Введем на плоскости прямоугольную декартову систему координат с осями x и y. Тогда каждому комплексному числу $\alpha=a+bi$ будет отвечать точка с координатами a, b. Эту точку чаще всего обозначают той же буквой α, что и само число; вместо слов «число α» говорят «точка α».

При таком способе изображения комплексных чисел любому действительному числу, т. е. числу вида $a+0i$, отвечает точка $(a; 0)$, лежащая на оси x. Таким образом, приходим к уже известному способу изображения действительных чисел точками числовой прямой x. В связи с этим ось x называют *действительной осью*. Комплексным же числам вида $0+bi$ («чисто мнимым») отвечают точки $(0; b)$ оси y; по этой причине ось y называют *мнимой осью*. На рис. 208 указаны изображения некоторых комплексных чисел.

Плоскость, точки которой указанным выше образом интерпретируются как изображения комплексных чисел, называется *комплексной плоскостью*. Можно считать, что комплексная плоскость — это множество **C** комплексных чисел, рассматриваемых в двойном качестве: как числа и как точки координатной плоскости.

Наряду с изображением комплексных чисел точками плоскости применяется и другой способ изображения — с помощью векторов на плоскости. Числу $a+bi$ ставят в соответствие вектор с координатами a, b, причем, как правило, этот вектор считается отложенным от начала. Другими словами, числу $a+bi$ сопоставляется радиус-вектор точки $(a;b)$ (рис. 209). «Точечный» и «векторный» способы изображения комплексных чисел применяются одинаково часто.

Изображение комплексных чисел с помощью векторов имеет то преимущество, что оно хорошо «увязано» с операцией сложения комплексных чисел. Пусть числам $\alpha_1=a_1+b_1i$, $\alpha_2=a_2+b_2i$ соответствуют векторы $\overrightarrow{OA_1}=(a_1;\ b_1)$, $\overrightarrow{OA_2}=(a_2;\ b_2)$. Тогда числу $\alpha_1+\alpha_2=(a_1+a_2)+(b_1+b_2)i$ соответствует вектор

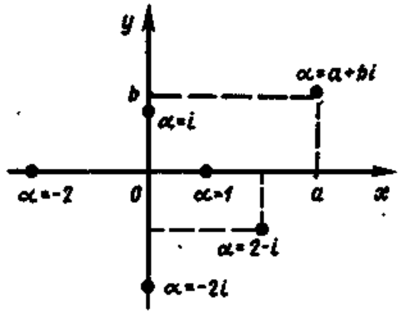


Рис. 208

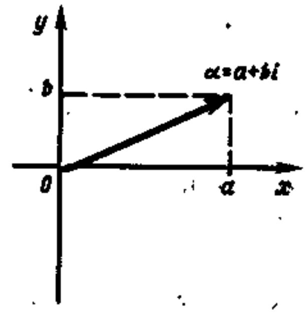


Рис. 209

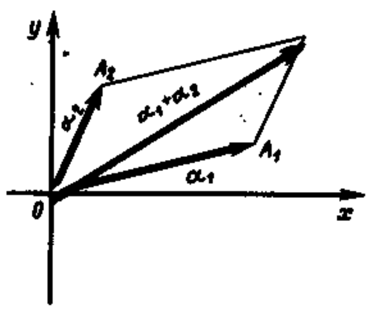


Рис. 210

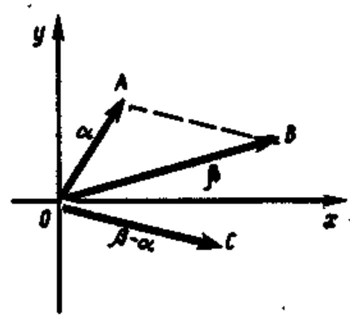


Рис. 211

с координатами a_1+a_2, b_1+b_2, т. е. вектор $\overrightarrow{OA_1}+\overrightarrow{OA_2}$. Таким образом, *сложение комплексных чисел геометрически сводится к сложению соответствующих им векторов*. Напомним, что сложение векторов осуществляется по правилу параллелограмма (рис. 210).

Поскольку сложение комплексных чисел сводится к сложению векторов, это же должно быть верно и по отношению к вычитанию: ведь вычитание (как для комплексных чисел, так и для векторов) есть операция, обратная сложению. Следовательно, если вектор $\overrightarrow{OA}$ изображает комплексное число α, а вектор $\overrightarrow{OB}$ — число β, то вектор $\overrightarrow{OB}-\overrightarrow{OA}=\overrightarrow{AB}$ является изображением числа $\beta-\alpha$ (рис. 211). Разумеется, чтобы получить точку, изображающую число $\beta-\alpha$, этот вектор нужно отложить от начала (точка C на рис. 211).

§ 61. Тригонометрическая форма комплексного числа и ее применения

1. Модуль и аргумент комплексного числа

Пусть на плоскости задана прямоугольная декартова система координат. В этой системе произвольное комплексное число $\alpha = a + bi$ изображается точкой A, имеющей координаты a, b. Положим $\alpha \neq 0$, т. е. будем считать, что точка A отлична от начала O системы координат. Тогда для этой точки определены полярные координаты r, φ, где $r = OA$, а φ есть угол между положительным направлением оси x и направлением вектора $\overrightarrow{OA}$ (рис. 212). Очевидно,

$$r = \sqrt{a^2 + b^2} \tag{1}$$

а также

$$\cos\varphi = \frac{a}{r}, \ \sin\varphi = \frac{b}{r}. \tag{2}$$

Так как точка A отлична от начала, то $r > 0$; что касается числа φ, то оно определено с точностью до слагаемого вида $2\pi k$, где k — любое целое число.

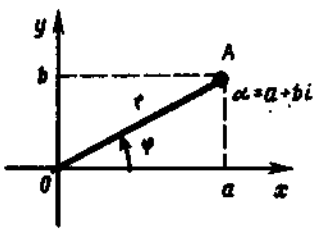


Рис. 212

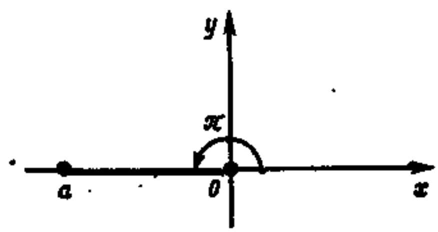


Рис. 213

Определение 1. Пусть $\alpha = a + bi$ — отличное от нуля комплексное число. Действительное число r, определенное равенством (1), называется *модулем* комплексного числа α, а число φ, определенное (с указанной выше точностью) равенствами (2), — *аргументом* числа α.

Условимся распространять понятие модуля на случай $\alpha = 0$; в этом случае будем считать число r равным нулю, а φ — каким угодно действительным числом.

Модуль комплексного числа α обозначается символом $|\alpha|$, а аргумент — символом $\operatorname{Arg}\alpha$. Значение аргумента, заключенное в промежутке $(-\pi, \pi]$, обозначают обычно $\arg\alpha$: $-\pi < \arg\alpha \leqslant \pi$.

Если α — действительное число, т. е. $\alpha = a + 0i$, то $|\alpha| = \sqrt{a^2 + 0^2}$ и, следовательно, $|\alpha| = |a|$. Таким образом, понятие модуля комплексного числа является обобщением понятия модуля действительного числа.

Число $\arg\alpha$ также можно считать обобщением известного понятия, а именно — понятия знака действительного числа. В самом деле, на действительной оси x из начала O выходят два луча, которые мы отмечаем знаками + и —. На комплексной же плоскости из начала координат можно провести не два, а бесчисленное множество лучей; чтобы различать эти лучи, мы приписываем каждому из них определенное значение полярного угла φ.

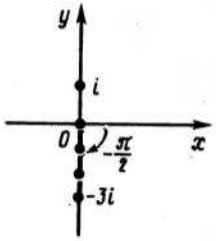


Рис. 214

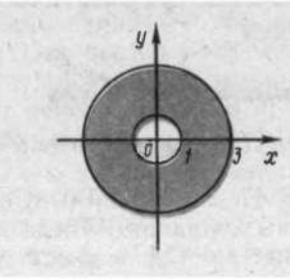


Рис. 215

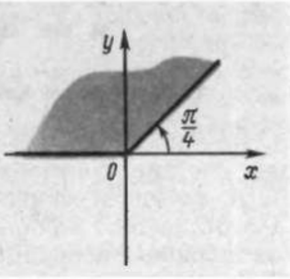


Рис. 216

Нетрудно видеть, что для действительного числа $a+0i$, где $a\neq 0$, аргумент всегда равен 0 или π (разумеется, с точностью до слагаемого $2\pi k$): 0 в случае, когда $a>0$, и π в случае, когда $a<0$ (рис. 213).

Формулы (1) и (2) позволяют для заданного комплексного числа $a+bi$ находить модуль r и аргумент φ. Обратно, если заданы два действительных числа r и φ, причем $r\geqslant 0$, то существует комплексное число $a+bi$, для которого r и φ являются соответственно модулем и аргументом. Число $a+bi$ находится с помощью равенств

$$a=r\cos\varphi,\ \ b=r\sin\varphi. \qquad (3)$$

Примеры. 1. Найти модуль и аргумент числа:

$$\text{а) } \alpha=-3i;\ \text{б) } \alpha=-1+i;\ \text{в) } \alpha=\frac{1}{2}-\frac{\sqrt{3}}{2}i.$$

○ а) Из положения точки $\alpha=-3i$ на комплексной плоскости непосредственно ясно (рис. 214), что $|\alpha|=3$, $\arg\alpha=-\pi/2$.

б) Как и в предыдущем случае, непосредственно находим $|\alpha|=\sqrt{2}$, $\arg\alpha=$ $=-3\pi/4$.

в) Имеем $|\alpha|=\sqrt{\frac{1}{4}+\frac{3}{4}}=1$; далее по формулам (2) находим $\cos\varphi=1/2$, $\sin\varphi=-\sqrt{3}/2$, т. е. $\varphi=-\frac{\pi}{3}+2\pi k$ $(k\in\mathbf{Z})$. ●

2. Описать геометрически множество M на комплексной плоскости, состоящее из всех точек α, для которых выполняется условие: а) $|\alpha|=2$; б) $1\leqslant|\alpha|\leqslant 3$; в) $\arg\alpha=$ $=\pi/4$; г) $\pi/4\leqslant\arg\alpha\leqslant\pi$.

○ а) Множество M есть окружность радиуса 2 с центром в начале координат.

б) Множество M есть кольцо между двумя концентрическими окружностями, имеющими центр в начале координат (рис. 215); радиусы окружностей соответственно равны 1 и 3.

в) Множество M есть луч, выходящий из начала координат и образующий угол $\pi/4$ с положительным направлением оси x.

г) Множество M выделено на рис. 216 цветом. ●

2. Модуль суммы и разности двух комплексных чисел

Докажем, что справедливы следующие два неравенства:

$$|\alpha+\beta|\leqslant|\alpha|+|\beta| \qquad (4)$$

(модуль суммы двух комплексных чисел не превосходит суммы их модулей);

$$|\alpha-\beta|\geqslant|\alpha|-|\beta| \qquad (5)$$

(модуль разности двух комплексных чисел не меньше разности модулей этих чисел).

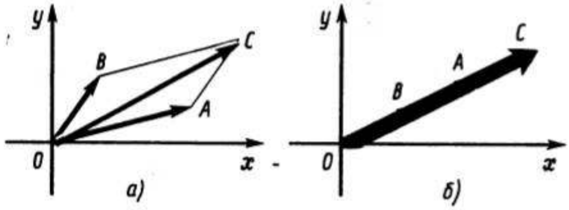


Рис. 217

□ Будем исходить из геометрического смысла сложения комплексных чисел. Пусть числу α отвечает вектор $\overrightarrow{OA}$, числу β — вектор $\overrightarrow{OB}$, тогда числу $\alpha+\beta$ отвечает вектор $\overrightarrow{OA}+\overrightarrow{OB}$ (рис. 217, *а*). Из треугольника OAC имеем

$$OC \leqslant OA + AC, \tag{6}$$

или, учитывая, что $AC = OB$,

$$OC \leqslant OA + OB.$$

Это равносильно неравенству (4). Отметим, что знак равенства в (6), а значит, и в (4), достигается лишь в том случае, когда векторы $\overrightarrow{OA}$ и $\overrightarrow{OB}$ имеют одинаковые направления; тогда треугольник OAC «сплющивается» в отрезок OC и точка A лежит на этом отрезке (рис. 217, *б*).

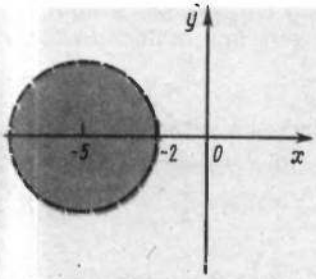


Рис. 218

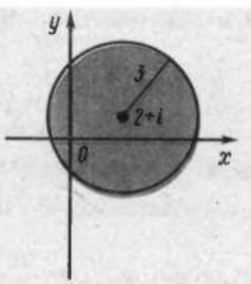


Рис. 219

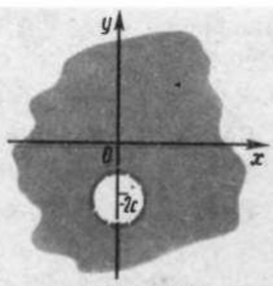


Рис. 220

Аналогично доказывается справедливость неравенства (5). Разности $\alpha-\beta$ соответствует вектор $\overrightarrow{OA}-\overrightarrow{OB}=\overrightarrow{BA}$, и неравенство (5) выражает тот факт, что в треугольнике OAB сторона BA не меньше, чем разность двух остальных сторон. ■

Используя метод математической индукции, можно распространить неравенство (4) на любое число n слагаемых:

$$|\alpha_1+\alpha_2+\dots+\alpha_n| \leqslant |\alpha_1|+|\alpha_2|+\dots+|\alpha_n|,$$

каковы бы ни были комплексные числа $\alpha_1, \alpha_2, \dots, \alpha_n$.

Установим геометрический смысл модуля разности двух комплексных чисел α и β. Выше мы отмечали, что разности $\alpha-\beta$ соответствует вектор $\overrightarrow{BA}$ (см. рис. 217, *а*). Так как модуль комплексного числа равен длине вектора, изображающего это число, то можем записать $|\alpha-\beta| = BA$. Иначе говоря, ***модуль разности двух комплексных чисел равен расстоянию между точками, изображающими эти числа.***

Это утверждение будет неоднократно использоваться в дальнейшем.

Пример. Описать геометрически множество M, состоящее из точек α комплексной плоскости, для которых выполняется условие: а) $|\alpha+5|<3$; б) $|\alpha-(2+i)|<3$; в) $|\alpha+2i|>1$.

○ а) Перепишем данное неравенство в виде $|\alpha-(-5)|<3$. Это означает, что расстояние от точки α до точки -5 должно быть меньше 3. Следовательно, точка α находится внутри круга радиуса 3 с центром в точке -5. Множество M есть внутренность указанного круга (рис. 218).

б), в) Решения изображены соответственно на рис. 219 и 220. ●

3. Тригонометрическая форма комплексного числа

Пусть комплексное число $\alpha=a+bi$ имеет модуль r и аргумент φ. Используя формулы (3), получим

$$\alpha=r(\cos\varphi+i\sin\varphi). \tag{7}$$

Определение 2. Представление комплексного числа α в виде (7), где r и φ — действительные числа, причем $r\geqslant 0$, называется *тригонометрической формой* числа α.

В случае, когда α — действительное число, т. е. $\alpha=a+0i$, тригонометрическая форма имеет вид

$$\alpha=|a|(\cos\varphi+i\sin\varphi),$$

где $\varphi=0$ при $a>0$ и $\varphi=\pi$ при $a<0$; при $a=0$ число φ может быть каким угодно.

В связи с определением тригонометрической формы напомним снова формулы для нахождения модуля и аргумента комплексного числа:

$$r=\sqrt{a^2+b^2},\ \cos\varphi=\frac{a}{r},\ \sin\varphi=\frac{b}{r}.$$

Пример. Представить в тригонометрической форме комплексное число: а) i; б) $-3i$; в) $1+i$.

○ а) Имеем $r=1$, $\varphi=\pi/2$; следовательно, $i=1\left(\cos\frac{\pi}{2}+i\sin\frac{\pi}{2}\right)$.

б) Здесь $r=3$, $\varphi=\frac{3\pi}{2}$, откуда $-3i=3\left(\cos\frac{3\pi}{2}+i\sin\frac{3\pi}{2}\right)$.

в) Здесь $r=\sqrt{2}$, $\varphi=\frac{\pi}{4}$, значит, $1+i=\sqrt{2}\left(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}\right)$. ●

4. Умножение и деление комплексных чисел в тригонометрической форме

Тригонометрическую форму комплексного числа удобно использовать при выполнении операций умножения и деления комплексных чисел.

Пусть

$$\alpha_1=r_1(\cos\varphi_1+i\sin\varphi_1),\ \alpha_2=r_2(\cos\varphi_2+i\sin\varphi_2) \tag{8}$$

— комплексные числа, заданные в тригонометрической форме. Имеем

$$\alpha_1\alpha_2=r_1r_2(\cos\varphi_1+i\sin\varphi_1)(\cos\varphi_2+i\sin\varphi_2).$$

Выполняя умножение чисел, заключенных в скобки, получим

$$\alpha_1\alpha_2=r_1r_2((\cos\varphi_1\cos\varphi_1-\sin\varphi_1\sin\varphi_2)+i(\cos\varphi_1\sin\varphi_2+\sin\varphi_1\cos\varphi_2)),$$

или

$$\alpha_1\alpha_2 = r_1r_2(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)). \tag{9}$$

Формула (9) дает правило умножения комплексных чисел в тригонометрической форме: *для умножения α_1 на α_2 модули этих чисел следует перемножить, а аргументы сложить.*

То же самое можно выразить по-другому. Из равенства (9) следует

$$|\alpha_1\alpha_2| = r_1r_2 = |\alpha_1|\,|\alpha_2|, \tag{10}$$

т. е. *модуль произведения равен произведению модулей,* а

$$\operatorname{Arg}\alpha_1\alpha_2 = \varphi_1+\varphi_2 = \operatorname{Arg}\alpha_1+\operatorname{Arg}\alpha_2, \tag{11}$$

т. е. *аргумент произведения равен сумме аргументов сомножителей.*

Впрочем, следует заметить, что равенство (11) имеет место лишь с точностью до прибавления к любому из чисел $\operatorname{Arg}\alpha_1$, $\operatorname{Agr}\alpha_2$, $\operatorname{Arg}\alpha_1\alpha_2$ слагаемого вида $2\pi k$, где k — целое.

Используя метод математической индукции, можно распространить формулу (9) на любое число n сомножителей:

$$\alpha_1\alpha_2\ldots\alpha_n = r_1r_2\ldots r_n(\cos(\varphi_1+\varphi_2+\cdots+\varphi_n)+i\sin(\varphi_1+\varphi_2+\cdots\varphi_n)), \tag{12}$$

где r_i, φ_i $(i=1, 2, \ldots, n)$ — модуль и аргумент числа α_i.

Обратимся теперь к делению комплексных чисел. Пусть α_1 и α_2 заданы в тригонометрической форме (8), причем $r_1\neq 0$ (т. е. $\alpha_1\neq 0$). Рассмотрим число

$$\xi = \frac{r_2}{r_1}(\cos(\varphi_2-\varphi_1)+i\sin(\varphi_2-\varphi_1)).$$

Согласно правилу умножения комплексных чисел в тригонометрической форме, имеем

$$\alpha_1\xi = r_1\frac{r_2}{r_1}(\cos(\varphi_1+\varphi_2-\varphi_1)+i\sin(\varphi_1+\varphi_2-\varphi_1)) = r_2(\cos\varphi_2+i\sin\varphi_2),$$

т. е. $\alpha_1\xi=\alpha_2$ или $\xi=\alpha_2/\alpha_1$. Итак,

$$\frac{\alpha_2}{\alpha_1} = \frac{r_2}{r_1}(\cos(\varphi_2-\varphi_1)+i\sin(\varphi_2-\varphi_1)). \tag{13}$$

Отсюда получаем следующее правило: *для нахождения частного α_2/α_1 следует модуль числа α_2 разделить на модуль числа α_1, а из аргумента числа α_2 вычесть аргумент числа α_1*, или, что то же самое:

$$\left|\frac{\alpha_2}{\alpha_1}\right| = \frac{r_2}{r_1} = \frac{|\alpha_2|}{|\alpha_1|},$$

т. е. *модуль частного равен отношению модулей,* а

$$\operatorname{Arg}\frac{\alpha_2}{\alpha_1} = \varphi_2-\varphi_1 = \operatorname{Arg}\alpha_2-\operatorname{Arg}\alpha_1,$$

т. е. *аргумент частного равен разности аргументов числителя и знаменателя.*

5. Геометрический смысл умножения на комплексное число

Пусть α и ξ — произвольные комплексные числа. Выясним, как связаны геометрически числа ξ и $\alpha\xi$; иначе говоря, каков геометрический «механизм» перехода от ξ к $\alpha\xi$.

Положим $\xi = x + yi$. Если $\alpha = r \in \mathbb{R}$, то $\xi = r(x+yi) = rx + ryi$. Следовательно, если M — точка, соответствующая числу ξ, а M' — точка, соответствующая $\alpha\xi$, то $\overrightarrow{OM} = (x;\ y)$, $\overrightarrow{OM'} = (rx;\ ry)$. Отсюда имеем $\overrightarrow{OM'} = r\overrightarrow{OM}$. При $r > 0$ это означает, что *точка M' получается из M в результате гомотетии с центром O и коэффициентом гомотетии r* (рис. 221).

В другом частном случае, когда $\alpha = \cos\varphi + i\sin\varphi$, т. е. когда $|\alpha| = 1$, имеем $|\alpha\xi| = |\xi|$, $\operatorname{Arg}\alpha\xi = \operatorname{Arg}\xi + \varphi$. Это означает, что *$M'$ получается из M поворотом плоскости вокруг O на угол* φ (рис. 222).

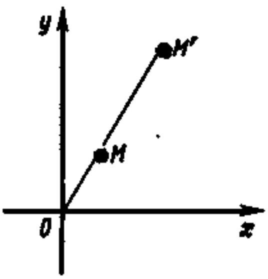


Рис. 221

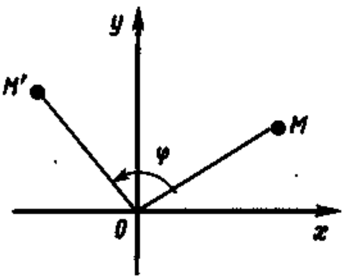


Рис. 222

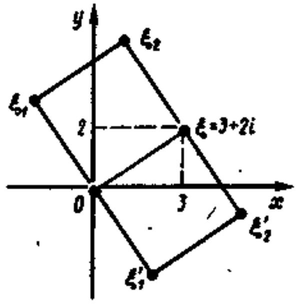


Рис. 223

В общем случае можем записать $\alpha = r(\cos\varphi + i\sin\varphi)$, где $r > 0$, откуда следует

$$|\alpha\xi| = r|\xi|,\quad \operatorname{Arg}\alpha\xi = \operatorname{Arg}\xi + \varphi,$$

т. е. *точка $\alpha\xi$ получается из ξ в результате последовательного применения двух преобразований: поворота вокруг O на угол φ и гомотетии центром O и коэффициентом r* (порядок, в котором производятся эти преобразования, безразличен).

Пример. Одной из вершин квадрата является начало координат, другой — точка $\xi = 3 + 2i$. Найти две остальные вершины.

○ Очевидно, задача имеет два решения: в одном из них неизвестные вершины — точки ξ_1 и ξ_2 (рис. 223), в другом — точки ξ_1' и ξ_2'.

Точка ξ_1 получается из ξ поворотом вокруг начала на угол $\pi/2$ в положительном направлении. Следовательно,

$$\xi_1 = \left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right)(3+2i) = i(3+2i) = -2+3i.$$

Теперь число ξ_2 можно найти как сумму чисел ξ и ξ_1, т. е.

$$\xi_2 = (3+2i) + (-2+3i) = 1+5i.$$

Во втором случае имеем

$$\xi_1' = \left(\cos\left(-\frac{\pi}{2}\right) + i\sin\left(-\frac{\pi}{2}\right)\right)(3+2i) = -i(3+2i) = 2-3i,$$

$$\xi_2' = (3+2i) + (2-3i) = 5-i. \ ●$$

6. Формула Муавра

Пусть комплексное число α представлено в тригонометрической форме: $\alpha = r(\cos\varphi + i\sin\varphi)$. Полагая в формуле (12) $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \alpha$, получим

$$\alpha^n = r^n(\cos n\varphi + i\sin n\varphi),$$

или

$$[r(\cos\varphi + i\sin\varphi)]^n = r^n(\cos n\varphi + i\sin n\varphi). \qquad (14)$$

Эта формула называется *формулой Муавра**.

Пример. Найти $(1+i)^{50}$.

○ В данном случае $\alpha = 1+i$. Представление числа α в тригонометрической форме имеет вид $\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$. Отсюда следует

$$\alpha^{50} = (\sqrt{2})^{50}\left(\cos 50\frac{\pi}{4} + i\sin 50\frac{\pi}{4}\right).$$

Так как $50\frac{\pi}{4} = 12\pi + \frac{\pi}{2}$, то $\cos\left(50\frac{\pi}{4}\right) = \cos\frac{\pi}{2} = 0$, $\sin\left(50\frac{\pi}{4}\right) = \sin\frac{\pi}{2} = 1$. Следовательно,

$$(1+i)^{50} = (\sqrt{2})^{50}(0 + 1\cdot i) = 2^{25}i. \ \bullet$$

Заметим, что если бы для нахождения числа $(1+i)^{50}$ мы решили воспользоваться формулой бинома Ньютона, то вычисления были бы намного более громоздкими.

7. Извлечение квадратного корня из комплексного числа

Квадратным корнем из комплексного числа α называется любое комплексное число ξ такое, что $\xi^2 = \alpha$. Как обычно, квадратный корень из α обозначается $\sqrt{\alpha}$. Ясно, что $\sqrt{0} = 0$.

Пусть $\alpha \neq 0$. Положим $\alpha = a + bi$; искомое число $\xi = \sqrt{\alpha}$ запишем в виде $\xi = x + yi$. Имеем

$$(x+yi)^2 = a + bi, \text{ или } (x^2 - y^2) + 2xyi = a + bi.$$

Согласно определению равенства комплексных чисел отсюда следует

$$\begin{cases} x^2 - y^2 = a, \\ 2xy = b. \end{cases} \qquad (15)$$

Для двух неизвестных действительных чисел x и y получили систему из двух уравнений. Можно показать (сделайте это самостоятельно), что эта система при $\alpha \neq 0$ (т. е. когда $a^2 + b^2 \neq 0$) имеет два различных решения, причем если одно из них есть $x = x_0$, $y = y_0$, то второе есть $x = -x_0$, $y = -y_0$. Следовательно, $\sqrt{\alpha}$ при $\alpha \neq 0$ имеет два значения: $x_0 + y_0 i$ и $-(x_0 + y_0 i)$, отличающиеся друг от друга множителем -1.

Пример. Вычислить $\sqrt{i}$.

○ Пусть $\sqrt{i} = x + yi$; тогда $(x+yi)^2 = i$, т. е. $(x^2 - y^2) + 2xyi = i$. Отсюда имеем

$$\begin{cases} x^2 - y^2 = 0, \\ 2xy = 1. \end{cases}$$

---

* А. Муавр (1667—1754) — французский математик.

Из первого уравнения следует $x = \pm y$, после чего из второго уравнения находим $x=y$, т. е. $2x^2 = 1$. Это уравнение имеет решения $x_1 = 1/\sqrt{2}$ и $x_2 = -1/\sqrt{2}$. Соответственно получим $y_1 = 1/\sqrt{2}$ и $y_2 = -1/\sqrt{2}$. Итак, $\sqrt{i}$ имеет два значения: $\frac{1}{\sqrt{2}}(1+i)$ и $-\frac{1}{\sqrt{2}}(1+i)$. ●

Возможность извлечь квадратный корень из произвольного комплексного числа α означает, что *в области комплексных чисел разрешимо любое квадратное уравнение*. Действительно, уравнение $\xi^2+p\xi+q=0$ (ξ — неизвестное) легко преобразуется к виду $\left(\xi+\frac{p}{2}\right)^2=\frac{p^2}{4}-q$. Отсюда находим $\xi+\frac{p}{2}=$ $=\sqrt{\frac{p^2}{4}-q}$ или $\xi=-\frac{p}{2}-\sqrt{\frac{p^2}{4}-q}$, где в правой части подразумевается любое из двух значений квадратного корня.

Например, уравнение $\xi^2+2\xi+3=0$ ($p=2$, $q=3$) имеет решение $\xi=-1+$ $+\sqrt{-2}$. Двумя различными значениями для $\sqrt{-2}$ являются $\sqrt{2}i$ и $-\sqrt{2}i$ (для $\sqrt{2}$ имеется в виду арифметическое значение). Таким образом, данное уравнение имеет корни $\xi_1=-1+\sqrt{2}i$, $\xi_2=-1-\sqrt{2}i$.

8. Извлечение корня любой степени из комплексного числа

Пусть требуется найти все значения корня n-й степени (n — любое натуральное) из комплексного числа α, не равного нулю (ясно, что $\sqrt[n]{0}=0$). Попытка решить эту задачу тем же способом, каким в п. 7 были найдены значения $\sqrt{\alpha}$, не приводит к успеху, так как вместо системы (15) получается система (из двух уравнений с двумя неизвестными) n-й степени; при $n>2$ решение такой системы — дело чрезвычайно трудное. Будем решать задачу нахождения $\sqrt[n]{\alpha}$, используя тригонометрическую форму комплексного числа.

Пусть α задано в тригонометрической форме: $\alpha=r(\cos\varphi+i\sin\varphi)$. Число $\xi=\sqrt[n]{\alpha}$ будем искать также в тригонометрической форме: $\xi=\rho(\cos\psi+i\sin\psi)$, где $\rho\geqslant 0$. Из условия $\xi^n=\alpha$, используя формулу Муавра (14), находим

$$\rho^n(\cos n\psi+i\sin n\psi)=r(\cos\varphi+i\sin\varphi).$$

Отсюда следует, что

$$\rho^n=r,\ n\psi=\varphi+2\pi k\quad (k\in\mathbf{Z}).$$

Из первого равенства находим, что $\rho=\sqrt[n]{r}$ (арифметическое значение), а из второго — что $\psi=\frac{\varphi+2\pi k}{n}$.

Таким образом, получаем следующее представление:

$$\sqrt[n]{\alpha}=\sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n}+i\sin\frac{\varphi+2\pi k}{n}\right) \tag{16}$$

(k — любое целое).

Может показаться, что эта формула дает для $\sqrt[n]{\alpha}$ бесчисленное множество значений (поскольку k — любое целое). В действительности же для $\sqrt[n]{\alpha}$ имеется ровно n различных значений и, чтобы получить эти значения, достаточно в правой части формулы (16) положить n равным 0, 1, 2, ..., $n-1$.

В самом деле, точки

$$\xi_k=\sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n}+i\sin\frac{\varphi+2\pi k}{n}\right)\qquad (k\in\mathbf{Z})$$

располагаются на окружности радиуса $\sqrt[n]{r}$ с центром в начале координат.

При этом если значение k изменяется на 1, то угол $\frac{\varphi+2\pi k}{n}$ изменяется на величину $\frac{2\pi}{n}$, т. е. на $\frac{1}{n}$ часть полного угла 2π. Это означает, что точки ξ_0, ξ_1, ξ_2, ..., ξ_{n-1} делят указанную окружность на n равных частей (на рис. 224 $n=6$); в частности, эти точки различны. Значения же k, отличные от 0, 1, 2, ..., $n-1$, не дают новых точек: например, $\xi_n=\xi_0$, $\xi_{n+1}=\xi_1$ и т. д.

Итак, *корень n-й степени из не равного нулю комплексного числа α имеет n различных значений, определяемых формулой* (16) *при* $k=0, 1, \ldots, n-1$, *где r и φ — модуль и аргумент числа α. Точки, соответствующие этим значениям,*

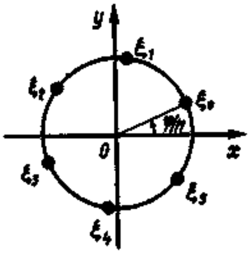


Рис. 224

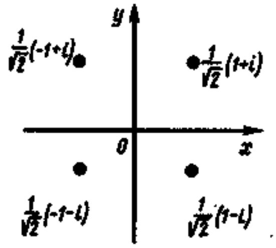


Рис. 225

расположены на окружности радиуса $\sqrt[n]{r}$ с центром в начале координат и делят окружность на n равных частей.

Пример. Найти все значения $\sqrt[4]{-1}$.

○ Имеем $-1=1(\cos\pi+i\sin\pi)$, так что $r=1$, $\varphi=\pi$. По формуле (16) находим

$$\sqrt[4]{-1}=\sqrt[4]{1}\left(\cos\frac{\pi+2\pi k}{4}+i\sin\frac{\pi+2\pi k}{4}\right).$$

Полагая k последовательно равным 0, 1, 2, 3, найдем все четыре значения $\sqrt[4]{-1}$. Этими значениями являются:

при $k=0$ $\quad \cos\frac{\pi}{4}+i\sin\frac{\pi}{4}=\frac{1}{\sqrt{2}}(1+i)$;

при $k=1$ $\quad \cos\frac{\pi+2\pi}{4}+i\sin\frac{\pi+2\pi}{4}=\frac{1}{\sqrt{2}}(-1+i)$;

при $k=2$ $\quad \cos\frac{\pi+2\pi\cdot 2}{4}+i\sin\frac{\pi+2\pi\cdot 2}{4}=\frac{1}{\sqrt{2}}(-1-i)$;

при $k=3$ $\quad \cos\frac{\pi+2\pi\cdot 3}{4}+i\sin\frac{\pi+2\pi\cdot 3}{4}=\frac{1}{\sqrt{2}}(1-i)$.

Точки, соответствующие полученным числам, изображены на рис. 225. ●

§ 62. Многочлены в комплексной области

Начнем с замечания по поводу дальнейших обозначений. В предыдущих параграфах комплексные числа записывались в основном с помощью греческих букв α, β, ξ и т. д. В дальнейшей части настоящей главы мы отступаем от такой записи и обозначаем комплексные числа как греческими, так и латинскими буквами. Переменные величины, принимающие комплексные значения, будем как правило, обозначать z и w.

1. Теорема о существовании корня

Напомним, что основной причиной введения комплексных чисел послужил тот факт, что не всякое квадратное уравнение можно решить, оставаясь в области действительных чисел. Комплексные числа полностью снимают эту трудность. Более того, как было показано в п. 7 § 61, не только квадратные уравнения с действительными коэффициентами, но и квадратные уравнения с любыми комплексными коэффициентами разрешимы в области комплексных чисел.

Естественно возникает вопрос о разрешимости уравнений, степень которых выше, чем 2. Оказывается, что для решения таких уравнений также достаточно комплексных чисел. Справедливо следующее утверждение.

Теорема 1 (основная теорема алгебры комплексных чисел). *Уравнение*

$$z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_{n-1} z + a_n = 0,$$

где n — любое натуральное число и $a_1, a_2, \ldots, a_n$ — какие угодно комплексные числа, имеет хотя бы один комплексный корень.

Мы примем эту теорему без доказательства.

2. Разложение многочлена с комплексными коэффициентами на множители первой и второй степени

Пусть

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n \quad (1)$$

— многочлен степени n с комплексными коэффициентами $a_1, \ldots, a_n$ и старшим коэффициентом, равным 1. Согласно теореме 1, $f(z)$ имеет некоторый комплексный корень α_1. Как известно из курса алгебры, отсюда следует, что $f(z)$ делится без остатка на $z - \alpha_1$, т. е.

$$f(z) = (z - \alpha_1) f_1(z),$$

где $f_1(z)$ — многочлен степени $n-1$ (со старшим коэффициентом 1). Применив такое же рассуждение к многочлену $f_1(z)$, получим $f_1(z) = (z - \alpha_2) f_2(z)$ и, следовательно,

$$f(z) = (z - \alpha_1)(z - \alpha_2) f_2(z),$$

где $f_2(z)$ — многочлен степени $n-2$ (со старшим коэффициентом 1). Продолжая рассуждать так же, получим

$$f(z) = (z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_n) f_n(z),$$

где $f_n(z)$ — многочлен степени 0 со старшим коэффициентом 1, т. е. $f_n(z) = 1$. Итак,

$$f(z) = (z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_n). \quad (2)$$

Мы доказали следующую теорему.

Теорема 2. *Многочлен $f(z)$ степени n с комплексными коэффициентами и со старшим коэффициентом, равным 1, разлагается в произведение n множителей вида $z - \alpha$.*

В случае, когда старший коэффициент многочлена $f(z)$ равен какому-то числу a (необязательно 1), имеем

$$f(z)=a(z-\alpha_1)(z-\alpha_2)\dots(z-\alpha_n).$$

Теорема 3. *Представление многочлена $f(z)$ в виде произведения* (2) *единственно с точностью до перестановки сомножителей.*

□ Пусть имеет место равенство

$$(z-\alpha_1)(z-\alpha_2)\dots(z-\alpha_n)=(z-\beta_1)(z-\beta_2)\dots(z-\beta_n). \qquad (3)$$

При $z=\alpha_1$ левая часть обращается в нуль; следовательно, равна нулю и правая часть. Это означает, что какое-то из чисел $\beta_1, \beta_2, \dots \beta_n$ равно α_1. Пусть, например, $\beta_1=\alpha_1$. Сокращая обе части равенства (3) на $z-\alpha_1$, получим

$$(z-\alpha_2)\dots(z-\alpha_n)=(z-\beta_2)\dots(z-\beta_n).$$

Далее точно так же докажем, что α_2 равно одному из чисел $\beta_2, \dots, \beta_n$ и т. д. Продолжая это рассуждение, заключаем, что набор чисел $\alpha_1, \alpha_2, \dots, \alpha_n$ лишь порядком отличается от набора чисел $\beta_1, \beta_2, \dots, \beta_n$.

Каждое из чисел $\alpha_1, \alpha_2, \dots, \alpha_n$ в равенстве (2) является, очевидно, корнем многочлена $f(z)$. Обратно, любой корень $f(z)$ совпадает с одним из чисел $\alpha_1, \alpha_2, \dots, \alpha_n$ в равенстве (2). Действительно, если α — корень $f(z)$, то при $z=\alpha$ из равенства (2) следует

$$0=(\alpha-\alpha_1)(\alpha-\alpha_2)\dots(\alpha-\alpha_n),$$

т. е. число α равно одному из чисел $\alpha_1, \alpha_2, \dots, \alpha_n$ ■

3. Кратность корня. Сумма кратностей всех корней многочлена

В разложении (2) некоторые из множителей могут оказаться одинаковыми. Произведение одинаковых множителей $z-\alpha_i$ можно записать в виде степени $(z-\alpha_i)^{k_i}$. В результате приходим к равенству вида

$$f(z)=(z-\beta_1)^{k_1}(z-\beta_2)^{k_2}\dots(z-\beta_s)^{k_s}, \qquad (4)$$

где числа $\beta_1, \beta_2, \dots, \beta_s$ попарно различны. При этом, разумеется, выполняется равенство

$$k_1+k_2+\dots+k_s=n, \qquad (5)$$

поскольку число всех множителей в правой части соотношения (4) равно n.

Определение. Корень β многочлена $f(z)$ называется *корнем кратности* k, если $f(z)$ делится на $(z-\beta)^k$, но не делится на $(z-\beta)^{k+1}$.

Согласно этому определению, равенство (5) означает, что, например, кратность корня β_1 равна k_1. Действительно, как показывает это равенство, $f(z)$ делится на $(z-\beta_1)^{k_1}$, а многочлен

$$\frac{f(z)}{(z-\beta_1)^{k_1}}=(z-\beta_2)^{k_2}\dots(z-\beta_s)^{k_s}$$

уже не делится на $z-\beta_1$ (в противном случае этот многочлен при $z=\beta_1$ обращался бы в нуль, а это противоречит тому, что $\beta_1\neq\beta_2, \dots, \beta_1\neq\beta_s$).

Равенство (5) может быть теперь «прочитано» следующим образом: сумма кратностей всех корней многочлена $f(z)$ равна степени многочлена. Обычно это утверждение формулируют в виде следующей теоремы.

Теорема 4. *Число комплексных корней многочлена $f(z)$ равно степени этого многочлена, если каждый корень считать столько раз, какова его кратность.*

Например, многочлен $(z-1)^3(z+1)=z^4-2z^3+2z-1$, степень которого равна 4, имеет четыре корня 1, 1, 1, -1 (с учетом кратностей); из них различными являются только два: 1 и -1.

§ 63. Многочлены с действительными коэффициентами

1. Попарная сопряженность комплексных корней

Теорема 1. *Пусть $f(z)$ — многочлен с действительными коэффициентами. Если комплексное число α является корнем $f(z)$, то сопряженное число $\bar{\alpha}$ также является корнем $f(z)$.*

□ Пусть $f(z)=a_0z^n+a_1z^{n-1}+\dots+a_{n-1}z+a_n$. По условию имеем $f(\alpha)=0$, т. е. $a_0\alpha^n+a_1\alpha^{n-1}+\dots+a_{n-1}\alpha+a_n=0$. Найдем число $f(\bar{\alpha})$:

$$f(\bar{\alpha})=a_0\bar{\alpha}^n+a_1\bar{\alpha}^{n-1}+\dots+a_{n-1}\bar{\alpha}+a_n.$$

Так как коэффициенты $a_0, a_1, \dots, a_n$ — действительные числа, то $a_0=\bar{a}_0$, $a_1=\bar{a}_1, \dots, a_n=\bar{a}_n$. Отсюда

$$f(\bar{\alpha})=\bar{a}_0\bar{\alpha}^n+\bar{a}_1\bar{\alpha}^{n-1}+\dots+\bar{a}_{n-1}\bar{\alpha}+\bar{a}_n.$$

Воспользовавшись свойствами сопряжения комплексных чисел (см. п. 6 § 60), можем записать

$$f(\bar{\alpha})=\overline{a_0\alpha^n}+\overline{a_1\alpha^{n-1}}+\dots+\overline{a_{n-1}\alpha}+\bar{a}_n=$$

$$=\overline{a_0\alpha^n+a_1\alpha^{n-1}+\dots+a_{n-1}\alpha+a_n}=\overline{f(\alpha)}=\bar{0}=0.$$

Полученное равенство $f(\bar{\alpha})=0$ означает, что число $\bar{\alpha}$ является корнем многочлена $f(z)$. ■

Пример. Найти многочлен наименьшей степени с действительными коэффициентами, имеющий корни $1-i$ и 2 (а также, быть может, еще и другие корни).

○ Пусть $f(z)$ — искомый многочлен. Тогда среди его корней обязательно должны быть числа $1-i$, $1+i$, 2. Это означает, что в разложение многочлена $f(z)$ на множители первой степени должно входить произведение $(z-(1-i))(z-(1+i))(z-2)$. Но

$$(z-(1-i))(z-(1+i))=((z-1)+i)((z-1)-i)=(z-1)^2+1$$

есть многочлен с действительными коэффициентами. Таким образом, $f(z)$ должен делиться на многочлен $((z-1)^2+1)(z-2)=z^3-4z^2+6z-4$. Отсюда следует, что многочлен наименьшей степени, удовлетворяющий поставленному условию, есть $c(z^3-4z^2+6z-4)$, где c — любое не равное нулю действительное число. ●

2. Разложение многочлена на множители первой и второй степени

Теорема 2. *Любой многочлен $f(z)$ с действительными коэффициентами разлагается в произведение многочленов первой и второй степени с действительными коэффициентами. Это разложение может быть выбрано так, чтобы каждый множитель второй степени не имел действительных корней.*

□ Доказательство проведем индукцией по степени n многочлена $f(z)$.

При $n=1$ утверждение, очевидно, справедливо. Предположим, что утверждение верно для всех n, меньших некоторого натурального числа k; докажем тогда, что оно верно и при $n=k$.

Пусть $f(z)$ — многочлен степени k с действительными коэффициентами. Согласно теореме 1 § 62 существует комплексный корень α многочлена $f(z)$. Возможны два случая.

I случай: число α — действительное. Тогда, разделив $f(z)$ на $z-\alpha$, получим равенство

$$f(z)=(z-\alpha)f_1(z), \tag{1}$$

где $f_1(z)$ — многочлен степени $k-1$, также имеющий действительные коэффициенты. В силу предположения индукции многочлен $f_1(z)$ разлагается в произведение многочленов первой и второй степени с действительными коэффициентами, причем каждый множитель второй степени не имеет действительных корней. Вследствие равенства (1) то же самое будет верно и для $f(z)$.

II случай: число α не является действительным, т. е. $\alpha=a+bi$, где $a\in\mathbb{R}$, $b\in\mathbb{R}$, $b\neq0$. Согласно теореме 1, число $\bar{\alpha}=a-bi$ также является корнем $f(z)$. Имеем $f(z)=(z-\alpha)g(z)$, где $g(z)$ — многочлен степени $k-1$. Из равенства $f(\bar{\alpha})=0$ следует $(\bar{\alpha}-\alpha)g(\bar{\alpha})=0$; так как $\bar{\alpha}-\alpha=2bi\neq0$, то $g(\bar{\alpha})=0$. Следовательно, многочлен $g(z)$ делится на $z-\bar{\alpha}$, т. е. $g(z)=(z-\bar{\alpha})h(z)$. Итак,

$$f(z)=(z-\alpha)(z-\bar{\alpha})h(z).$$

Положим $\varphi(z)=(z-\alpha)(z-\bar{\alpha})$. Тогда

$$\varphi(z)=(z-(a+bi))(z-(a-bi))=((z-a)-bi)((z-a)+bi)=(z-a)^2+b^2,$$

т. е. $\varphi(z)$ — многочлен с действительными коэффициентами. Значит, и $h(z)$ как результат деления $f(z)$ на $\varphi(z)$ есть также многочлен с действительными коэффициентами. Степень этого многочлена равна $k-2$. В силу предположения индукции $h(z)$ должен разлагаться на множители первой и второй степени с действительными коэффициентами, причем каждый множитель второй степени не имеет действительных корней. Вследствие равенства $f(z)=\varphi(z)h(z)$ то же самое будет верно и для $f(z)$. ■

Пример. Разложить на множители первой и второй степени с действительными коэффициентами многочлен $f(z)=z^4+1$.

○ Имеем

$$z^4+1=z^4+2z^2+1-2z^2=(z^2+1)^2-(\sqrt{2}z)^2=(z^2+1-\sqrt{2}z)(z^2+1+\sqrt{2}z). \bullet$$

Отметим, что никаких общих приемов для разложения многочлена с действительными коэффициентами на множители первой и второй степени не существует. В каждом конкретном случае такое разложение ищется «индивидуально».

3. Разложение рациональной функции на простейшие дроби

Напомним, что *рациональной функцией* действительной переменной называется любая функция вида $P(x)/Q(x)$, где $P(x)$ и $Q(x)$ — многочлены с действительными коэффициентами. Если при этом степень $P(x)$ меньше степени $Q(x)$, то дробь $P(x)/Q(x)$ называют *правильной*.

Очевидно, любая рациональная функция может быть представлена в виде суммы многочлена и правильной дроби; чтобы получить такое представление, достаточно разделить $P(x)$ на $Q(x)$ с остатком.

Особое значение имеют рациональные функции следующего специального вида:

$$\frac{A}{(x-a)^n}, \quad \frac{Cx+D}{(x^2+px+q)^n},$$

где n — любое натуральное число, а многочлен x^2+px+q не имеет действительных корней (это равносильно тому, что дискриминант p^2-4q отрицателен). Рациональные функции такого вида называют *простейшими дробями.*

В § 38 при интегрировании рациональных функций мы использовали прием, заключавшийся в разложении на простейшие дроби. Сформулируем теорему, дающую обоснование такого метода интегрирования (ее доказательство мы не приводим).

Теорема 3. *Правильная дробь $P(x)/Q(x)$ может быть представлена в виде суммы простейших дробей, знаменатели которых $(x-a)^s$ и $(x^2+px+q)^t$ входят в разложение многочлена $Q(x)$ на множители первой и второй степени.*

Например, для правильной дроби

$$\frac{10x^3-5x^2+23x+5}{(x-1)^2(x^2+2x+3)^2}$$

утверждение теоремы означает, что эту дробь можно представить в виде

$$\frac{A}{(x-1)^2}+\frac{B}{x-1}+\frac{Cx+D}{(x^2+2x+3)^2}+\frac{Ex+F}{x^2+2x+3}.$$

Заметим, что нахождение числителей простейших дробей, на которые разлагается данная дробь $P(x)/Q(x)$, проще всего проводится методом неопределенных коэффициентов, изложенным в § 38. Так, для указанной выше дроби после соответствующих выкладок (рекомендуем выполнить их самостоятельно) получается следующее разложение:

$$\frac{10x^3-5x^2+23x+5}{(x-1)^2(x^2+2x+3)^2}=\frac{1}{(x-1)^2}+\frac{5x-1}{(x^2+2x+3)^2}+\frac{-1}{x^2+2x+3}.$$

§ 64. Функции комплексной переменной. Предел, непрерывность, дифференцируемость

1. Сведение функций комплексной переменной к функциям действительной переменной

Понятие функции комплексной переменной вводится по аналогии с понятием функции действительной переменной. Пусть G — некоторое множество точек на комплексной плоскости. Задать *функцию комплексной переменной* (короче — *комплексную функцию*), определенную на G, значит указать правило, сопоставляющее каждому числу $z \in G$ некоторое комплексное число w (зависящее от z). Как и в случае функции действительной переменной, будем писать $w=f(z)$.

Положим $z=x+iy$ и $w=u+iv$. Число w, а, значит, и каждое из чисел u и v зависит от x и y, т. е. является действительной функцией от x и y: $u=u(x, y)$, $v=v(x, y)$.

Обратно, если заданы две действительные функции u и v, определенные на некотором множестве G точек координатной плоскости, то выражение $u(x, y)+iv(x, y)$ представляет собой функцию комплексной переменной $z=x+iy$, определенную на множестве G, рассматриваемом как часть комплексной плоскости.

Таким образом, задание на множестве G комплексной функции $f(z)$ равносильно заданию на G двух действительных функций $u(x, y)$ и $v(x, y)$. Связь между функциями f, u, v устанавливается с помощью равенства

$$f(x+iy)=u(x, y)+iv(x, y). \tag{1}$$

Примеры. 1. Для функции $w=z^2$ (определенной на всем множестве **C**) имеем

$$w=(x+iy)^2=(x^2-y^2)+2xyi.$$

Здесь $u(x, y)=x^2-y^2$, $v(x, y)=2xy$.

2. Для $w=\operatorname{Re}z$ имеем $w=x+0i$, т. е. $u(x, y)=x$, $v(x, y)=0$.

3. Пусть $u(x, y)=\frac{x}{x^2+y^2}$, $v(x, y)=-\frac{y}{x^2+y^2}$ — две действительные функции, определенные на $\mathbf{C}\setminus\{0\}$. Соответствующая им комплексная функция имеет вид

$$w=u(x, y)+iv(x, y)=\frac{x-iy}{x^2+y^2}=\frac{\bar z}{z\bar z}=\frac{1}{z}.$$

2. Предел функции комплексной переменной в точке

Пусть z_0 — комплексное число и ε — положительное действительное число. Условимся называть *ε-окрестностью* числа z_0 множество всех комплексных чисел z, удовлетворяющих условию

$$|z-z_0|<\varepsilon. \tag{2}$$

Геометрически ε-окрестность представляет собой внутренность круга радиуса ε с центром z_0 (рис. 226).

Условимся также называть *проколотой ε-окрестностью* точки z_0 круг (2) без его центра, т. е. множество точек z таких, что $|z-z_0|<\varepsilon$, $z\neq z_0$.

Определение 1. Пусть $f(z)$ — комплексная функция, определенная в некоторой окрестности точки z_0, кроме быть может, самой точки z_0. Число c называется *пределом функции* $f(z)$ при $z\to z_0$, если для любого $\varepsilon>0$ существует такая проколотая окрестность точки z_0, для всех точек z которой выполняется неравенство $|f(z)-c|<\varepsilon$.

В этом случае, как обычно, пишут $c=\lim\limits_{z\to z_0}f(z)$.

Понятие предела комплексной функции сводится к понятию предела действительной функции с помощью следующей теоремы.

Теорема 1. *Пусть* $z_0=x_0+iy_0$, $c=a+ib$. *Тогда равенство* $\lim\limits_{z\to z_0}f(z)=c$ *равносильно двум равенствам:*

$$\lim_{\substack{x\to x_0\\ y\to y_0}}u(x, y)=a,\quad \lim_{\substack{x\to x_0\\ y\to y_0}}v(x, y)=b, \tag{3}$$

где $z=x+iy$, $f(z)=u+iv$.

□ Пусть $\lim\limits_{z\to z_0}f(z)=c$. По определению предела, для всякого $\varepsilon>0$ существует такое $\delta>0$, что из $|Z-z_0|<\delta$ (и $z\neq z_0$) следует $|f(z)-f(z_0)|<e$. Последнее неравенство означает

$$\sqrt{(u(x, y)-a)^2+(v(x, y)-b)^2}<\varepsilon,$$

откуда, в свою очередь, вытекают два неравенства

$$|u(x,\ y)-a|<\varepsilon,\ |v(x,\ y)-b|<\varepsilon. \tag{4}$$

Итак, для всякого $\varepsilon>0$ существует такое $\delta>0$, что из неравенства $|z-\varepsilon_0|<\delta$ (и $z\neq z_0$) следуют оба неравенства (4). Это равносильно равенствам (3).

Обратно, пусть выполняются оба равенства (3). Тогда для любого $\varepsilon>0$ найдется такое $\delta>0$, что из $|z-z_0|<\delta$ (и $z\neq z_0$) следуют оба неравенства (4). Из этих неравенств, в свою очередь, вытекает

$$\sqrt{(u(x,\ y)-a)^2+(v(x,\ y)-b)^2}<\sqrt{\varepsilon^2+\varepsilon^2}=\varepsilon\sqrt{2}.$$

Таким образом, при $z\to z_0$ число $|f(z)-c|$ становится как угодно малым, т. е. $\lim\limits_{z\to z_0} f(z)=c$. ■

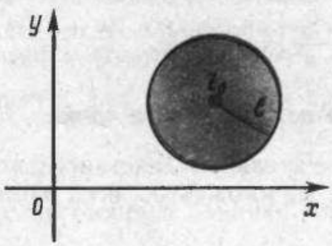


Рис. 226

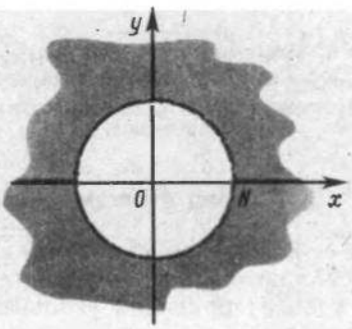


Рис. 227

С помощью теоремы 1 известные свойства пределов действительных функций переносятся на комплексные функции. В частности, если комплексные функции $f(z)$ и $g(z)$ при $z\to z_0$ имеют пределы, соответственно равные c и d, то:

1^0. *Функция $f(z)+g(z)$ при $z\to z_0$ имеет предел, равный $c+d$ (предел суммы равен сумме пределов).*

2^0. *Функция $f(z)g(z)$ при $z\to z_0$ имеет предел, равный cd (предел произведения равен произведению пределов).*

3^0. *В случае $d\neq 0$ функция $\frac{f(z)}{g(z)}$ при $z\to z_0$ имеет предел, равный c/d (предел частного равен отношению пределов числителя и знаменателя, если предел знаменателя отличен от нуля).*

3. Расширенная комплексная плоскость

Многие факты, касающиеся комплексных функций, можно сформулировать в более общем виде, если ввести понятие бесконечности. Условимся дополнить множество **C** еще одним «числом», называемым *бесконечностью* и обозначаемым ∞. Соответствующая этому числу (воображаемая) точка также обозначается символом ∞. Множество $\mathbf{C}\cup\{\infty\}$ называется *расширенной комплексной плоскостью* и обозначается $\widehat{\mathbf{C}}$.

Определенные ранее операции над комплексными числами распространим на все множество $\widehat{\mathbf{C}}$ с помощью следующих соглашений:

$$\infty + c = \infty, \qquad \text{где } c \in \mathbf{C};$$
$$\infty \cdot c = \infty, \qquad \text{где } c \in \mathbf{C},\ c \neq 0;$$
$$\infty \cdot \infty = \infty;$$
$$\frac{c}{\infty} = 0, \qquad \text{где } c \in \mathbf{C},\ c \neq 0.$$

Распространим на $\bar{\mathbf{C}}$ и понятие окрестности. А именно, *окрестностью точки* ∞ будем считать множество точек $z \in \mathbf{C}$, удовлетворяющих условию $|z| > N$, где N — заданное неотрицательное число; условимся относить к этому множеству и саму точку ∞. *Проколотой окрестностью точки* ∞ будем называть любую окрестность точки ∞ без самой этой точки.

Геометрически окрестность точки ∞ представляет собой внешность круга радиуса N с центром в начале координат (рис. 227). Ясно, что с увеличением N такая окрестность становится «все меньше».

4. Предел на бесконечности и бесконечный предел

Приведенное в п. 2 определение предела функции при $z \to z_0$ не охватывает случаи, когда одно из чисел z_0, c (или оба) равно ∞. Поэтому дадим более общее определение предела, включающее и эти случаи.

Прежде всего расширим само понятие комплексной функции, предусмотрев в нем возможность бесконечных значений для z и $f(z)$.

Пусть G — подмножество расширенной комплексной плоскости $\bar{\mathbf{C}}$. Комплексной функцией, определенной на G, называется любое отображение f множества G в $\bar{\mathbf{C}}$.

Определение 2. Пусть функция $f(z)$ определена в некоторой окрестности точки $z_0 \in \bar{\mathbf{C}}$, кроме, быть может, самой точки z_0. Число $c \in \bar{\mathbf{C}}$ называется *пределом* $f(z)$ при $z \to z_0$, если для любой окрестности $V(c)$ точки c найдется такая проколотая окрестность $U(z_0)$ точки z_0, что для всех $z \in U(z_0)$ выполняется условие $f(z) \in V(c)$.

Подчеркнем еще раз, что в этом определении любое из чисел z_0, c может быть равно ∞. Для большей отчетливости конкретизируем данное определение в каждом из таких случаев.

1. Равенство $\lim\limits_{z \to \infty} f(z) = c$ при конечном c означает следующее: для любого $\varepsilon > 0$ найдется такое $N > 0$, что при всех $z \in \mathbf{C}$, удовлетворяющих условию $|z| > N$, справедливо неравенство $|f(z) - c| < \varepsilon$.

2. Равенство $\lim\limits_{z \to z_0} f(z) = \infty$ при конечном z_0 означает, что для любого $N > 0$ найдется такое $\delta > 0$, что при всех z, удовлетворяющих условиям $|z - z_0| < \delta$, $z \neq z_0$, выполняется неравенство $|f(z)| > N$.

3. Равенство $\lim\limits_{z \to \infty} f(z) = \infty$ означает, что для любого $N > 0$ найдется такое $M > 0$, что при всех $z \in \mathbf{C}$, удовлетворяющих условию $|z| > M$, выполняется неравенство $|f(Z)| > N$.

Предоставляем проверить самостоятельно следующие свойства пределов:

$$\text{если } \lim_{z \to z_0} f(z) \neq 0, \ \lim_{z \to z_0} g(z) = 0, \text{ то } \lim_{z \to z_0} \frac{f(z)}{g(z)} = \infty;$$

$$\text{если } \lim_{z \to z_0} f(z) \neq \infty, \ \lim_{z \to z_0} g(z) = \infty, \text{ то } \lim_{z \to z_0} \frac{f(z)}{g(z)} = 0.$$

Отметим, что для предела отношения двух многочленов при $z \to \infty$, как и в случае действительной переменной, имеют место следующие формулы:

$$\lim_{z\to\infty}\frac{a_nz^n+\cdots+a_0}{b_mz^m+\cdots+b_0}=\begin{cases}0, & \text{если } n<m,\\ \dfrac{a_n}{b_m}, & \text{если } n=m,\\ \infty, & \text{если } n>m,\end{cases}$$

где предполагается, что $a_n\neq 0$, $b_m\neq 0$.

5. Непрерывность функции комплексной переменной

Определение 3. Пусть функция комплексной переменной задана на множестве G и пусть $z_0\in G$. Функцию f называют *непрерывной в точке* z_0, если $\lim\limits_{z\to z_0}f(z)=f(z_0)$.

Из свойств пределов (см. п. 2) вытекают следующие утверждения.

Сумма и произведение функций, непрерывных в точке z_0, непрерывны в этой точке. Если функция f непрерывна в точке z_0 и при этом $f(z_0)\neq 0$, то и функция $1/f$ непрерывна в z_0.

В частности, отсюда следует, что любая дробно-рациональная функция, т. е. функция вида

$$\frac{a_nz^n+a_{n-1}z^{n-1}+\cdots+a_0}{b_mz^m+b_{m-1}z^{m-1}+\cdots+b_0}\quad (a_n\neq 0,\ b_m\neq 0),$$

непрерывна во всех точках, где ее знаменатель отличен от нуля.

На функции комплексной переменной переносится без изменений теорема о непрерывности сложной функции:

Пусть f и g — две функции комплексной переменной. Если g непрерывна в точке z_0, а f непрерывна в точке $g(z_0)$, то сложная функция $w=f(g(z))$ непрерывна в точке z_0.

Примеры. 1. Функция $f(z)=\operatorname{Re}(z)$ непрерывна в любой точке $z_0\in\mathbf{C}$, так как из $x+iy\to x_0+iy_0$ следует $x\to x_0$, т. е. $\lim\limits_{z\to z_0}f(z)=f(z_0)$.

2. Функция $f(z)=\bar z$ непрерывна в любой точке $z_0\in\mathbf{C}$, так как из $x+iy\to x_0+iy_0$ следует $x-iy\to x_0-iy_0$, т. е. $\lim\limits_{z\to z_0}f(z)=f(z_0)$.

3. Функция $f(z)=\dfrac{1}{z+i}$ непрерывна во всех точках z, где ее знаменатель не равен нулю, т. е. при всех z, отличных от $-i$.

6. Дифференцирование функции комплексной переменной

Приводимые ниже определения вводятся по аналогии со случаем функций действительной переменной.

Определение 4. Пусть функция $f(z)$ комплексной переменной определена в некоторой окрестности точки z_0. *Производной* этой функции в точке z_0 называется число

$$f'(z_0)=\lim_{z\to z_0}\frac{\Delta f}{\Delta z},$$

где $\Delta z=z-z_0$, $\Delta f=f(z)-f(z_0)$.

Определение 5. Функция $f(z)$ называется *дифференцируемой в точке* z_0, если ее приращение Δf в этой точке можно представить в виде

$$\Delta f=c\Delta z+\alpha\Delta z,$$

где c — некоторое комплексное число, не зависящее от Δz, а α, вообще говоря, зависит от Δz, причем $\lim\limits_{z\to z_0}\alpha=0$.

Как и в случае функций действительной переменной, справедлива следующая теорема.

Теорема 2. *Для дифференцируемости функции $f(z)$ в точке z_0 необходимо и достаточно существование производной $f'(z_0)$.*

Доказательство опускаем, так как оно полностью повторяет доказательство аналогичной теоремы для функций действительной переменной.

Примеры. 1. Найти производную функции $f(z)=z^2$.

○ Имеем $\Delta f=(z_0+\Delta z)^2-z_0^2=(2z_0+\Delta z)\Delta z$. Отсюда

$$\frac{\Delta f}{\Delta z}=\frac{(2z_0+\Delta z)\Delta z}{\Delta z}=2z_0+\Delta z.$$

Это отношение при $\Delta z\to 0$ имеет пределом число $2z_0$. Итак, $f'(z)=2z$. ●

2. Показать, что функция $f(z)=\operatorname{Re}(z)$ не дифференцируема ни в одной точке.

○ Полагая $z_0=x_0+iy_0$, $\Delta z=\Delta x+i\Delta y$, находим

$$\frac{\Delta f}{\Delta z}=\frac{(x_0+\Delta x)-x_0}{\Delta x+i\Delta y}=\frac{\Delta x}{\Delta x+i\Delta y}=\begin{cases}0, & \text{если } \Delta x=0,\\ 1, & \text{если } \Delta y=0.\end{cases}$$

Отсюда видно, что предела (единого) при $\Delta z\to 0$ для отношения $\frac{\Delta f}{\Delta z}$ не существует. Следовательно, функция $f(z)=\operatorname{Re}(z)$ не дифференцируема ни в одной точке. ●

Так как вывод правил дифференцирования суммы, произведения и частного основан на свойствах пределов, а эти свойства для функций действительной и комплексной переменных формулируются одинаково, то для функций комплексной переменной справедливы следующие формулы:

$$(f+g)'=f'+g';\ (cf)'=cf',\ \text{где } c=\text{const}\in\mathbf{C};$$

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

Кроме того, для функций комплексной переменной остается справедливым правило дифференцирования сложной функции:

$$[f(g(z))]'=f'(g(z))g'(z).$$

Из указанных свойств производной точно так же, как и в случае действительной переменной, выводится дифференцируемость степенной функции z^n (с натуральным показателем n) и формула

$$(z^n)'=nz^{n-1}.$$

Вообще, любой многочлен от z представляет собой функцию, дифференцируемую в каждой точке комплексной плоскости; любая дробно-рациональная функция также дифференцируема в каждой точке, где ее знаменатель отличен от нуля.

Аналогично случаю функций действительной переменной вводятся производные высших порядков: второго, третьего и т. д.

7. Условия существования производной (условия Даламбера — Эйлера)

Поставим вопрос: каким условиям должны удовлетворять функции $u(x, y)$ и $v(x, y)$ действительных переменных x и y для того, чтобы функция

$$f(z)=u(x, y)+iv(x, y) \tag{5}$$

комплексной переменной $z=x+iy$ была дифференцируема в данной точке $z_0=x_0+iy_0$? Ответ дает следующее утверждение.

Теорема 3. *Пусть функции $u(x, y)$ и $v(x, y)$ определены в некоторой окрестности точки $(x_0; y_0)$ и дифференцируемы в этой точке (как функции действительных переменных x и y). Тогда для дифференцируемости в точке $z_0=x_0+iy_0$ комплексной функции $f(z)$, определенной равенством* (5), *необходимо и достаточно выполнение в этой точке условий*

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}. \tag{6}$$

□ Необходимость. Пусть функция $f(z)$ дифференцируема в точке z_0, т. е. существует предел

$$\lim_{\Delta z\to 0}\frac{\Delta f}{\Delta z}=\lim_{\Delta z\to 0}\frac{\Delta u+i\Delta v}{\Delta x+i\Delta y}, \tag{7}$$

равный $f'(z_0)$. В частности, если положить $\Delta y=0$, то получим

$$\lim_{\Delta x\to 0}\frac{\Delta u+i\Delta v}{\Delta x}=f'(z_0), \tag{8}$$

где $\Delta u=u(x_0+\Delta x,\ y_0)-u(x_0,\ y_0)$, $\Delta v=v(x_0+\Delta x,\ y_0)-v(x_0,\ y_0)$. Выраженне под знаком предела в равенстве (8) есть $\frac{\Delta u}{\Delta x}+i\frac{\Delta v}{\Delta x}$. Так как по условию функции $u(x, y)$ и $v(x, y)$ дифференцируемы в точке $(x_0; y_0)$, то существуют пределы $\lim\limits_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=\frac{\partial u}{\partial x}$, $\lim\limits_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=\frac{\partial v}{\partial x}$, и из соотношения (8) следует

$$\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=f'(z_0). \tag{9}$$

Аналогично, полагая $\Delta x=0$, из равенства (7) получим

$$\lim_{\Delta y\to 0}\frac{\Delta u+i\Delta v}{i\Delta y}=f'(z_0),$$

что равносильно

$$\frac{\partial v}{\partial y}-i\frac{\partial u}{\partial y}=f'(z_0). \tag{10}$$

Сравнивая (9) и (10), приходим к соотношениям (6).

Достаточность. Из дифференцируемости функций $u(x, y)$ и $v(x, y)$ в точке $(x_0; y_0)$ следует, что в этой точке

$$\Delta u=\frac{\partial u}{\partial x}\Delta x+\frac{\partial u}{\partial y}\Delta y+\alpha\sqrt{\Delta x^2+\Delta y^2},$$

$$\Delta v=\frac{\partial v}{\partial x}\Delta x+\frac{\partial v}{\partial y}\Delta y+\beta\sqrt{\Delta x^2+\Delta y^2},$$

где величины α и β (зависящие от Δx, Δy) стремятся к нулю при $\Delta x\to 0$, $\Delta y\to\to 0$ (см. § 45). Отсюда имеем

$$\frac{\Delta f}{\Delta z}=\frac{\Delta u+i\Delta v}{\Delta x+i\Delta y}=\frac{\frac{\partial u}{\partial x}\Delta x+\frac{\partial u}{\partial y}\Delta y+i\left(\frac{\partial v}{\partial x}\Delta x+\frac{\partial v}{\partial y}\Delta y\right)}{\Delta x+i\Delta y}+\frac{\alpha+i\beta}{\Delta x+i\Delta y}\sqrt{\Delta x^2+\Delta y^2}.$$

Если в правой части $\frac{\partial u}{\partial x}$ и $\frac{\partial v}{\partial y}$ заменить по формулам (6), а также заметить, что величина

$$\xi=\frac{\sqrt{\Delta x^2+\Delta y^2}}{\Delta x+i\Delta y}=\sqrt{\frac{(\Delta x+i\Delta y)(\Delta x-i\Delta y)}{(\Delta x+i\Delta y)^2}}=\sqrt{\frac{\Delta x-i\Delta y}{\Delta x+i\Delta y}}$$

имеет модуль, равный 1 $\left(\text{так как } \left|\frac{\bar z}{z}\right|=1 \text{ для любого } z\in\mathbf{C},\ z\neq 0\right)$, и тем самым ограничена, то получим

$$\frac{\Delta f}{\Delta z}=\frac{\frac{\partial u}{\partial x}\Delta x-\frac{\partial v}{\partial x}\Delta y+i\frac{\partial v}{\partial x}\Delta x+i\frac{\partial u}{\partial x}\Delta y}{\Delta x+i\Delta y}+(\alpha+i\beta)\xi=\frac{\frac{\partial u}{\partial x}(\Delta x+i\Delta y)+i\frac{\partial v}{\partial x}(\Delta x+i\Delta y)}{\Delta x+i\Delta y}+$$
$$+(\alpha+i\beta)\xi=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}+(\alpha+i\beta)\xi.$$

Отсюда вытекает существование предела

$$\lim_{\Delta z\to 0}\frac{\Delta f}{\Delta z}=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x},$$

т. е. дифференцируемость функции $f(z)$ в точке z_0. ■

Условия (6) называются *условиями Даламбера — Эйлера**.

Примеры. 1. Показать, что функция $f(z)=\bar z$ не дифференцируема ни в одной точке.

○ Имеем $f(z)=f(x+iy)=x-iy$, откуда $u(x,y)=x$, $v(x,y)=-y$, $\frac{\partial u}{\partial x}=1$, $\frac{\partial v}{\partial y}=-1$. Условия Даламбера — Эйлера не выполняются. ●

2. Проверить выполнение условий Даламбера — Эйлера для функции $f(z)=z^2$.

○ Имеем $f(z)=(x+iy)^2=(x^2-y^2)+2xyi$, откуда $u(x,y)=x^2-y^2$, $v(x,y)=2xy$ и $\frac{\partial u}{\partial x}=2x$, $\frac{\partial v}{\partial y}=2x$, $\frac{\partial u}{\partial y}=-2y$, $\frac{\partial v}{\partial x}=2y$. Мы видим, что условия Даламбера — Эйлера выполняются; следовательно функция z^2 дифференцируема.

3. Найти мнимую часть дифференцируемой функции $f(z)$, если известна ее действительная часть $u(x,y)=x^3-3xy^2-x+1$.

○ Имеем $f(z)=u(x,y)+iv(x,y)$. Из условий Даламбера — Эйлера следует

$$\frac{\partial v}{\partial y}=\frac{\partial u}{\partial x}=3x^2-3y^2-1,\quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}=6xy.$$

Таким образом, обе частные производные функции $v(x,y)$ известны. Это позволяет найти и саму функцию $v(x,y)$ (задачи подобного рода рассматривались в § 45). Опуская соответствующие вычисления, приведем окончательный результат: $v(x,y)=3x^2y-y^3-y+C$, где C — произвольная постоянная. Рекомендуем проверить, что $f(z)=z^3-z+1$. ●

8. Область. Дифференцируемость функции в области

Определение 6. Открытое множество G на плоскости называется *областью*, если любые две точки z_1 и z_2, принадлежащие G, можно соединить с помощью кривой, целиком лежащей в G.

Так, множество G, изображенное на рис. 228, является областью; в противоположность этому множество H на рис. 229 — не область.

---

* Л. Эйлер (1707—1783) — русский математик, уроженец Швейцарии. Условия (6) называют также *условиями Коши — Римана*.

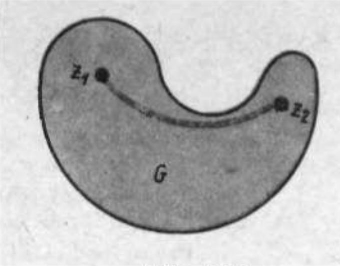


Рис. 228

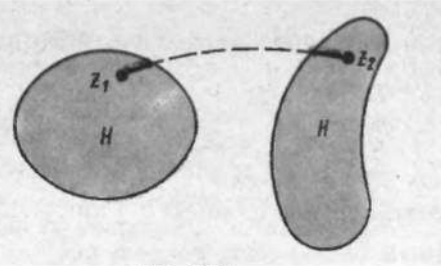


Рис. 229

Определение 7. Комплексная функция $f(z)$, определенная в области G комплексной плоскости, называется *дифференцируемой в этой области,* если она дифференцируема в каждой точке $z \in G$. Если при этом производная $f'(z)$ непрерывна в каждой точке $z \in G$, то функция $f(z)$ называется *непрерывно дифференцируемой* в G.

Можно показать, что из дифференцируемости функции в области G вытекает ее непрерывная дифференцируемость в G. В случае функции действительной переменной такое утверждение неверно.

§ 65. Степенные ряды в комплексной области

1. Числовые ряды с комплексными членами

Теория рядов с комплексными членами во многом повторяет уже известную нам теорию рядов с действительными членами. Поэтому исходные понятия, касающиеся рядов с комплексными членами, приведем без пояснений.

Пусть $a_1, a_2, \ldots$ — бесконечная последовательность комплексных чисел. *Рядом*, составленным из членов $a_1, a_2, \ldots$, называется выражение

$$a_1 + a_2 + \cdots + a_n + \cdots . \tag{1}$$

Определение 1. Ряд (1) называется *сходящимся*, если существует предел последовательности его частичных сумм A_n, где

$$A_n = a_1 + a_2 + \cdots + a_n \quad (n = 1,\ 2,\ \ldots).$$

В противном случае, т. е. если не существует указанного предела, ряд называется *расходящимся*.

В случае сходящегося ряда число $A = \lim\limits_{n \to \infty} A_n$ называют *суммой ряда*; говорят также, что *ряд сходится к сумме A*.

Определение 2. Ряд (1) называется *абсолютно сходящимся*, если сходится ряд, составленный из модулей его членов, т. е. ряд

$$|a_1| + |a_2| + \cdots + |a_n| + \cdots . \tag{2}$$

Определение 3. Если дан ряд (1), то ряд

$$a_{n+1} + a_{n+2} + \cdots \tag{3}$$

называется *остатком* ряда (1) после n-го члена.

Укажем некоторые свойства числовых рядов с комплексными членами. Доказательства этих свойств, за единственным исключением (свойство 6^0), полностью повторяют доказательства соответствующих свойств для рядов с действительными членами (см. § 50).

1^0. *Если ряд* (1) *сходится, то его общий член* a_n *стремится к нулю при* $n \to \infty$.

2^0. *Если ряды $a_1+a_2+\cdots$ и $b_1+b_2+\cdots$ сходятся соответственно к суммам A и B, то ряд $(a_1+b_1)+(a_2+b_2)+\cdots$ сходится к сумме $A+B$.*

3^0. *Если ряд $a_1+a_2+\cdots$ сходится к сумме A, то ряд $\lambda a_1+\lambda a_2+\cdots$ сходится к сумме λA.*

4^0. *Если ряд сходится и имеет сумму A, то ряд, полученный из него произвольной перестановкой членов, также сходится и имеет ту же сумму A.*

5^0. *Если ряд* (1) *сходится, то сходится и ряд* (3) *(где n — любое натуральное), причем*

$$A=a_1+\cdots+a_n+R_n,$$

где A — сумма ряда (1), *а R_n — сумма ряда* (3).

6^0. *Абсолютно сходящийся ряд является сходящимся.*

□ Пусть ряд (1) сходится абсолютно, т. е. сходится ряд, составленный из модулей его членов. Положим $a_n=p_n+iq_n$ $(n=1,\ 2,\ \ldots)$, где p_n и q_n — действительные числа. Тогда

$$|p_n|\leqslant\sqrt{p_n^2+q_n^2}=|a_n|,\quad |q_n|\leqslant\sqrt{p_n^2+q_n^2}=|a_n|.$$

Согласно признаку сравнения рядов с положительными членами (см. теорему 2 § 51), оба ряда $|p_1|+|p_2|+\cdots$ и $|q_1|+|q_2|+\cdots$ сходятся. Следовательно, сходятся и ряды $p_1+p_2+\cdots$ и $q_1+q_2+\cdots$. Очевидно, имеем $A_n=P_n+iQ_n$, где A_n, P_n, Q_n — соответственно частичные суммы рядов $a_1+a_2+\cdots$, $p_1+p_2+\cdots$, $q_1+q_2+\cdots$. Так как существует каждый из пределов $\lim\limits_{n\to\infty}P_n$ и $\lim\limits_{n\to\infty}Q_n$, то существует и предел $\lim\limits_{n\to\infty}A_n$, т. е. ряд (1) сходится. ■

7^0. *Если ряды $a_1+a_2+\cdots$ и $b_1+b_2+\cdots$ абсолютно сходятся, то сходится абсолютно и ряд, составленный из членов a_ib_j $(i=1,\ 2,\ \ldots,\ j=1,\ 2,\ \ldots)$, расположенных в каком угодно порядке, причем сумма такого ряда равна AB, где A — сумма первого ряда, а B — сумма второго* (см. теорему 2 § 52).

2. Степенные ряды в комплексной области

Определение 4. *Степенным рядом* в комплексной области называется ряд вида

$$a_0+a_1(z-z_0)+a_2(z-z_0)^2+\cdots, \tag{4}$$

где a_0, a_1, a_2, ... и z_0 — фиксированные комплексные числа, а z — переменная, которая может принимать любые комплексные значения.

Определение 5. Пусть z^* — некоторое комплексное число. Говорят, что ряд (4) *сходится* в точке z^*, если при подстановке в него вместо z числа z^* получается сходящийся ряд с комплексными членами. В противном случае говорят, что ряд (4) *расходится* в точке z^*.

Определение 6. Множество точек z, в которых ряд (4) сходится, называется *областью сходимости* этого ряда.

Как и в случае степенных рядов в действительной области, справедливо следующее утверждение.

Теорема 1 (теорема Абеля). *Если степенной ряд* (4) *сходится в точке z^*, то он сходится, и притом абсолютно, в любой точке z, которая лежит внутри окружности с центром z_0, проходящей через z^** (рис. 230), *т. е. для всех z таких, что $|z-z_0|<|z^*-z_0|$.*

Из этой теоремы в точности так же, как для степенных рядов в действительной области, вытекает, что для степенных рядов (4) возможны лишь следующие случаи:

1) ряд (4) сходится только при $z = z_0$;

2) ряд сходится при всех значениях z; в этом случае его сходимость при любом z является абсолютной;

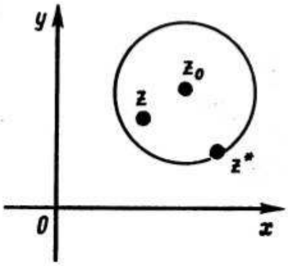


Рис. 230

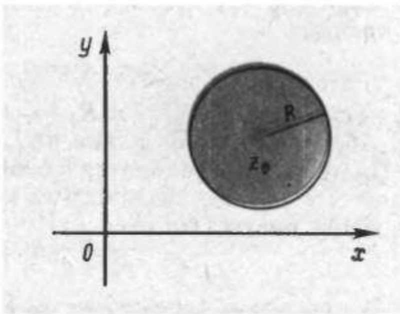


Рис. 231

3) существует такое $R > 0$, что ряд сходится при любом значении z, для которого $|z - z_0| < R$ (рис. 231), и расходится при любом z, для которого $|z - z_0| > R$; в этом случае сходимость для всех z таких, что $|z - z_0| < R$, является абсолютной.

В случае 3 число R называется *радиусом сходимости* степенного ряда (4), а круг $|z - z_0| < R$ — *кругом сходимости* ряда. Что касается граничной окружности $|z - z_0| = R$, то в некоторых ее точках ряд может сходиться, в других — расходиться.

Понятием радиуса сходимости удобно пользоваться и в случаях 1 и 2, полагая $R = 0$ в случае 1 и $R = \infty$ в случае 2.

Существенное отличие по сравнению со степенными рядами в действительной области заключается в том, что для степенного ряда в действительной области существует интервал сходимости, в то время как для степенного ряда в комплексной области — круг сходимости.

3. Вычисление радиуса сходимости степенного ряда

Радиус сходимости степенного ряда чаще всего находят с помощью признака сходимости Даламбера. Из него, как и в случае степенных рядов в действительной области, вытекает следующее утверждение.

Теорема 2. *Пусть для ряда* (4) *существует*

$$\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = p. \qquad (5)$$

Тогда радиус сходимости ряда определяется формулой $R = 1/p$. *Сюда включаются и случаи* $p = 0$ *и* $p = \infty$; *если* $p = 0$, *то* $R = \infty$, *а если* $p = \infty$, *то* $R = 0$.

Пример. Найти круг сходимости ряда:

а) $1 + (z - z_0) + (z - z_0)^2 + (z - z_0)^3 + \cdots$;

б) $z + \frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \cdots$;

в) $1+z+\frac{z^2}{2!}+\frac{z^3}{3!}+\cdots$;

г) $1!z+2!z^2+3!z^3+\cdots$.

○ а) Для данного ряда число p в формуле (5) очевидно, равно 1; следовательно, $R=1$. Ряд сходится внутри круга $|z-z_0|<1$ и расходится вне этого круга. В точках окружности $|z-z_0|=1$ ряд также расходится, поскольку его общий член не стремится к нулю.

б) В этом случае имеем $p=\lim\limits_{n\to\infty}\left(\frac{1}{n+1}:\frac{1}{n}\right)=1$; значит, $R=1$. Ряд сходится внутри круга $|z|<1$ и расходится вне этого круга. Что касается граничной окружности $|z|=1$, то в некоторых ее точках ряд сходится (например, при $z=-1$), в некоторых — расходится (например, при $z=1$).

в) Здесь $p=\lim\limits_{n\to\infty}\left(\frac{1}{n!}:\frac{1}{(n-1)!}\right)=\lim\limits_{n\to\infty}\frac{1}{n}=0$; поэтому $R=\infty$. Ряд сходится при любом z.

г) Находим $p=\lim\limits_{n\to\infty}\frac{(n+1)!}{n!}=\lim\limits_{n\to\infty}(n+1)=\infty$; следовательно, $R=0$. Ряд сходится только в точке $z_0=0$. ●

4. Свойства степенного ряда внутри его круга сходимости

Пусть степенной ряд (4) имеет радиус сходимости R. Как и в случае степенных рядов с действительными членами, справедливы следующие утверждения.

Теорема 3. *В любом круге $|z-z_0|<r$, где $r<R$, ряд (4) правильно сходится.*

Для степенных рядов в действительной области это было доказано в § 55. Из этого утверждения, как и в § 55, вытекает, что *сумма $F(z)$ ряда (4) является непрерывной функцией внутри его круга сходимости.*

Теорема 4. *Ряд*

$$a_1(z-z_0)+2a_2(z-z_0)+3a_3(z-z_0)^2+\dots, \tag{6}$$

полученный почленным дифференцированием ряда (4), имеет тот же радиус сходимости R, что и ряд (4). Сумма ряда (6) внутри его круга сходимости совпадает с $F'(z)$.

Последнее утверждение теоремы означает, что при $|z-z_0|<R$ справедливо равенство

$$[a_0+a_1(z-z_0)+a_2(z-z_0)^2+\dots]'=(a_0)'+(a_1(z-z_0))'+(a_2(z-z_0)^2)'+\dots,$$

т. е. *степенной ряд внутри его круга сходимости можно дифференцировать почленно.*

Разумеется, ряд (6) можно продифференцировать еще раз; тогда получим ряд

$$2\cdot 1a_2+3\cdot 2a_3(z-z_0)+4\cdot 3a_4(z-z_0)^2+\dots,$$

сходящийся в круге $|z-z_0|<R$ к функции $F''(z)$. Вообще, ряд, полученный из (4) полученным дифференцированием n раз, сходится в круге $|z-z_0|<R$ к функции $F^{(n)}(z)$ $(n=1, 2, \dots)$.

Доказательство теоремы 4 мы не приводим.

Пример. Как мы уже знаем, степенной ряд

$$1+z+z^2+z^3+\dots$$

сходится в круге $|z|<1$ к сумме $F(z)=\frac{1}{1-z}$. Следовательно, «продифференцированный» ряд

$$1+2z+3z^2+\dots$$

сходится в том же круге к функции $F'(z)=\frac{1}{(1-z)^2}$. В частности, отсюда вытекает, что при $|z|<1$ справедливо соотношение

$$z+2z^2+3z^3+\dots=\frac{z}{(1-z)^2}.$$

5. Ряд Тейлора

Если ряд (4) продифференцировать почленно n раз, то получим ряд, первый член которого равен $n(n-1)\dots2\cdot1a_n$; следующий член ряда содержит множитель $z-z_0$, следующий за ним член — множитель $(z-z_0)^2$ и т. д. Сумма этого ряда, как уже отмечалось выше, при всех z из круга $|z-z_0|<R$ равна $F^{(n)}(z)$. В частности, при $z=z_0$ сумма указанного ряда равна $F^{(n)}(z_0)$. Отсюда получаем равенство $n!a_n=F^{(n)}(z_0)$ или

$$a_n=\frac{F^{(n)}(z_0)}{n!}, \tag{7}$$

где $n=1, 2, \dots$, очевидно, это равенство справедливо и при $n=0$, поскольку $F(z_0)=a_0$.

Определение 7. Пусть функция $f(z)$ имеет в точке z_0 производные всех порядков. Тогда степенной ряд $\sum\limits_{n=0}^{\infty}\frac{f^{(n)}(z_0)}{n!}(z-z_0)^n$ называется *рядом Тейлора* для функции $f(z)$ в точке z_0.

Полученной выше формуле (7) мы можем теперь дать следующее истолкование:

Теорема 5. *Если функция $F(z)$ разлагается в некотором круге $|z-z_0|<R$ в ряд по степеням $z-z_0$, то этот ряд не может быть ничем иным, как ее рядом Тейлора в точке z_0.*

В частности, для функции, тождественно равной нулю, имеется лишь единственное разложение в ряд Тейлора (в произвольной точке z_0), а именно,

$$0+0(z-z_0)+0(z-z_0)^2+\cdots.$$

Выясним, что дает применение последней теоремы к функции вида

$$F(z)=a_0+a_1z+a_2z^2+\cdots+a_mz^m,$$

т. е. к многочлену m-й степени от z. Прежде всего ясно, что такая функция представляется в виде сходящегося (при всех z) степенного ряда в любой точке z_0: для этого достаточно заметить, что при любом $n=0, 1, 2, \dots, m$ справедливо равенство $a_nz^n=a_n(z_0+(z-z_0))^n$, а затем раскрыть выражение $(z_0+(z-z_0))^n$ по формуле бинома Ньютона. Следовательно, разложение $F(z)$ в степенной ряд в точке z_0 непременно является ее рядом Тейлора (в точке z_0). Если еще заметить, что все производные функции $F(z)$ порядка выше m равны нулю, то придем к формуле

$$F(z)=\sum_{n=1}^{m}\frac{F^{(n)}(z_0)}{n!}(z-z_0)^n, \tag{8}$$

справедливой для всякого многочлена $F(z)$ степени m. Как и в случае функций действительной переменной, формула (8) называется *формулой Тейлора для многочленов*.

§ 66. Аналитические функции

1. Аналитичность функции в точке

Определение 1. Функция $f(z)$ комплексной переменной z называется *аналитической в точке* z_0, если существует такое $r>0$, что в круге $|z-z_0|<r$ функцию $f(z)$ можно представить в виде суммы степенного ряда по степеням $z-z_0$.

Другими словами, функция $f(z)$ аналитична в точке z_0, если существуют такие коэффициенты a_0, a_1, a_2, ..., что при любом z, удовлетворяющем условию $|z-z_0|<r$, справедливо разложение

$$f(z)=a_0+a_1(z-z_0)+a_2(z-z_0)^2+\dots .$$

Как уже отмечалось, стоящий справа ряд не может быть ничем иным, кроме ряда Тейлора для $f(z)$ в точке z_0.

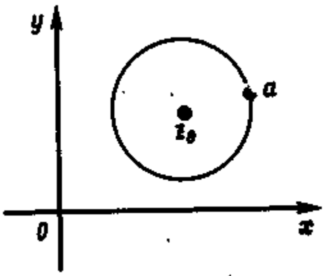


Рис. 232

Примеры. 1. Формула Тейлора для многочленов m-й степени (см. равенство (8) § 65) показывает, что многочлен есть аналитическая функция (в любой точке).

2. Показать, что функция $f(z)=\dfrac{1}{z-a}$ аналитична в любой точке z_0, $z_0\neq a$.

○ Имеем

$$\frac{1}{z-a}=\frac{1}{(z_0-a)+(z-z_0)}=\frac{1}{z_0-a}\cdot\frac{1}{1+\frac{z-z_0}{z_0-a}}=-\frac{1}{a-z_0}\cdot\frac{1}{1-\frac{z-z_0}{a-z_0}}.$$

Используя формулу суммы бесконечной геометрической прогрессии со знаменателем $\left|\frac{z-z_0}{a-z_0}\right|<1$, получим

$$\frac{1}{1-\frac{z-z_0}{a-z_0}}=1+\frac{z-z_0}{a-z_0}+\left(\frac{z-z_0}{a-z_0}\right)^2+\dots .$$

Следовательно, при $|z-z_0|<|a-z_0|$ имеем

$$\frac{1}{z-a}=-\frac{1}{a-z_0}-\frac{1}{(a-z_0)^2}(z-z_0)-\frac{1}{(a-z_0)^3}(z-z_0)^2-\dots .$$

что и доказывает аналитичность функции $\dfrac{1}{z-a}$ в точке z_0. Заметим, что круг сходимости полученного ряда есть $|z-z_0|<|a-z_0|$, т. е. круг с центром z_0 и радиусом $|a-z_0|$; граничная окружность этого круга проходит через точку a (рис. 232). Указанный круг является наименьшим из всех кругов с центром z_0, внутри которых функция $\dfrac{1}{z-a}$ является непрерывной. ●

2. Свойства функций, аналитических в точке

1^0. *Если функции $f(z)$ и $g(z)$ аналитичны в точке z_0, то и функция $f(z)+g(z)$ аналитична в z_0.*

☐ Так как в некотором круге с центром z_0 справедливы равенства $f(z)=$ $=\sum_{n=0}^{\infty} a_n(z-z_0)^n$, $g(z)=\sum_{n=0}^{\infty} b_n(z-z_0)^n$, то в этом же круге имеем

$$f(z)+g(z)=\sum_{n=0}^{\infty}(a_n+b_n)(z-z_0)^n. \blacksquare$$

2^0. *Если функции $f(z)$ и $g(z)$ аналитичны в точке z_0, то и функция $f(z)g(z)$ аналитична в точке z_0.*

Доказательство аналогично.

3^0. *Если функция $f(z)$ аналитична в точке z_0, то функция $f(z-a)$ аналитична в точке z_0+a.*

☐ Из равенства $f(z)=\sum_{n=0}^{\infty} a_n(z-z_0)^n$, справедливого при любом z таком, что $|z-z_0|<r$, следует равенство

$$f(z-a)=\sum_{n=0}^{\infty} a_n(z-(z_0+a))^n,$$

верное для всех z, удовлетворяющих условию $|(z-a)-z_0|<r$ или, что то же, условию $|z-(z_0+a)|<r$. $\blacksquare$

В дальнейшем мы рассмотрим еще два свойства аналитических функций: аналитичность частного $f(z)/g(z)$ и аналитичность сложной функции $f(g(z))$ (см. п. 2 § 69).

Согласно теореме 2 § 62, любой многочлен $Q(z)$ можно представить в виде произведения $a(z-\alpha_1)...(z-\alpha_n)$. Отсюда имеем

$$\frac{1}{Q(z)}=\frac{1}{a(z-\alpha_1)...(z-\alpha_n)},$$

и так как любая функция вида $\frac{1}{z-\alpha}$ аналитична (во всякой точке z, отличной от α), то на основании свойства 2^0 приходим к заключению, что функция $1/Q(z)$ аналитична во всякой точке z, где $Q(z)\neq 0$. Отсюда вытекает следующее утверждение.

Любая дробно-рациональная функция $P(z)/Q(z)$ аналитична в каждой точке плоскости, где ее знаменатель не обращается в нуль.

Пример. Функция $\frac{z^2+1}{z^4+1}$ аналитична во всех точках плоскости, кроме четырех точек — корней уравнения $z^4+1=0$. Этими точками являются: $z_1=\frac{\sqrt{2}}{2}(1+i)$, $z_2=$ $=\frac{\sqrt{2}}{2}(1-i)$, $z_3=\frac{\sqrt{2}}{2}(-1+i)$, $z_4=\frac{\sqrt{2}}{2}(-1-i)$ (см. пример в п. 8 § 61).

3. Аналитичность функции в области. Теорема единственности

Определение 2. Функция $f(z)$, определенная в области G, называется *аналитической в этой области,* если она аналитична в каждой точке $z\in G$.

Теорема (теорема единственности). *Пусть функции $f(z)$ и $g(z)$ аналитичны в некоторой области G на плоскости и $f(p_k)=g(p_k)$ $(k=1, 2, ...)$, где p_1, p_2, p_3, ... — бесконечная последовательность точек из G, сходящаяся к точке $p\in G$, причем все p_k отличны от p. Тогда $f(z)=g(z)$ для всех $z\in G$.*

Доказательство теоремы мы опускаем.

Следствие. *Если L — некоторая линия, лежащая в области G, и равенство $f(z)=g(z)$, где f и g — две аналитические функции в G, имеет место во всех точках этой линии, то $f(z)=g(z)$ для всех $z\in G$.*

Иначе говоря, *не существует двух разных аналитических функций в области G, совпадающих вдоль линии $L\subset G$*, т. е. *аналитическая функция в области G однозначно определяется своими значениями на линии $L\subset G$.*

Это — весьма важное свойство аналитических функций.

4. Аналитическое продолжение функций

Пусть на множестве $A\subset \mathbf{C}$ задана функция f комплексной переменной. Рассмотрим какую-либо область G, содержащую A. *Аналитическим продолжением функции* f на область G называется функция F, определенная и аналитическая в области G и совпадающая на множестве A с заданной функцией f.

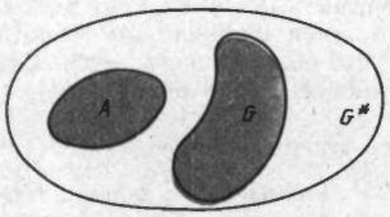


Рис. 233

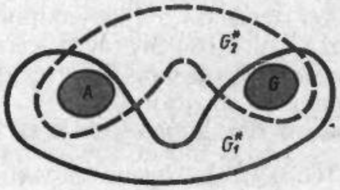


Рис. 234

Пример. Функция $f(z)=1+z+z^2+z^3+...$ определена в круге $A:|z|<1$. Ее можно аналитически продолжить на область G, состоящую из всех точек z, $z\neq 1$. Продолженная функция $F(z)$ есть $\frac{1}{1-z}$.

Может оказаться, что функция f допускает не одно, а несколько аналитических продолжений с множества A на область G. Однако при определенных предположениях относительно A такое продолжение возможно лишь единственным образом. Согласно п. 3, это, например, имеет место, если множество A представляет собой некоторую линию. Этим соображением мы будем часто пользоваться в дальнейшем.

Возможен и другой подход к понятию аналитического продолжения. Пусть на множестве A определена функция f_A. Рассмотрим некоторую область G. Если $G\supset A$, то, как было отмечено в п. 3, аналитическое продолжение функции f_A на область G возможно лишь единственным образом. Однако может оказаться, что A не содержится целиком в G (или даже не пересекается с G). В этом случае функцию f_G, определенную на G, будем называть *формальным аналитическим продолжением* функции f_A, если существуют такая область G^*, содержащая G и A (рис. 233), и такая аналитическая функция f_{G^*} в области G^*, что $f_{G^*}=f_G$ на множестве G и $f_{G^*}=f_A$ на множестве A.

При таком подходе к понятию аналитического продолжения становится возможным своеобразное явление, называемое *ветвлением*. Множества A и G можно охватить, вообще говоря, с помощью различных областей G^* (на рис. 234 показаны две такие области G^*_1 и G^*_2). Возможны случаи, когда аналитические функции f_G, полученные формальным продолжением одной и той же функции f_A с помощью различных областей G^*, не совпадают, иначе говоря, когда функцию f_A можно продолжить формально аналитически на область G неединственным образом. Полученные таким путем различные функции f_G являются как бы разными «ветвями» продолжения для f_A.

§ 67. Элементарные аналитические функции комплексной переменной

1. Отображение, задаваемое функцией. Область однолистности

Пусть функция $w=f(z)$ комплексной переменной z задана в некоторой области G. Положим, как обычно, $z=x+iy$, $w=u+iv$. Наряду с координатной плоскостью xy (плоскостью переменной z) рассмотрим координатную плоскость uv (плоскость переменной w; рис. 235). Функциональная зависимость $\omega=f(z)$ означает, что задано отображение $z\mapsto w$ области G на некоторое множество точек в плоскости переменной w.

Отображение, задаваемое функцией $w=f(z)$, не обязательно является взаимно однозначным: может оказаться, что для двух различных точек z_1, z_2 области G справедливо равенство $f(z_1)=f(z_2)$. Если $f(z_1)\neq f(z_2)$ при всех $z_1\neq z_2$ (где $z_1\in G$, $z_2\in G$), то функция $f(z)$ называется *однолистной* (обратимой) *в области* G; в этом случае также говорят, что область G является *областью однолистности* функции $f(z)$. Например, функция $w=2z$ однолистна на всей комплексной плоскости $\mathbb{C}$.

Приведем без доказательства следующее утверждение.

Теорема 1. *Пусть функция $w=f(z)$ аналитична и однолистна в области G и пусть W — множество всех значений w, когда z пробегает G. Тогда W также является областью (в плоскости переменной w), причем обратная функция $z=g(w)$ аналитична в области W.*

В случае однолистной аналитической функции $w=f(z)$ отображение $z\mapsto w$ устанавливает взаимно однозначное соответствие между двумя областями: областью G, на которой задана функция, и областью W всех значений функции.

Из определения однолистности следует, что если G — область однолистности функции f, то любая область, являющаяся частью G, также является областью однолистности для f. В связи с этим представляют особый интерес такие области однолистности, которые являются максимальными, — в том смысле, что их нельзя расширить с сохранением свойства однолистности. Для каждой из рассматриваемых ниже элементарных функций будут указаны такие области.

2. Линейная функция

Функция

$$w=az+b, \tag{1}$$

где a и b — фиксированные комплексные числа, причем $a\neq 0$, называется *линейной*. Очевидно, что она аналитична на всей комплексной плоскости.

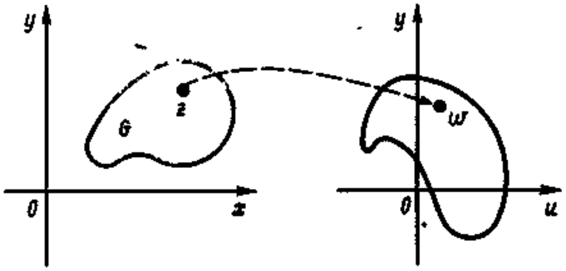


Рис. 235

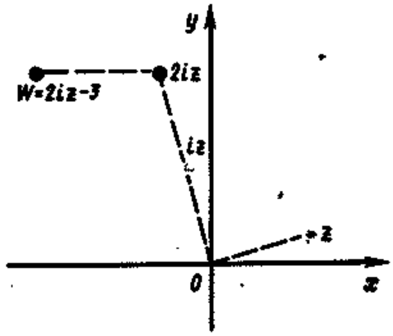


Рис. 236

Линейная функция осуществляет взаимно однозначное отображение плоскости переменной z на плоскость переменной w: действительно, из равенства (1) при данном значении z находится единственное значение w, а при данном w — единственное z. Поэтому любая область является областью однолистности для функции (1); максимальной же областью однолистности служит вся плоскость.

Геометрический смысл отображения, задаваемого функцией (1), можно установить на основании сказанного в п. 5 § 61. Отображение $z \mapsto az$ есть композиция двух отображений: поворота вокруг начала на угол, равный $\arg a$, и гомотетии с центром в начале координат и коэффициентом гомотетии $|a|$. Следовательно, отображение $z \mapsto az + b$ является композицией трех отображений: двух указанных выше и параллельного переноса на вектор, соответствующий числу b.

На рис. 236 показана точка z и ее образ w для случая $w = 2iz - 3$; обе точки отнесены к одной и той же плоскости и одной и той же системе координат.

3. Функция $w = 1/z$

Функцию $w = 1/z$ удобно рассматривать на расширенной комплексной плоскости $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. В этом случае отображение, задаваемое функцией, является взаимно однозначным отображением $\hat{\mathbb{C}}$ на себя (точке $z = \infty$ отвечает $w = 0$, а точке $z = 0$ — точка $w = \infty$).

Ранее было показано (см. пример 2 п. 1 § 66), что функция $w = 1/z$ аналитична в любой точке $z \neq 0$.

Важным свойством функции $w = 1/z$ является то, что *задаваемое ею отображение переводит любую прямую плоскости в прямую или окружность; любая окружность также переводится в прямую или окружность.*

□ Запишем сначала уравнения произвольной прямой и произвольной окружности в комплексной форме.

Прямая на плоскости задается уравнением вида

$$Ax + By + C = 0, \tag{2}$$

где A, B, C — фиксированные действительные числа, причем $A^2 + B^2 \neq 0$. Положим $z = x + iy$; тогда $\bar{z} = x - iy$ и $x = \frac{1}{2}(z + \bar{z})$, $y = \frac{1}{2i}(z - \bar{z})$. Подставляя эти выражения для x и y в уравнение (2), получим

$$\frac{1}{2}A(z + \bar{z}) + \frac{1}{2i}B(z - \bar{z}) + C = 0,$$

откуда, группируя вместе члены с z, а также с $\bar z$, имеем

$$\left(\frac{1}{2}A+\frac{1}{2i}B\right)z+\left(\frac{1}{2}A-\frac{1}{2i}B\right)\bar z+C=0.$$

Полагая $\frac{1}{2}A+\frac{1}{2i}B=a$, приходим к уравнению вида

$$az+\bar a\bar z+c=0, \tag{3}$$

где $c\in\mathbb{R}$.

Окружность на плоскости задается уравнением вида

$$x^2+y^2+Ax+By+C=0,$$

т. е.

$$z\bar z+az+\bar a\bar z+c=0, \tag{4}$$

где $c\in\mathbb{R}$.

Применим теперь отображение $w=1/z$ к прямой (3). Тогда на плоскости переменной w получим линию

$$a\frac{1}{w}+\bar a\frac{1}{\bar w}+c=0 \text{ или } cw\bar w+a\bar w+\bar a w=0 \quad (c\in\mathbb{R}).$$

При $c=0$ имеем уравнение прямой (проходящей через начало); при $c\neq0$ уравнение приводится к виду $w\bar w+b\bar w+\bar b w=0$ и определяет окружность (проходящую через начало).

Если применить отображение $w=1/z$ к окружности (4), то получим линию

$$\frac{1}{w}\cdot\frac{1}{\bar w}+\frac{a}{w}+\frac{\bar a}{\bar w}+c=0 \text{ или } cw\bar w+a\bar w+\bar a w+1=0,$$

т. е. прямую или окружность. ■

Пример. Выяснить, во что переходит при отображении $w=1/z$ прямая $2x-3y+5=0$.

○ Запишем уравнение прямой в комплексной форме:

$$2\frac{1}{2}(z+\bar z)-3\frac{1}{2i}(z-\bar z)+5=0 \text{ или } \left(1+\frac{3}{2}i\right)z+\left(1-\frac{3}{2}i\right)\bar z+5=0.$$

При отображении $w=1/z$ эта линия переходит в линию

$$\left(1+\frac{3}{2}i\right)\frac{1}{w}+\left(1-\frac{3}{2}i\right)\frac{1}{\bar w}+5=0.$$

Таким образом, для w получаем уравнение

$$5w\bar w+\left(1+\frac{3}{2}i\right)\bar w+\left(1-\frac{3}{2}i\right)w=0.$$

После перехода к действительным переменным u и v, где $u+iv=w$, оно принимает вид

$$5(u^2+v^2)+\left(1+\frac{3}{2}i\right)(u-iv)+\left(1-\frac{3}{2}i\right)(u+iv)=0$$

или, после упрощений,

$$u^2+v^2+\frac{2}{5}u+\frac{3}{5}v=0.$$

Записав полученное уравнение в виде

$$\left(u+\frac{1}{5}\right)^2+\left(v+\frac{3}{10}\right)^2=\frac{13}{100},$$

мы видим, что оно определяет окружность с центром $(-1/5;\ -3/10)$ и радиусом $\sqrt{13}/10$. ●

4. Функция $w=\frac{az+b}{cz+d}$

Функция

$$w=\frac{az+b}{cz+d}, \tag{5}$$

где a, b, c, d — фиксированные комплексные числа, причем $ad-bc\neq 0$, называется *дробно-линейной*. Условие $ad-bc\neq 0$ вызвано тем, что в случае $ad-bc=0$ имеем $a/c=b/d$; тогда функция w постоянна (тождественно равна a/c).

Если $c=0$, то функция (5) является, очевидно, линейной. Отбросим пока этот случай, т. е. будем считать $c\neq 0$. Тогда можем записать

$$w=\frac{\frac{a}{c}(cz+d)+\left(b-\frac{a}{c}d\right)}{cz+d}=\frac{a}{c}+\frac{p}{cz+d}, \tag{6}$$

где $p=b-\frac{a}{c}d$.

Из формулы (6) вытекают такие свойства.

1^0. *Дробно-линейная функция аналитична во всех точках z, где ее знаменатель $cz+d$ не обращается в нуль.*

Действительно, аналитичность функции $\frac{1}{(z+q)}$ при $z\neq -q$ была доказана ранее (см. пример 2 п. 1 § 66); отсюда вытекает аналитичность функции $\frac{1}{(cz+d)}$ при $z=-\frac{d}{c}$, а значит, и аналитичность функции в правой части равенства (6).

2^0. *Дробно-линейное отображение переводит прямую и окружность снова в одну из этих линий (при этом прямая может отображаться в окружность или, наоборот, окружность в прямую).*

В самом деле, как следует из равенства (6), отображение $z\mapsto w$ является композицией четырех отображений: $z\mapsto t=cz+d$, затем $t\mapsto s=\frac{1}{t}$, далее $s\mapsto h=ps$ и, наконец, $h\mapsto w=\frac{a}{c}+h$. Три из названных четырех отображений являются линейными, а одно ($t\mapsto s$) есть отображение, рассмотренное в п. 3. Отсюда и вытекает указанное свойство отображения (6).

Сформулируем еще одно свойство дробно-линейных отображений.

3^0. *Пусть z_1, z_2, z_3 — любая тройка попарно различных точек, а w_1, w_2, w_3 — другая тройка попарно различных точек. Тогда существует дробно-линейное отображение, и притом единственное, при котором*

$$z_1\mapsto w_1,\ z_2\mapsto w_2,\ z_3\mapsto w_3. \tag{7}$$

Добавим, что при этом одно из чисел z_1, z_2, z_3, а также одно из чисел w_1, w_2, w_3 может быть равно ∞.

Приведем набросок доказательства. Рассмотрим соотношение

$$\frac{w-w_1}{w-w_2}:\frac{w_3-w_1}{w_3-w_2}=\frac{z-z_1}{z-z_2}:\frac{z_3-z_1}{z_3-z_2}. \tag{8}$$

Оно приводится к виду

$$\frac{w-w_1}{w-w_2}=k\,\frac{z-z_1}{z-z_2};$$

если выразить отсюда w через z, то получим дробно-линейную зависимость $w=\frac{az+b}{cz+d}$. С другой стороны, из соотношения (8) нетрудно вывести, что при $z=z_1$ должно быть выполнено равенство $w=w_1$, при $z=z_2$ — равенство $w=w_2$ и, наконец, при $z=z_3$ — равенство $w=w_3$.

Единственность дробно-линейного отображения, удовлетворяющего условию (7), доказывать не будем.

Пример. Записать дробно-линейное отображение, при котором $0\mapsto -1$, $i\mapsto 0$, $1\mapsto -i$.

○ В данном случае соотношение (8) принимает вид

$$\frac{w-(-1)}{w-0}:\frac{-i-(-1)}{-i-0}=\frac{z-0}{z-i}:\frac{1-0}{1-i},$$

или, после упрощений,

$$\frac{w+1}{w}=\frac{z}{z-i}\cdot 2.$$

Отсюда находим искомое отображение: $w=\frac{z-i}{z+i}$. ●

5. Степенная функция с натуральным показателем

Рассмотрим функцию $w=z^n$, где n — фиксированное натуральное число. Как всякий многочлен от z, функция z^n аналитична в любой точке. При $z=0$ имеем, очевидно, $w=0$. Пусть теперь $z\neq 0$. Представим число z в тригонометрической форме: $z=r(\cos\varphi+i\sin\varphi)$. Используя формулу Муавра, получим $z^n=r^n(\cos\varphi+i\sin n\varphi)$, т. е. число z^n имеет модуль r^n и аргумент $n\varphi$.

Итак, отображение, задаваемое функцией $w=z^n$, геометрически означает следующее: если векторы $\overrightarrow{OA}$ и $\overrightarrow{OB}$ изображают соответственно числа z и w, то вектор $\overrightarrow{OB}$ образует с действительной осью угол, в n раз больший, чем вектор $\overrightarrow{OA}$, а длина вектора $\overrightarrow{OB}$ равна длине вектора $\overrightarrow{OA}$, возведенной в n-ю степень (рис. 237).

Из сказанного следует, что область G, определяемая условиями $0<\operatorname{Arg} z<2\pi/n$ (заштрихованный угол на рис. 238), является областью однолистности функции $f(z)=z^n$. В самом деле, пусть z_1 и z_2 — две различные точки в области G; r_1 и φ_1 — модуль и аргумент числа z_1, а r_2 и φ_2 — модуль и аргумент z_2. Так как $z_1\in G$ и $z_2\in G$, то

$$0<\varphi_1<2\pi/n,\ 0<\varphi_2<2\pi/n. \tag{9}$$

Условие $z_1\neq z_2$ означает, что $r_1\neq r_2$ или $\varphi_1\neq\varphi_2$. В обоих случаях имеем $z_1^n\neq z_2^n$. Действительно, если $r_1\neq r_2$, то $r_1^n\neq r_2^n$, а значит, $z_1^n\neq z_2^n$; если же $\varphi_1\neq\varphi_2$, например, если $\varphi_2>\varphi_1$, то в силу неравенств (9) имеем $0<n\varphi_2-n\varphi_1<2\pi$, что снова влечет за собой $z_1^n\neq z_2^n$.

Заметим, что область G как область однолистности функции $f(z)=z^n$ максимальна: при любом расширении G в ней обязательно должны оказаться две точки z_1 и z_2 такие, что $r_2=r_1$, $\varphi_2-\varphi_1=2\pi/n$, откуда будет следовать $z_1^n=z_2^n$. Разумеется, любая часть области G также является областью однолистности.

Функция $w=z^n$ отображает область G на область W, определяемую условиями $0<\operatorname{Arg} w<2\pi$; при этом граница области G, состоящая из двух лучей $\operatorname{Arg} z=0$ и $\operatorname{Arg} z=2\pi/n$, отображается в один луч $\operatorname{Arg} z=0$, т. е. в границу области W (условие $\operatorname{Arg} w=2\pi$ дает тот же луч, что и $\operatorname{Arg} w=0$).

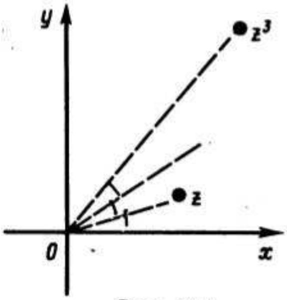


Рис. 237

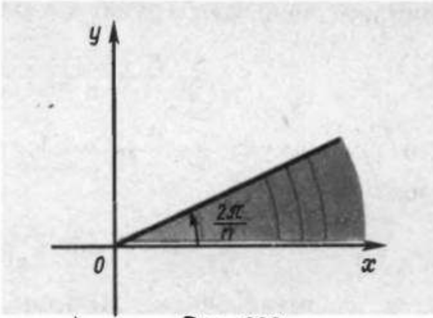


Рис. 238

На рис. 238 показаны заключенные внутри G дуги окружностей с центром в начале координат, т. е. окружностей $|z| = r$, где $r = \text{const} > 0$; им соответствуют окружности в W, из которых выколоты точки, лежащие на положительном луче Ox действительной оси.

6. Функция e^z

Как мы уже знаем, функция e^x при действительных значениях x разлагается в степенной ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

Это разложение подсказывает путь для введения функции e^z комплексной переменной z.

Положим

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \qquad (10)$$

для комплексных значений z. Поскольку написанный ряд сходится при всех значениях z (см. пример п. 3 § 65), функция e^z определена для всех z.

Следующая теорема устанавливает основное свойство функции e^z.

Теорема 2. *Для любых комплексных s и t справедливо равенство* $e^s e^t = e^{s+t}$.

□ Имеем

$$e^s e^t = \left(1 + s + \frac{s^2}{2!} + \frac{s^3}{3!} + \cdots\right)\left(1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right).$$

Перепишем это равенство следующим образом:

$$e^s e^t = \left(\frac{s^0}{0!} + \frac{s^1}{1!} + \frac{s^2}{2!} + \frac{s^3}{3!} + \cdots\right)\left(\frac{t^0}{0!} + \frac{t^1}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right).$$

Так как оба ряда, заключенные в скобки, сходятся абсолютно, то их можно перемножить по обычному правилу умножения суммы на сумму; при этом безразлично, в каком порядке будут записаны произведения членов первого ряда на члены второго. Расположим их так:

$$\frac{s^0}{0!}\frac{t^0}{0!} + \left(\frac{s^0}{0!}\frac{t^1}{1!} + \frac{s^1}{1!}\frac{t^0}{0!}\right) + \left(\frac{s^0}{0!}\frac{t^2}{2!} + \frac{s^1}{1!}\frac{t^1}{1!} + \frac{s^2}{2!}\frac{t^0}{0!}\right) + \cdots,$$

т. е. сначала сгруппируем произведения, в которых сумма показателей степени (при s и t) равна 0, затем произведения с суммой показателей 1, 2 и т. д.

Рассмотрим, например, группу из членов четвертой степени:

$$\frac{s^0}{0!}\frac{t^4}{4!}+\frac{s^1}{1!}\frac{t^3}{3!}+\frac{s^2}{2!}\frac{t^2}{2!}+\frac{s^3}{3!}\frac{t^1}{1!}+\frac{s^4}{4!}\frac{t^0}{0!}.$$

Используя формулу $\frac{n!}{k!(n-k)!}=C_n^k$, перепишем это выражение следующим образом:

$$\frac{1}{4!}(C_4^0s^0t^4+C_4^1s^1t^3+C_4^2s^2t^2+C_4^3s^3t^1+C_4^4s^4t^0).$$

Согласно формуле бинома Ньютона, последнее выражение равно $\frac{1}{4!}(s+t)^4$. Подобным же образом может быть преобразована группа членов любой степени n $(n=0, 1, 2, \ldots)$. В итоге получаем

$$e^se^t=\frac{(s+t)^0}{0!}+\frac{(s+t)^1}{1!}+\frac{(s+t)^2}{2!}+\cdots=e^{s+t}. \blacksquare$$

Почленно дифференцируя ряд (10), находим $(e^z)'=e^z$, т. е. производная функции e^z совпадает с самой функцией (как и в действительной области).

Другие свойства функции e^z будут рассмотрены в пп. 8—11.

7. Функции sin z и cos z

Функции $\sin z$ и $\cos z$ определяются тем же способом, что и e^z. Исходя из уже известных разложений $\sin x$ и $\cos x$ в степенные ряды в действительной области, положим

$$\sin z=z-\frac{z^3}{3!}+\frac{z^5}{5!}-\frac{z^7}{7!}+\cdots,$$

$$\cos z=1-\frac{z^2}{2!}+\frac{z^4}{4!}-\frac{z^6}{6!}+\cdots$$

для комплексных значений z. Указанные степенные ряды сходятся при всех $z\in\mathbf{C}$ (что следует, например, из признака Даламбера) и тем самым функции $\sin z$ и $\cos z$ определены на всей комплексной плоскости.

Непосредственно из определения функций $\sin z$ и $\cos z$ вытекают равенства $\sin(-z)=-\sin z$, $\cos(-z)=\cos z$.

Почленным дифференцированием рядов для $\sin z$ и $\cos z$ получаем формулы $(\sin z)'=\cos z$, $(\cos z)'=-\sin z$.

8. Формула Эйлера

Функции e^z, $\sin z$, $\cos z$ связаны замечательным соотношением

$$e^{iz}=\cos z+i\sin z, \tag{11}$$

которое называют *формулой Эйлера*. Эта формула справедлива для любого $z\in\mathbf{C}$.

Действительно, по определению функции e^z имеем

$$e^{iz}=1+(iz)+\frac{(iz)^2}{2!}+\frac{(iz)^3}{3!}+\cdots=\left(1-\frac{z^2}{2!}+\frac{z^4}{4!}-\cdots\right)+i\left(z-\frac{z^3}{3!}+\frac{z^5}{5!}-\cdots\right)=$$
$$=\cos z+i\sin z.$$

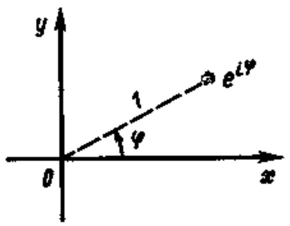


Рис. 239

Отметим частный случай, когда z — действительное число. Тогда формула (11) примет вид

$$e^{i\varphi} = \cos\varphi + i\sin\varphi,$$

где $\varphi \in \mathbf{R}$. Отсюда следует, что $e^{i\varphi}$ *есть комплексное число с модулем* 1 *и аргументом* φ (рис. 239).

9. Некоторые применения формулы Эйлера

Пусть $x+iy$ — произвольное комплексное число. Имеем $e^{x+iy} = e^{x}y^{iy}$, откуда с помощью формулы Эйлера получаем

$$e^{x+iy} = e^{x}(\cos y + i\sin y). \tag{12}$$

Используя эту формулу, можно непосредственно находить e^{z}.

Например, $e^{2+3i} = e^{2}(\cos 3 + i\sin 3) \approx -e^{2}$, поскольку $\cos 3 \approx \cos\pi = -1$, $\sin 3 \approx \sin\pi = 0$.

С помощью формулы Эйлера можно получить более удобные выражения для функций $\sin z$ и $\cos z$:

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \tag{13}$$

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}. \tag{14}$$

Чтобы установить эти формулы, достаточно записать равенства

$$e^{iz} = \cos z + i\sin z,\ \ e^{-iz} = \cos(-z) + i\sin(-z);$$

складывая эти равенства, а также вычитая из первого второе, приходим к формулам (13) и (14).

В частности, при $a \in \mathbf{R}$ имеем

$$\cos ia = \frac{e^{i(ia)} + e^{-i(ia)}}{2} = \frac{e^{-a} + e^{a}}{2}.$$

Отсюда видно, что $\cos z$ может принимать как угодно большие действительные значения. Например, $\cos 3i = (e^{-3} + e^{3})/2 = 13, \ldots$. Напомним, что в действительной области $\cos x$ всегда заключен в пределах от -1 до $+1$.

Исходя из полученных выражений для $\sin z$ и $\cos z$, можно доказать формулы сложения в комплексной области:

$$\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2,$$

$$\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2,$$

хорошо знакомые из тригонометрии (в случае действительных z_1 и z_2).

10. Аналитичность функций e^z, $\sin z$, $\cos z$

Из самого определения функции e^z как ряда (10) следует аналитичность функции в точке $z=0$. Используя формулу (10), легко доказать, что e^z аналитична в произвольной точке z_0. Действительно, полагая $e^{z_0}=c$, из равенства $e^z=e^{z_0}e^{z-z_0}$, находим

$$e^z=c\left(1+(z-z_0)+\frac{(z-z_0)^2}{2!}+\cdots\right)=c+c(z-z_0)+c\frac{(z-z_0)^2}{2!}+\cdots,$$

т. е. функция e^z разлагается в сходящийся всюду степенной ряд с центром z_0. Это и доказывает аналитичность e^z в точке z_0.

Функция e^{az} ($a=\text{const}\in\mathbb{C}$), как нетрудно установить, также является аналитической в произвольной точке z_0; отсюда и из формул (13), (14) вытекает аналитичность функций $\cos z$ и $\sin z$ на всей комплексной плоскости.

Поскольку при действительных значениях z функции e^z, $\sin z$, $\cos z$ обращаются в функции e^x, $\sin x$, $\cos x$ действительной переменной x, можно сказать, что *каждая из функций* e^z, $\sin z$, $\cos z$ *является аналитическим продолжением соответствующей функции* e^x, $\sin x$, $\cos x$, *где* $x\in\mathbb{R}$, *на всю комплексную плоскость.* Как известно (см. п. 3 § 66), такое продолжение возможно лишь единственным образом.

Используя соображения аналитичности, для функций $\sin z$ и $\cos z$ в комплексной области можно доказать обычные формулы приведения:

$$\sin\left(z+\frac{\pi}{2}\right)=\cos z,\quad \cos\left(z+\frac{\pi}{2}\right)=-\sin z,$$

$$\sin(z+\pi)=-\sin z,\quad \cos(z+\pi)=-\cos z.$$

Докажем, например, первую из них. При действительных значениях x имеем $\sin\left(x+\frac{\pi}{2}\right)=\cos x$. Это показывает, что две аналитические (во всей плоскости) функции комплексной переменной z, а именно $\sin\left(z+\frac{\pi}{2}\right)$ и $\cos z$, принимают одинаковые значения в точках действительной оси. Отсюда в силу теоремы единственности вытекает совпадение этих функций.

11. Периоды функций e^z, $\sin z$, $\cos z$

Из формулы Эйлера, в частности, следует, что

$$e^{2\pi i}=\cos 2\pi+i\sin 2\pi=1+0i=1.$$

Отсюда при любом z имеем

$$e^{z+2\pi i}=e^z e^{2\pi i}=e^z.$$

Мы получили весьма неожиданное свойство функции e^z: она периодична с периодом $2\pi i$. Заметим, что функция e^x при $x\in\mathbb{R}$ не является периодической; появление периода связано с выходом в комплексную область.

Ясно, что любое из чисел $2\pi ki$, где $k\in\mathbb{Z}$, также является периодом функции e^z. Убедимся, что других периодов функция не имеет.

В самом деле, если T — какой-нибудь период для e^z, то при любом z выполняется равенство $e^{z+T}=e^z$. В частности, при $z=0$ имеем $e^T=1$. Полагая $T=t_1+it_2$, отсюда находим

$$e^{t_1}(\cos t_2+i\sin t_2)=1,$$

т. е. $t_1=0$, $t_2=2\pi k$, где $k\in\mathbb{Z}$. Следовательно, $T=2\pi ki$.

Итак, *функция e^z периодична с периодом $2\pi ki$, где $k \in \mathbf{Z}$. Наименьший по модулю период для e^z есть $2\pi i$ или $-2\pi i$.*

Отметим также, что если $z_1 - z_2 \neq 2\pi ki$, где $k \in \mathbf{Z}$ то $e^{z_1} \neq e^{z_2}$.

Что касается функций $\sin z$ и $\cos z$, то их периодом является, как и в действительной области, число 2π. Действительно, по формуле (13) имеем

$$\cos(z + 2\pi) = \frac{e^{i(z+2\pi)} + e^{-i(z+2\pi)}}{2} = \frac{e^{iz} + e^{-iz}}{2} = \cos z$$

и аналогично для $\sin z$. Любой из других периодов для $\sin z$ и $\cos z$ равен $2\pi k$, где $k \in \mathbf{Z}$. В самом деле, пусть число T является, например, периодом для $\cos z$. Тогда в силу соотношения $\sin z = -\cos(z + \pi/2)$ число T есть период и для $\sin z$; значит, как показывает формула Эйлера (11), число iT есть период для функции e^z. Отсюда следует, что $T = 2\pi k$, где $k \in \mathbf{Z}$.

12. Область однолистности для функции e^z

Рассмотрим область (полосу) G, определяемую условиями

$$-\pi < \mathrm{Im}(z) < \pi \tag{15}$$

(рис. 240). Очевидно, в ней нет точек z_1, z_2 таких, что $z_1 \neq z_2$, $z_1 - z_2 = 2\pi ki$, $k \in \mathbf{Z}$. Следовательно, для любых двух различных точек z_1, z_2 области G имеем $e^{z_1} \neq e^{z_2}$. Это означает, что *область G является областью однолистности функции e^z*.

На рис. 240 показан вертикальный отрезок $z = x_0 + iy$, где $-\pi < y < \pi$, в области G. Образом этого отрезка при отображении $z \mapsto w = e^z$ является, как

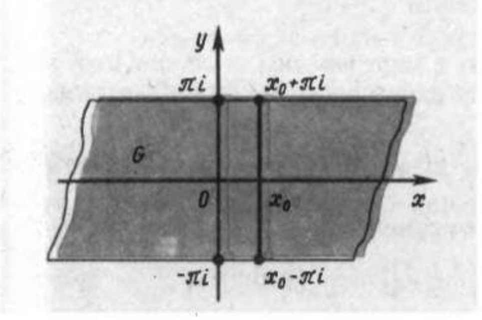


Рис. 240

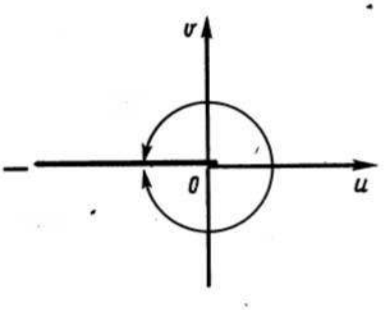


Рис. 241

показывает формула (12), множество точек $w = e^{x_0}(\cos y + i \sin y)$, где $-\pi < y < \pi$, т. е. окружность (в плоскости переменной w) с центром в начале координат и радиусом e^{x_0}, причем из этой окружности исключена точка

$$w_0 = e^{x_0}(\cos\pi + i\sin\pi) = e^{x_0}(\cos(-\pi) + i\sin(-\pi)) = -e^{x_0}.$$

Образом же всей области G является область W, которая получается из плоскости w удалением луча $\mathbf{R}$, состоящего из точек $u + 0i$, где

$$u \in \mathbf{R},\ u \leqslant 0 \tag{16}$$

(рис. 241). Заметим, что область G как область однолистности функции e^z является максимальной.

13. Функция $\ln z$

Рассмотрим снова области G и W из предыдущего пункта. Отображение G на W, устанавливаемое с помощью функции $w=e^z$, является, как было сказано, взаимно однозначным. Согласно теореме 1, обратная функция $z=g(w)$, определенная в области W, аналитична в W. Обозначим эту обратную функцию $\ln w$. Итак, *$z=\ln w$ есть функция, обратная к $w=e^z$, при условии, что последняя функция рассматривается в области G, определяемой условиями* (15). *Функция $\ln w$ аналитична в области $W=C\setminus\mathbf{R}_-$, где $\mathbf{R}_-$ есть луч, определяемый условиями* (16).

При $z\in\mathbf{R}$ функция e^z совпадает с действительной функцией e^x; область определения последней есть $\mathbf{R}_+$. Поэтому можно сказать, что *функция $\ln w$ является аналитическим продолжением функции $\ln u$, где $u\in\mathbf{R}_+$, на область W.*

Укажем простое аналитическое выражение для $\ln w$. Пусть

$$w=re^{i\varphi}, \tag{17}$$

где $r>0$; будем считать, что угол φ заключен в пределах $-\pi<\varphi<\pi$ (это соответствует $w\in W$). Переписав равенство (17) в виде

$$w=e^{\ln r}e^{i\varphi}=e^{\ln r+i\varphi},$$

получим $\ln w=\ln r+i\varphi$. Итак,

$$\ln w=\ln r+i\varphi,\ \text{где}\ w=re^{i\varphi},\ -\pi<\varphi<\pi.$$

Это и есть требуемая формула для $\ln w$.

Например,

$$\ln(1+i)=\ln\left(\sqrt{2}e^{i\frac{\pi}{4}}\right)=\ln\sqrt{2}+i\frac{\pi}{4}.$$

Подчеркнем, что согласно данному выше определению функции $\ln w$ значение логарифма не существует только для действительных отрицательных w.

§ 68. Интеграл от функции комплексной переменной

1. Интеграл по отрезку действительной оси

Пусть на отрезке $[a, b]$ задана функция, принимающая комплексные значения:

$$f(t)=u(t)+iv(t),\ \text{где}\ a\leqslant t\leqslant b.$$

Определим *интеграл от этой функции по отрезку* $[a, b]$ формулой

$$\int_a^b f(t)dt=\int_a^b u(t)dt+i\int_a^b v(t)dt.$$

Пример. Найти $\int_0^3(1+it^2)dt$.

○ $\int_0^3(1+it^2)dt=\int_0^3 dt+i\int_0^3 t^2dt=3+i\cdot 9.$ ●

Рассмотренные ранее общие свойства интегралов в действительной области остаются верными и для интегралов от комплекснозначных функций. Перечислим эти свойства:

1°. $\int_a^b \lambda f(t)dt=\lambda\int_a^b f(t)dt$, *где λ — любое комплексное число.*

2^0. $\int\limits_a^b (f_1(t)+f_2(t))dt=\int\limits_a^b f_1(t)dt+\int\limits_a^b f_2(t)dt.$

3^0. $\int\limits_a^b f(t)dt=\int\limits_a^c f(t)dt+\int\limits_c^b f(t)dt.$

4^0. *Правило подстановки:*

$$\int\limits_a^b f(t)dt=\int\limits_\alpha^\beta f(\varphi(t))\varphi'(t)dt,$$

где $\varphi(t)$ — действительная функция такая, что $\varphi(\alpha)=a$, $\varphi(\beta)=b$.

Доказательства рекомендуем провести самостоятельно в качестве упражнений.

Важное значение имеет следующее свойство:

5^0. $\left|\int\limits_a^b f(t)dt\right|\leqslant\int\limits_a^b |f(t)|dt.$

Для доказательства достаточно заметить, что если $\sum\limits_{i=1}^n f(\xi_i)\Delta t_i$ есть интегральная сумма для интеграла $\int\limits_a^b f(t)dt$, то

$$\left|\sum_{i=1}^n f(\xi_i)\Delta t_i\right|\leqslant\sum_{i=1}^n |f(\xi_i)|\Delta t_i$$

(модуль суммы нескольких чисел не превосходит суммы их модулей).

Отметим еще равенство $\int\limits_a^b f(t)dt=-\int\limits_b^a f(t)dt$, которое при $a>b$ служит определением интеграла, стоящего в левой части.

2. Интеграл по кривой

Пусть в некоторой области G задана функция $w=f(z)$ комплексной переменной. Положим, как обычно, $z=x+iy$, $w=u+iv$. Возьмем кривую Γ, лежащую в G. Будем считать эту кривую *ориентированной,* т. е. выберем на ней одно из двух возможных направлений движения. Предположим, что $x=x(t)$, $y=y(t)$ — параметрические уравнения кривой Γ, где параметр t изменяется от a до b; значению $t=a$ отвечает начальная точка кривой, значению $t=b$ — конечная.

Определение. *Интеграл от функции $w=f(z)$ по ориентированной кривой* Γ определяется равенством

$$\int\limits_\Gamma f(z)dz=\int\limits_a^b f(z(t))z'(t)dt,$$

где $z(t)=x(t)+iy(t)$ и, следовательно, $z'(t)=x'(t)+iy'(t)$.

Итак, по определению имеем

$$\int\limits_\Gamma f(z)dz=\int\limits_a^b [u(x(t),\ y(t))+iv(x(t),\ y(t)]\,(x'(t)+iy'(t))dt$$

или, в более краткой записи,

$$\int\limits_\Gamma f(z)dz=\int\limits_a^b (u+iv)(x'+iy')dt. \qquad (1)$$

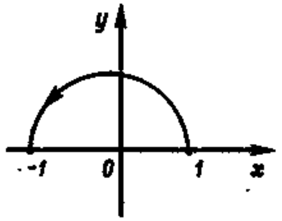


Рис. 242

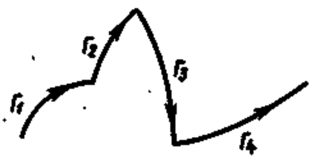


Рис. 243

Рис. 244

Примеры. 1. Вычислить $\int\limits_{\Gamma} z^2 dz$, где Γ — полуокружность $|z|=1$, $\operatorname{Im}(z)\geqslant 0$, проходимая в направлении против часовой стрелки (рис. 242).

○ Уравнение кривой Γ имеет вид $z=e^{i\varphi}$, где параметр φ изменяется от 0 до π (см. п. 8 § 67). Следовательно,

$$\int\limits_{\Gamma} z^2 dz=\int\limits_0^{\pi} e^{2it} ie^{it}dt=i\int\limits_0^{\pi} e^{3it}dt=i\int\limits_0^{\pi}(\cos 3t+i\sin 3t)dt=i\int\limits_0^{\pi}\cos 3t dt-\int\limits_0^{\pi}\sin 3t dt=-\frac{2}{3}. \bullet$$

2. Найти $\int\limits_{\Gamma}(z-a)^m dz$, где Γ — окружность $|z-a|=r$, проходимая против часовой стрелки, а m — целое число.

○ Любое комплексное число с модулем r представимо в виде $re^{i\varphi}$, где $0\leqslant\varphi\leqslant 2\pi$. Отсюда имеем $z-a=re^{i\varphi}$ или $z=a+re^{i\varphi}$, где $0\leqslant\varphi\leqslant 2\pi$. Поэтому

$$\int\limits_{\Gamma}(z-a)^m dz=\int\limits_0^{2\pi} r^m e^{im\varphi} ire^{i\varphi}d\varphi=ir^{m+1}\int\limits_0^{2\pi} e^{i(m+1)\varphi}d\varphi=ir^{m+1}\int\limits_0^{2\pi}[\cos(m+1)\varphi+i\sin(m+1)\varphi]d\varphi.$$

Если $m\neq -1$, то оба интеграла $\int\limits_0^{2\pi}\cos(m+1)\varphi d\varphi$ и $\int\limits_0^{2\pi}\sin(m+1)\varphi d\varphi$ равны нулю. Если же $m=-1$, то в качестве ответа получаем $i\int\limits_0^{2\pi}d\varphi=2\pi i$. Итак,

$$\int\limits_{\Gamma}(z-a)^m dz=0, \text{ если } m\neq -1;\quad \int\limits_{\Gamma}\frac{dz}{z-a}=2\pi i. \bullet$$

Перечислим свойства комплексного интеграла по кривой; они вытекают из соответствующих свойств интеграла по отрезку.

1^0. $\int\limits_{\Gamma}\lambda f(z)dz=\lambda\int\limits_{\Gamma}f(z)dz \quad (\lambda=\text{const}\in\mathbb{C})$.

2^0. $\int\limits_{\Gamma}[f(z)+g(z)]dz=\int\limits_{\Gamma}f(z)dz+\int\limits_{\Gamma}g(z)dz$.

3^0. *Пусть* $\Gamma=\Gamma_1+\Gamma_2+\ldots+\Gamma_n$, *т. е.* Γ *состоит из кривых* $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$, *последовательно «пристыкованных» друг к другу* (рис. 243). *Тогда*

$$\int\limits_{\Gamma}f(z)dz=\int\limits_{\Gamma_1}f(z)dz+\int\limits_{\Gamma_2}f(z)dz+\ldots+\int\limits_{\Gamma_n}f(z)dz.$$

4^0. *Если* Γ^- *означает ту же кривую, что и* Γ, *но ориентированную противоположным образом, то* $\int\limits_{\Gamma^-}f(z)dz=-\int\limits_{\Gamma}f(z)dz$.

Добавим к этим свойствам еще одно, дающее оценку модуля интеграла по кривой.

5°. *Справедливо неравенство*

$$\left|\int_\Gamma f(z)dz\right| \leqslant Ml(\Gamma), \qquad (2)$$

где M — максимум функции $|f(z)|$ на Γ, а $l(\Gamma)$ — длина Γ.

□ Неравенство (2) вытекает из следующей цепочки равенств и неравенств:

$$\left|\int_\Gamma f(z)dz\right| = \left|\int_a^b f(z(t))z'(t)dt\right| \leqslant \int_a^b |f(z(t))||z'(t)|dt \leqslant M\int_a^b |z'(t)|dt =$$

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

3. Интегральная теорема Коши

Условимся называть область *односвязной*, если она не имеет «дыр». Примером односвязной области является область, изображенная на рис. 228, а неодносвязной — область на рис. 226.

Теорема 1 (интегральная теорема Коши для односвязной области). *Пусть функция $f(z)$ непрерывно дифференцируема в области G. Пусть, далее, D есть односвязная область, которая вместе со своей границей Γ лежит внутри G* (рис. 244). *Тогда справедливо равенство $\int_\Gamma f(z)dz = 0$.*

□ Воспользуемся формулой (1), которую запишем в виде

$$\int_\Gamma f(z)dz = \int_a^b (ux' - vy')dt + i\int_a^b (uy' + vx')dt$$

или

$$\int_\Gamma f(z)dz = \int_\Gamma u(x, y)dx - v(x, y)dy + i\int_\Gamma u(x, y)dy + v(x, y)dx. \qquad (3)$$

Каждый из двух интегралов по Γ, входящих в правую часть, равен нулю. Действительно, криволинейный интеграл вида $\int_\Gamma P(x, y)dx + Q(x, y)dy$, где функции $P(x, y)$ и $Q(x, y)$ непрерывны в некоторой области G, а кривая Γ ограничивает односвязную область, лежащую (вместе с Γ) внутри G, равен нулю, если выполняется условие $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ (см. § 49). Но в данном случае указанное условие справедливо для каждого из двух интегралов, входящих в правую часть равенства (3): это вытекает из условий Даламбера — Эйлера $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$, $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, которым удовлетворяют функции $u(x, y)$ и $v(x, y)$, являющиеся действительной частью и коэффициентом при мнимой части для непрерывно дифференцируемой функции $f(z)$. Таким образом, $\int_\Gamma f(z)dz = 0 + 0 = 0$. ■

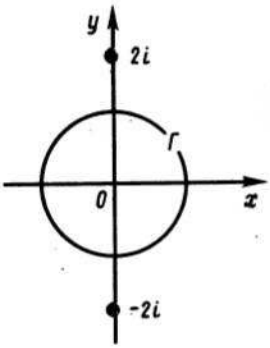


Рис. 245

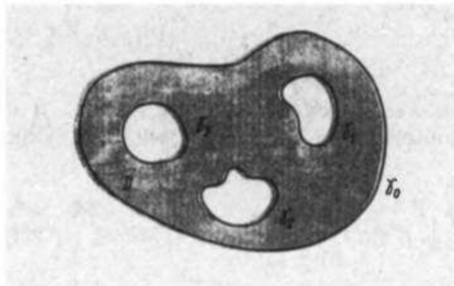


Рис. 246

Пример. Пусть Γ — окружность $|z|=1$. Интеграл $\int\limits_{\Gamma}\frac{dz}{z^2+4}$ равен нулю, так как подынтегральная функция $\frac{1}{z^2+4}$ непрерывно дифференцируема в области $|z|\neq 2i$, $z\neq -2i$, содержащей внутри себя круг $|z|<1$ вместе с его границей Γ (рис. 245).

Пусть теперь область D не является односвязной, а получается из некоторой односвязной области выкидыванием нескольких (также односвязных) частей; как говорят в подобных случаях, область D имеет «дыры» (рис. 246). Обозначим через γ_0 внешнюю границу области D и через $\gamma_1, \ldots, \gamma_n$ — внутренние границы, т. е. границы «дыр».

Теорема 2 (интегральная теорема Коши для неодносвязной области). *Пусть функция $f(z)$ непрерывно дифференцируема в некоторой области G и пусть область D вместе со своей границей лежит внутри G. Пусть, далее, γ_0 — внешняя граница D, а $\gamma_1, \ldots, \gamma_n$ — внутренние границы. Тогда справедлива формула*

$$\int\limits_{\gamma_0} f(z)dz=\int\limits_{\gamma_1} f(z)dz+\ldots+\int\limits_{\gamma_n} f(z)dz,$$

где контуры $\gamma_0, \gamma_1, \ldots, \gamma_n$ ориентированы в одном и том же направлении.

□ На рис. 247 изображена область D с одной «дырой»; граница последней обозначена γ_1. Превратим область D в односвязную, проведя разрез γ_{01}, соединяющий γ_0 с γ_1. Граница полученной односвязной области состоит из γ_0, γ_1 и кривой γ_{01}, проходимой дважды (в противоположных направлениях). Согласно теореме 1, интеграл от $f(z)$ по границе односвязной области должен быть равен нулю, т. е.

$$\int\limits_{\gamma_1}+\int\limits_{\gamma_{01}^{+}}+\int\limits_{\gamma_0^{-}}+\int\limits_{\gamma_{01}^{-}}=0$$

(подынтегральное выражение $f(z)dz$ не пишем). Но второй и четвертый интегралы взаимно уничтожаются, третий же интеграл равен $-\int\limits_{\gamma_0}$. В результате приходим к равенству $\int\limits_{\gamma_0}=\int\limits_{\gamma_1}$.

В случае, когда D имеет две «дыры» с границами γ_1 и γ_2 соответственно, проведем разрезы γ_{01} и γ_{02} (рис. 248). Такое же рассуждение, как и выше,

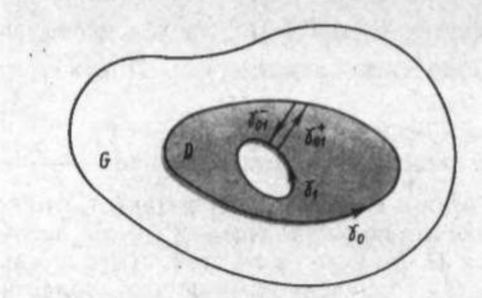


Рис. 247

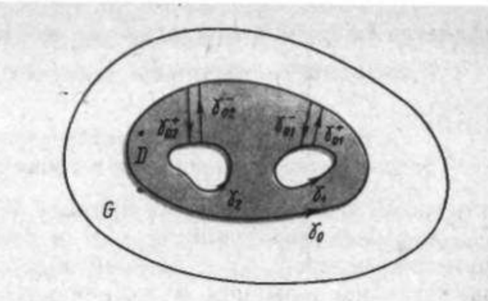


Рис. 248

приводит к равенству $\int\limits_{\gamma_0}=\int\limits_{\gamma_1}+\int\limits_{\gamma_2}$. Аналогично рассуждаем в случае любого количества «дыр». ■

Следствие. *Пусть в области, где функция $f(z)$ непрерывно дифференцируема, содержится кольцеобразная область D, ограниченная контурами γ_0 и γ_1. Тогда $\int\limits_{\gamma_0} f(z)dz=\int\limits_{\gamma_1} f(z)dz$, где контуры γ_0 и γ_1 ориентированы одинаковым образом.*

4. Интегральная формула Коши

Непрерывно дифференцируемые функции комплексной переменной обладают замечательным свойством, заключающимся в том, что значения функции на каком-либо контуре Γ однозначно определяют ее значения в точках, находящихся внутри Γ. Это свойство существенно облегчает изучение таких функций.

Пусть контур Γ ограничивает область D. Напомним, что направление обхода на Γ называется положительным, если при этом обходе область D остается слева.

Теорема 3. *Пусть функция $f(z)$ непрерывно дифференцируема в области G. Если D — односвязная область, лежащая вместе со своей границей Γ внутри области G, то для любой точки $a\in D$ справедливо равенство*

$$f(a)=\frac{1}{2\pi i}\int\limits_{\Gamma}\frac{f(z)}{z-a}dz, \qquad (4)$$

где контур Γ ориентирован положительно относительно D.

□ Рассмотрим функцию $\varphi(z)=\frac{f(z)-f(a)}{z-a}$. Будучи произведением двух непрерывно дифференцируемых функций $f(z)-f(a)$ и $\frac{1}{z-a}$ в области $G\setminus\{a\}$, функция $\varphi(z)$ также непрерывно дифференцируема в этой области. Пусть γ — окружность радиуса r с центром в точке a, целиком лежащая внутри D (рис. 249). При достаточно малом r кривые γ и Γ не пересекаются и, следовательно, ограничивают некоторую кольцеобразную область, целиком лежащую в G. На основании следствия из теоремы Коши (см. п. 3) можем записать $\int\limits_{\Gamma}\varphi(z)dz=$ $=\int\limits_{\gamma}\varphi(z)dz$, где контур γ ориентирован одинаково с Γ. Поскольку это равенство имеет место при любом (достаточно малом) значении r, интеграл $\int\limits_{\gamma}\varphi(z)dz$

является постоянным числом, не зависящим от r. Обозначим это число через c. Итак, $\int\limits_{\gamma}\varphi(z)dz=c$ (const) для всех (достаточно малых) r. Покажем, что $c=0$.

Так как функция $f(z)$ имеет в точке a производную, то отношение $\frac{f(z)-f(a)}{z-a}$ стремится к определенному пределу $f'(a)$ при $z\to a$. Отсюда вытекает, что если мы доопределим функцию $\varphi(z)$ в точке a, полагая $\varphi(a)=f'(a)$, то получим функцию, непрерывную во всей области G. Модуль $|\varphi(z)|$ этой функции также является непрерывной функцией в G. В ограниченной замкнутой области $\bar{D}$ (замыкании области D) функция $|\varphi(z)|$ достигает своего максимума M, т. е. $|\varphi(z)|\leqslant M$ для всех $z\in\bar{D}$. В силу неравенства (2) получим

$$\left|\int\limits_{\gamma}\varphi(z)dz\right|\leqslant M\cdot 2\pi r.$$

Таким образом, постоянное число c удовлетворяет неравенству $|c|\leqslant M\cdot 2\pi r$, где r — как угодно малое положительное число. Отсюда следует $c=0$.

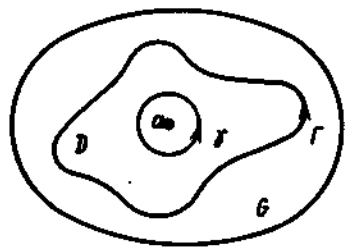


Рис. 249

Итак, мы показали, что интеграл $\int\limits_{\Gamma}\varphi(z)dz$ равен нулю, т. е.

$$\int\limits_{\Gamma}\frac{f(z)-f(a)}{z-a}dz=0.$$

Перепишем это равенство следующим образом:

$$\int\limits_{\Gamma}\frac{f(z)dz}{z-a}=\int\limits_{\Gamma}\frac{f(a)dz}{z-a}.$$

Стоящий справа интеграл равен $f(a)\int\limits_{\Gamma}\frac{dz}{z-a}$. Воспользовавшись снова следствием из теоремы Коши, запишем $\int\limits_{\Gamma}\frac{dz}{z-a}=\int\limits_{\gamma}\frac{dz}{z-a}$, где γ — любая (достаточно малая) окружность с центром a. Следовательно,

$$\int\limits_{\Gamma}\frac{f(z)dz}{z-a}=f(a)\int\limits_{\gamma}\frac{dz}{z-a}.$$

Если теперь учесть, что интеграл в правой части равен $2\pi i$ (см. пример 2 п. 2), то придем к равенству (4). ■

Формула (4) называется *интегральной формулой Коши.*

§ 69. Совпадение понятий аналитической и непрерывно дифференцируемой функции

1. Разложение непрерывно дифференцируемой функции в степенной ряд

Напомним, что согласно определению функция $f(z)$ называется аналитической в точке a, если в некотором круге с центром a она разлагается в степенной ряд по степеням $z-a$. Как мы знаем, любая функция, аналитическая в точке a, непрерывно дифференцируема в некоторой окрестности этой точки; более того, в любом круге, в котором $f(z)$ разлагается в степенной ряд, она имеет непрерывные производные всех порядков.

Из сказанного следует, что необходимым условием аналитичности функции $f(z)$ в некоторой области G является существование в этой области непрерывной производной $f'(z)$. Оказывается, что это условие также и достаточно. А именно, справедливо следующее утверждение.

Теорема (о разложении в степенной ряд). *Пусть функция $f(z)$ непрерывно дифференцируема в круге D: $|z-c|<R$. Тогда в этом круге она разлагается в степенной ряд по степеням $z-c$.*

Рис. 250

Следствием этой теоремы является высказанное выше утверждение: *если функция непрерывно дифференцируема в области G, то она аналитична в каждой точке $c \in G$* (и тем самым аналитична в области G).

□ Применим к функции $f(z)$ интегральную формулу Коши, взяв в качестве a произвольную точку $z \in D$ и в качестве Γ — окружность с центром c и радиусом r, охватывающую эту точку (рис. 250). Получим

$$f(z)=\frac{1}{2\pi i}\int_{\Gamma}\frac{f(\xi)\,d\xi}{\xi-z}. \qquad (1)$$

Задача заключается в том, чтобы интеграл, стоящий в правой части, представить в виде степенного ряда по степеням $z-c$.

Прежде всего запишем выражение $\frac{1}{\xi-z}$ следующим образом:

$$\frac{1}{\xi-z}=\frac{1}{(\xi-c)-(z-c)}=\frac{1}{(\xi-c)\left(1-\frac{z-c}{\xi-c}\right)}.$$

Дробь $\frac{z-c}{\xi-c}$ при $\xi \in \Gamma$ имеет постоянный модуль $\frac{|z-c|}{r}<1$. Поэтому можем записать разложение

$$\frac{1}{1-\frac{z-c}{\xi-c}}=1+\frac{z-c}{\xi-c}+\left(\frac{z-c}{\xi-c}\right)^2+\ldots,$$

справедливое для всех $\xi\in\Gamma$. Теперь подынтегральное выражение в равенстве (1) можно преобразовать так:

$$\frac{f(\xi)}{\xi-z}=\frac{f(\xi)}{\xi-c}+\frac{(z-c)f(\xi)}{(\xi-c)^2}+\frac{(z-c)^2f(\xi)}{(\xi-c)^3}+\ldots,$$

а само равенство (1) записать в виде

$$f(z)=\frac{1}{2\pi i}\oint\left[\sum_{n=0}^{\infty}\frac{f(\xi)(z-c)^n}{(\xi-c)^{n+1}}\right]d\xi.$$

Функциональный ряд, стоящий под знаком интеграла (члены ряда являются функциями от ξ), на окружности Γ сходится правильно, т. е. члены ряда по модулю не превосходят членов сходящегося числового ряда с положительными членами. Действительно, при $\xi\in\Gamma$ имеем

$$\left|\frac{f(\xi)(z-c)^n}{(\xi-c)^{n+1}}\right|=\frac{|f(\xi)|\,|z-c|^n}{|\xi-c|^{n+1}}\leqslant\frac{M|z-c|^n}{r^{n+1}},$$

где M — максимум модуля функции $f(\xi)$ на окружности Γ; так как $\frac{|z-c|}{r}<1$, то ряд $\sum\limits_{n=0}^{\infty}\frac{|z-c|^n}{r^{n+1}}$ с положительными членами сходится. Из правильной сходимости функционального ряда вытекает, что ряд можно интегрировать почленно. Тогда получаем

$$f(z)=\frac{1}{2\pi i}\sum_{n=0}^{\infty}\int_{\Gamma}(z-c)^n\frac{f(\xi)}{(\xi-c)^{n+1}}d\xi \quad\text{или}\quad f(z)=\sum_{n=0}^{\infty}d_n(z-c)^n, \tag{2}$$

где $d_n=\frac{1}{2\pi i}\int\limits_{\Gamma}\frac{f(\xi)d\xi}{(\xi-c)^{n+1}}$ $(n=0, 1, 2, \ldots)$. Поскольку разложение (2) имеет место при любом $z\in D$, теорема доказана. ■

2. Следствия теоремы о разложении в степенной ряд

О наиболее важном следствии фактически уже говорилось. Сформулируем его еще раз, но по-другому.

Класс функций, аналитических в области G, совпадает с классом функций, непрерывно дифференцируемых в G.

Действительно, утверждение «аналитичность $\Rightarrow$ непрерывная дифференцируемость» вытекает, как уже было отмечено, из свойств аналитических функций; обратное же утверждение «непрерывная дифференцируемость $\Rightarrow$ аналитичность» непосредственно следует из теоремы о разложении.

Сформулируем ряд других следствий теоремы о разложении. Каждое из них представляет собой утверждение об аналитичности некоторой функции $g(z)$ в предположении, что аналитична некоторая другая функция, связанная с $g(z)$.

1. *Если функция $f(z)$ аналитична в точке a, причем $f(a)\neq 0$, то функция $g(z)=1/f(z)$ также аналитична в точке a.*

□ Из аналитичности $f(z)$ в точке a следует, что эта функция непрерывно дифференцируема в некоторой окрестности точки a. Поскольку $f(a)\neq 0$, можно выбрать окрестность столь малой, чтобы во всех ее точках было $f(z)\neq 0$. Тогда и функция $g(z)=1/f(z)$ является непрерывно дифференцируемой в этой окрестности, а, значит, аналитической в точке a.

2. *Если функция $g(z)$ аналитична в точке a, а функция $f(w)$ аналитична в точке $b=g(a)$, то сложная функция $f(g(z))$ аналитична в точке a.*

□ Из условия вытекает, что функция $g(z)$ непрерывно дифференцируема в некоторой окрестности точки a, а функция $f(w)$ непрерывно дифференцируема в некоторой окрестности точки b. Тогда, как мы знаем, функция $f(g(z))$ непрерывно дифференцируема в некоторой окрестности точки a, а, значит, аналитична в точке a. ■

3. *Если функция $f(z)$ аналитична в точке a, причем $f'(a)\neq 0$, то в некоторой окрестности точки a существует обратная функция к $f(z)$, и эта обратная функция аналитична в точке $f(a)$.*

Доказательство этого утверждения также опирается на теорему о разложении, но требует более сложных рассуждений, поэтому его не приводим.

§ 70. Полюсы и вычеты аналитической функции

1. Полюсы

Функция $f(z)=\frac{1}{z-a}$ аналитична во всей плоскости, за исключением точки $z=a$. При этом $|f(z)|\to\infty$, когда $z\to a$. Введем такое определение.

Определение 1. Пусть функция $f(z)$ определена и аналитична в некоторой окрестности точки a, за исключением самой точки a. Если

$$\lim_{z\to a}|f(z)|=\infty, \tag{1}$$

то точка a называется *полюсом функции* $f(z)$.

Разумеется, $\frac{1}{z-a}$ — не единственная функция, имеющая полюс в точке a. Можно также взять $f(z)=\frac{1}{(z-a)^m}$, где m — любое натуральное число. Можно рассмотреть и функцию более общего вида:

$$f(z)=\frac{1}{(z-a)^m}\varphi(z), \tag{2}$$

где m — по-прежнему натуральное число, а $\varphi(z)$ — какая-либо функция, аналитическая в окрестности точки a (включая саму точку a), причем $\varphi(a)\neq 0$.

Определение 2. Пусть функция $f(z)$ определена и аналитична в окрестности точки a, исключая саму точку a. Точка a называется *полюсом m-го порядка* функции $f(z)$ (m — натуральное), если в некоторой окрестности точки a справедливо равенство (2), где $\varphi(z)$ — функция, аналитическая в этой окрестности, причем $\varphi(a)\neq 0$.

Пример. Показать, что точки $z=i$ и $z=-i$ являются полюсами третьего порядка функции $f(z)=\frac{1}{(z^2+1)^3}$.

○ Имеем $f(z)=\frac{1}{(z+i)^3(z-i)^3}$. Полагая $\varphi(z)=\frac{1}{(z+i)^3}$, убедимся, в том, что точка i является полюсом третьего порядка для $f(z)$. Аналогично, полагая $\psi(z)=\frac{1}{(z-i)^3}$, найдем, что точка $-i$ также является полюсом третьего порядка. ●

2. Вычет относительно полюса

Пусть a — полюс m-го порядка функции $f(z)$. Заменим в равенстве (2) функцию $\varphi(z)$ ее рядом Тейлора в точке a:

$$\varphi(z)=c_0+c_1(z-a)+c_2(z-a)^2+\dots. \tag{3}$$

Тогда получим

$$f(z)=\frac{c_0}{(z-a)^m}+\frac{c_1}{(z-a)^{m-1}}+\ldots+\frac{c_{m-1}}{z-a}+c_m+c_{m+1}(z-a)+\ldots. \qquad (4)$$

Пусть R — радиус сходимости ряда (3). Рассмотрим произвольную окружность γ: $|z-a|=r$, где $r<R$. Будем считать кривую γ ориентированной в направлении против хода часовой стрелки. Так как ряд (3) сходится правильно в замкнутом круге $|z-a|\leqslant r$ и, в частности, на окружности γ, то и ряд (4) сходится правильно на γ. Отсюда вытекает, что этот ряд можно почленно проинтегрировать вдоль γ (см. п. 3 § 54), т. е. справедливо равенство

$$\int_\gamma f(z)dz=\int_\gamma \frac{c_0}{(z-a)^m}dz+\ldots+\int_\gamma \frac{c_{m-1}}{z-a}dz+\int_\gamma c_m dz+\int_\gamma c_{m+1}(z-a)dz+\ldots.$$

В правой части все интегралы равны нулю, за исключением интеграла $\int_\gamma \frac{c_{m-1}}{z-a}dz=c_{m-1}\cdot 2\pi i$ (см. пример 2 п. 2 § 68). Следовательно, приходим к равенству

$$\int_\gamma f(z)dz=c_{m-1}\cdot 2\pi i. \qquad (5)$$

Определение 3. Пусть точка a является полюсом (некоторого порядка m) функции $f(z)$. Рассмотрим любой замкнутый круг с центром a, все точки которого, за исключением, быть может, самой точки a, принадлежат области аналитичности функции $f(z)$. Пусть γ — граничная окружность этого круга, ориентированная против хода часовой стрелки. Число $\frac{1}{2\pi i}\int_\gamma f(z)dz$ называется *вычетом* функции $f(z)$ относительно полюса a и обозначается $\mathrm{Res}(f;\,a)$.

Итак,

$$\mathrm{Res}(f;\,a)=\frac{1}{2\pi i}\int_\gamma f(z)dz,$$

где γ — любая достаточно малая окружность с центром a, ориентированная против хода часовой стрелки.

Из предыдущих рассуждений, в частности, из равенства (5), вытекает следующая теорема.

Теорема 1. *Если a — полюс некоторого порядка m функции $f(z)$, то справедлива формула $\mathrm{Res}(f;\,a)=c_{m-1}$, где c_{m-1} — коэффициент при $(z-a)^{m-1}$ в разложении (3) функции $\varphi(z)$ в ряд Тейлора.*

Так как из равенства (3) следует

$$c_{m-1}=\frac{1}{(m-1)!}\,\varphi(a)^{(m-1)},$$

то для нахождения вычета функции $f(z)$ относительно полюса a порядка m получаем формулу

$$\mathrm{Res}\,(f;\;a)=\frac{1}{(m-1)!}\,\frac{d^{m-1}[(z-a)^m f(z)]}{dz^{m-1}}\bigg|_{z=a}. \qquad (6)$$

3. Теорема о вычетах

Нахождение интегралов в комплексной области часто облегчается с помощью следующей теоремы.

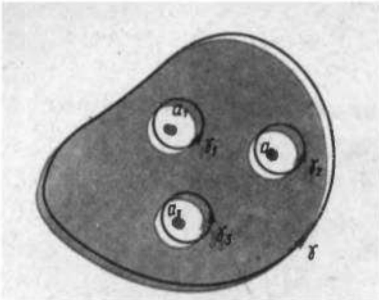


Рис. 251

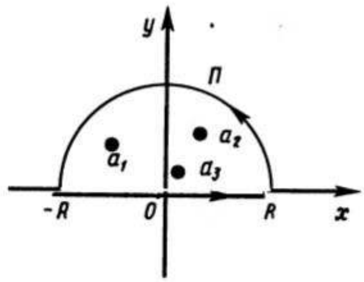


Рис. 252

Теорема 2 (о вычетах). *Пусть функция $f(z)$ аналитична во всех точках односвязной области G, за исключением нескольких точек, являющихся полюсами для $f(z)$. Пусть, далее, γ есть замкнутый контур, лежащий в G и не содержащий ни одного полюса для $f(z)$. Тогда справедливо равенство*

$$\int_{\gamma} f(z)\,dz = 2\pi i \sum_{k=1}^{n} \operatorname{Res}(f;\ a_k), \tag{7}$$

где $a_1, \ldots, a_n$ — полюсы функции $f(z)$, расположенные внутри контура γ; при этом γ считается ориентированным в направлении против хода часовой стрелки.

□ Для каждого полюса a_k рассмотрим замкнутый круг S_k: $|z-a_k| \leqslant r_k$ столь малого радиуса r_k, чтобы этот круг целиком лежал в G и не содержал других полюсов функции $f(z)$ (рис. 251); граничную окружность круга S_k обозначим γ_k. Если из области, охваченной контуром γ, выбросить круги $S_1, \ldots, S_n$, то получим неодносвязную область D. Применяя к области D и функции $f(z)$ теорему 2 § 68, находим $\int_{\gamma} f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \cdots + \int_{\gamma_n} f(z)\,dz$. Заменив каждое слагаемое $\int_{\gamma_k} f(z)\,dz$ правой части на $2\pi i \operatorname{Res}(f;\ a_k)$, придем к равенству (7). ■

4. Применение вычетов к нахождению некоторых определенных интегралов

Пусть $f(z)$ — дробно-рациональная функция, т. е. $f(z) = P(z)/Q(z)$, где $P(z)$ и $Q(z)$ — многочлены. Предположим, что степень $Q(z)$ по крайней мере на две единицы превосходит степень $P(z)$ и что $Q(z)$ не имеет действительных корней. Тогда справедлива формула

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi i \sum_{k=1}^{n} \operatorname{Res}(f;\ a_k), \tag{8}$$

где $a_1, \ldots, a_n$ — все полюсы функции $f(z)$, лежащие в верхней полуплоскости.

□ Для доказательства рассмотрим полукруг, изображенный на рис. 252. Его граница состоит из полуокружности П и отрезка $[-R, R]$ действительной оси. Радиус R выберем столь большим, чтобы все точки $a_1, a_2, \ldots, a_n$ оказались лежащими внутри указанного полукруга. Согласно теореме о вычетах,

$$\int_{-R}^{R} f(x)\,dx+\int_{\Pi} f(z)\,dz=2\pi i\sum_{k=1}^{n}\operatorname{Res}(f;\ a_k). \tag{9}$$

Покажем, что при $R\to\infty$ второй из написанных интегралов стремится к нулю.

Пусть

$$P(z)=a_0z^{p-1}+a_1z^{p-1}+\dots+a_p,\quad Q(z)=b_0z^q+b_1z^{q-1}+\dots+b_q,$$

где $a_0\neq 0$, $b_0\neq 0$; по условию $q-p\geqslant 2$. Имеем

$$f(z)=\frac{z^p\left(a_0+\frac{a_1}{z}+\dots+\frac{a_p}{z^p}\right)}{z^q\left(b_0+\frac{b_1}{z}+\dots+\frac{b_q}{z^q}\right)}.$$

Положим $\frac{a_1}{z}+\dots+\frac{a_p}{z^p}=\alpha(z)$, $\frac{b_1}{z}+\dots+\frac{b_q}{z^q}=\beta(z)$; тогда $f(z)=\frac{1}{z^{q-p}}\frac{a_0+\alpha(z)}{b_0+\beta(z)}$.

Очевидно, $\lim\alpha(z)=0$ и $\lim\beta(z)=0$ при $z\to\infty$. Поэтому дробь $\left|\frac{a_0+\alpha(z)}{b_0+\beta(z)}\right|$ при $z\to\infty$ имеет пределом число $\delta=\left|\frac{a_0}{b_0}\right|$. Отсюда, в частности, следует, что при достаточно большом $|z|$, скажем, при $|z|>R_0$, эта дробь будет меньше, чем $\delta/2$. Тогда при $|z|=R>R_0$ находим $|f(z)|<\frac{1}{|z|^{q-p}}\frac{\delta}{2}<\frac{\delta}{2}\frac{1}{R^2}$. Значит,

$$\left|\int_{\Pi} f(z)\,dz\right|\leqslant\int_{\Pi}|f(z)|dz<\frac{\delta}{2}\frac{1}{R^2}\pi R=\frac{\delta\pi}{2}\frac{1}{R},$$

откуда и вытекает равенство $\lim\limits_{R\to\infty}\int_{\Pi} f(z)\,dz=0$. Теперь из формулы (9) предельным переходом при $R\to 0$ получим формулу (8). ■

Примеры. 1. Вычислить интеграл $\int_{-\infty}^{\infty}\frac{dx}{1+x^2}$.

○ Положим $f(z)=\frac{1}{1+z^2}$. Функция $f(z)$ удовлетворяет условиям приложимости формулы (8). Равенство $f(x)=\frac{1}{(z-i)(z+i)}$ показывает, что полюсами функции $f(z)$ являются точки i и $-i$, причем вычеты $f(z)$ относительно этих полюсов равны соответственно $\frac{1}{2i}$ и $-\frac{1}{2i}$. По формуле (8) находим $\int_{-\infty}^{\infty}\frac{dx}{1+x^2}=2\pi i\cdot\operatorname{Res}(f;\ i)=2\pi i\frac{1}{2i}=\pi$. ●

2. Вычислить интеграл $\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)^2}$.

○ Положим $f(z)=\frac{1}{(1+z^2)^2}$. Единственный полюс функции $f(z)$, лежащий в верхней полуплоскости, есть точка i, причем порядок этого полюса равен 2. Так как $f(z)=\frac{1}{(z-i)^2}\varphi(z)$, где $\varphi(z)=\frac{1}{(z+i)^2}$, то на основании формулы (6) можно записать

$$\operatorname{Res}(f;\ i)=\frac{1}{1!}\frac{d}{dz}\left[\frac{(z-i)^2}{(z^2+1)^2}\right]\bigg|_{z=i}=\frac{d}{dz}\left[\frac{1}{(z+i)^2}\right]\bigg|_{z=i}=-2\frac{1}{(z+i)^3}\bigg|_{z=i}=\frac{-2}{(2i)^3}=\frac{1}{4i}.$$

Применяя затем формулу (8), получим $\int_{-\infty}^{\infty}\frac{dx}{(1+x^2)^2}=2\pi i\left(\frac{1}{4i}\right)=\frac{\pi}{2}$. ●

10

Дифференциальные уравнения

§ 71. Основные понятия

1. Дифференциальное уравнение и его порядок

До сих пор мы встречались с уравнениями вида $F(x)=0$, содержащими неизвестную величину x; задача заключалась в том, чтобы найти все значения величины x, удовлетворяющие заданному соотношению. Однако ряд важных задач — как самой математики, так и ее приложений — приводит к необходимости решать уравнения более сложного вида, где неизвестной является не величина x, а некоторая функция $y(x)$, причем в уравнение входят, наряду с x и $y(x)$, еще и производные y', y'', ... до какого-то порядка n. Приведем примеры таких уравнений:

$$y'+y-x^2=0,\ y'''=\sin x,\ y''+y=0.$$

Определение 1. Уравнение, связывающее независимую переменную x с неизвестной функцией $y(x)$ и ее производными до некоторого порядка n включительно, называется *дифференциальным уравнением* n-го порядка.

Все приведенные выше уравнения являются дифференциальными, причем первое из них имеет порядок 1, второе — порядок 3, третье — порядок 2.

Дифференциальное уравнение n-го порядка записывают обычно в виде

$$F(x,\ y,\ y',\ \ldots,\ y^{(n)})=0. \qquad (1)$$

В дальнейшем слово «дифференциальное» будем часто опускать и говорить просто «уравнение (1)».

2. Решения дифференциального уравнения

Определение 2. *Решением* дифференциального уравнения (1) называется любая функция $y=f(x)$, дифференцируемая по крайней мере n раз и такая, что при ее подстановке в уравнение (1) последнее обращается в тождество.

Так, для дифференциального уравнения

$$y'-y=0 \qquad (2)$$

одним из решений является функция $y=e^x$. Однако это решение — не единственное: любая функция вида

$$y = Ce^x \tag{3}$$

где C — постоянная, также является решением данного уравнения. Вскоре мы установим, что никаких других решений, кроме (3), данное уравнение не имеет. В этом смысле формула (3) определяет *общее решение* уравнения (2).

Поскольку в выражение (3) для y входит произвольная постоянная C, то говорят, что *множество решений уравнения* (2) *зависит от одной произвольной постоянной* C. Придавая C определенное числовое значение, мы будем получать конкретные или, как говорят, *частные решения* уравнения (2).

В качестве другого примера рассмотрим уравнение второго порядка

$$y'' = 0. \tag{4}$$

Все решения этого уравнения могут быть найдены непосредственно. Из соотношения (4) находим $y' = C_1$ и далее

$$y = C_1 x + C_2, \tag{5}$$

где C_1 и C_2 — постоянные. Обратно, при любых значениях постоянных C_1 и C_2 функция $y = C_1 x + C_2$ является решением уравнения (4). Таким образом, формула (5) определяет общее решение уравнения (4). Как видим, оно зависит от двух произвольных постоянных C_1 и C_2. При конкретных значениях C_1 и C_2 будем получать частные решения.

Понятия общего и частного решений дифференциального уравнения в дальнейшем будут уточнены. Однако одно важное обстоятельство можно отметить уже сейчас, исходя из приведенных примеров. А именно: общее решение зависит от стольких произвольных постоянных, каков порядок уравнения. Частные же решения получаются из общего при конкретных значениях этих постоянных.

Процесс отыскания решений дифференциального уравнения называют *интегрированием* этого *уравнения*. В зависимости от контекста несколько расплывчатый термин «интегрирование» может означать либо отыскание общего решения, либо нахождение того или иного частного решения.

§ 72. Дифференциальные уравнения первого порядка

1. Уравнение вида $y' = f(x, y)$ и его геометрический смысл

Наиболее общий вид дифференциального уравнения первого порядка есть

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

Если это уравнение разрешить относительно y', то оно запишется в виде

$$y' = f(x, y). \tag{1}$$

Такое уравнение мы и будем сейчас рассматривать.

Укажем прежде всего геометрический смысл уравнения (1). Возьмем какую-либо точку $(x_0; y_0)$, принадлежащую области определения D функции $f(x, y)$. Пусть $y = \varphi(x)$ — решение уравнения (1), график которого проходит через эту точку (т. е. $y_0 = \varphi(x_0)$). Чтобы найти значение производной $\varphi'(x_0)$, совсем необязательно знать функцию $\varphi(x)$, так как согласно уравнению (1) должно выполняться равенство

$$\varphi'(x_0) = f(x_0, y_0).$$

Таким образом, угловой коэффициент кривой $y = \varphi(x)$, проходящей через точ-

Рис. 253

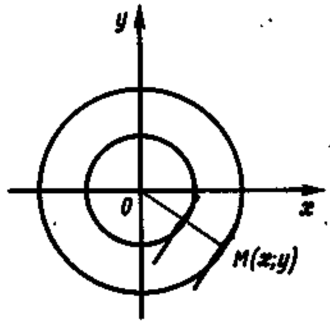


Рис. 254

ку $(x_0; y_0)$ и являющейся решением уравнения (1), равен (при $x = x_0$) числу $f(x_0, y_0)$.

Построим теперь для каждой точки (x_0, y_0) области D прямую, проходящую через эту точку и имеющую угловой коэффициент, равный $f(x_0, y_0)$. Будем говорить, что эта *прямая задает направление в точке* $(x_0; y_0)$. Функция $y = \varphi(x)$ тогда и только тогда является решением уравнения (1), когда ее график в каждой своей точке имеет заданное в этой точке направление, т. е. касается прямой, построенной для этой точки.

Определение 1. Пусть D — множество точек на плоскости. Говорят, что на этом множестве задано *поле направлений*, если для каждой точки $M \in D$ указана некоторая прямая $l(M)$, проходящая через эту точку.

Определение 2. Кривая γ, которая в каждой своей точке M имеет направление поля (т. е. касается прямой $l(M)$), называется *интегральной кривой* данного *поля направлений*. В случае, когда поле направлений отвечает дифференциальному уравнению (1), кривая γ называется *интегральной кривой уравнения* (1).

Обычно вместо прямой $l(M)$ рисуют маленький отрезок («штрих»), проходящий через M. На рис. 253 изображены некоторое поле направлений и две интегральные кривые этого поля.

Пример. Описать геометрически поле направлений для уравнения

$$y' = -\frac{x}{y}. \qquad (2)$$

○ Правая часть уравнения определена на множестве D, состоящем из всех точек $(x; y)$, где $y \neq 0$; следовательно, уравнение задает на этом множестве поле направлений. Что касается точек, не принадлежащих D, т. е. точек вида $(x; 0)$, то при $x \neq 0$ естественно считать, что и в таких точках уравнение задает определенное направление, а именно — направление, параллельное оси y (поскольку в этих точках $y' = \infty$). Таким образом, данное уравнение определяет поле направлений во всей плоскости, за исключением единственной точки $(0; 0)$. Это поле имеет простой геометрический смысл. Если $M(x; y)$ — точка, отличная от начала координат O, то прямая OM имеет угловой коэффициент $k = y/x$; перпендикулярная же ей прямая, проходящая через M, имеет угловой коэффициент $-1/k = -x/y$, что в силу (2) совпадает с y'. Таким образом, в каждой точке M, отличной от начала, направление поля перпендикулярно прямой OM (рис. 254). ●

Продолжим обсуждение примера. Из данного выше описания поля направлений, соответствующего уравнению (2), ясно, что интегральные кривые поля представляют собой окружности с центром в начале координат. К тому же заключению придем, решая данное уравнение непосредственно:

$$y' = -\frac{x}{y} \Leftrightarrow yy' + x = 0 \Leftrightarrow \frac{1}{2}(y^2)' + \frac{1}{2}(x^2)' = 0 \Leftrightarrow \frac{1}{2}x^2 + \frac{1}{2}y^2 = C.$$

Мы получили уравнение вида $x^2 + y^2 = 2C$, определяющее (при $C > 0$) окружность с центром в начале координат. В данном случае решение дифференциального уравнения привело к соотношению вида $F(x, y) = 0$, определяющему y как функцию от x неявно.

2. Задача Коши. Теорема существования и единственности решения

Мы уже отмечали, что дифференциальное уравнение имеет, как правило, бесконечное множество решений. Чтобы из этого множества выделить какое-то конкретное решение, необходимо задать дополнительное условие. Чаще всего такое условие ставится в форме следующей задачи, называемой *задачей Коши*:

Требуется найти решение $y(x)$ уравнения $y' = f(x, y)$, которое при заданном значении x_0 аргумента x принимает заданное значение y_0. Иначе говоря, требуется найти решение уравнения при *начальном условии* $y|_{x=x_0} = y_0$.

Интуиция подсказывает, что через каждую точку $(x_0; y_0)$ должна проходить единственная интегральная кривая, т. е. задача Коши должна иметь единственное решение. Как правило, дело обстоит именно так. Однако возможны и такие случаи, когда задача Коши не имеет решения либо имеет не одно, а много решений. Чтобы гарантировать существование и единственность решения задачи Коши, следует подчинить функцию $f(x, y)$ некоторым ограничениям. Точная формулировка этих ограничений дается в следующей теореме Коши.

Теорема (о существовании и единственности решения задачи Коши). *Если в некоторой окрестности начальной точки $(x_0; y_0)$ функция $f(x, y)$ определена, непрерывна и имеет непрерывную частную производную $\frac{\partial f}{\partial y}$, то существует такая окрестность точки $(x_0; y_0)$, в которой задача Коши для уравнения $y' = f(x, y)$ с начальным условием $y|_{x=x_0} = y_0$ имеет решение, и притом единственное.*

Эту теорему мы примем без доказательства.

Применим, например, теорему к уравнению $y' = \sqrt[3]{y^2}$. Здесь $f(x, y) = \sqrt[3]{y^2}$. Функция f определена на всей плоскости xy, однако ее частная производная $\frac{\partial f}{\partial y} = \frac{2}{3}\frac{1}{\sqrt[3]{y}}$ определена и непрерывна лишь в точках, где $y \neq 0$. Согласно теореме Коши, для каждой такой точки $(x_0; y_0)$ существует окрестность, в которой задача Коши с начальным условием $y|_{x=x_0} = y_0$ имеет решение, и притом единственное. Более подробное обсуждение этого примера будет дано в конце этого параграфа (см. п. 7).

3. Уточнение понятий общего и частного решений

Определение 3. Если задание начальной точки $(x_0; y_0)$ определяет единственное решение уравнения $y' = f(x, y)$, то такое решение называется *частным решением.*

Иначе говоря, частное решение — это решение, однозначно определяемое начальным условием.

Определение 4. Множество всех частных решений называется *общим решением* дифференциального уравнения.

Обычно общее решение записывается в виде $y=\varphi(x, C)$, где C — произвольная постоянная. Задание начального условия позволяет определить значение постоянной C; она находится из равенства $y_0=\varphi(x_0, C)$.

В некоторых случаях процесс решения уравнения приводит не к явному выражению $y=\varphi(x, C)$ для общего решения, а к соотношению вида $\Phi(x, y, C)=0$, определяющему y как неявно заданную функцию от x.

Определение 5. *Общим интегралом уравнения* $y'=f(x, y)$ называется соотношение $\Phi(x, y, C)=0$ (где C — произвольная постоянная), из которого при различных значениях C получаются все частные решения уравнения.

Например, для уравнения $y'=y$ условия теоремы о существовании и единственности решения выполняются во всей плоскости xy. Формула $y=Ce^x$ дает общее решение, так как любое начальное условие $y|_{x=x_0}=y_0$ удовлетворяется при подходящем выборе постоянной C. Действительно, для определения C имеем равенство $y_0=Ce^{x_0}$, откуда $C=y_0e^{-x_0}$.

Для уравнения $y'=-x/y$ условия теоремы о существовании и единственности решения выполнены во всей плоскости, за исключением точек оси x (где $y=0$). Выше мы установили, что общий интеграл имеет вид $x^2+y^2=C$. Каковы бы ни были числа x_0 и y_0, где $y_0\neq 0$, существует такое значение постоянной C (а именно, $C=x_0^2+y_0^2$), при котором функция $y(x)$, заданная неявно уравнением $x^2+y^2=C$, удовлетворяет условию $y|_{x=x_0}=y_0$.

В связи со сказанным можно внести некоторые уточнения в формулировку задачи о решении дифференциального уравнения. Слова «решить уравнение» обычно означают нахождение общего решения (или общего интеграла).

4. Уравнения с разделяющимися переменными

Наиболее простой тип дифференциального уравнения первого порядка — это уравнение вида

$$y'=f(x), \tag{3}$$

которое решается простым интегрированием обеих частей уравнения:

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

Если $F(x)$ — какая-либо первообразная функция для $f(x)$, то общее решение запишется в виде $y=F(x)+C$.

Рассмотрим теперь одно важное обобщение уравнения (3) — так называемое *уравнение с разделяющимися переменными*. Это — дифференциальное уравнение вида

$$y'=p(x)\,q(y), \tag{4}$$

где правая часть есть произведение функции от x на функцию от y (если функция $q(y)$ постоянна, то получается уравнение вида (3)).

Запишем уравнение (4) в виде

$$\frac{dy}{dx}=p(x)\,q(y).$$

Предполагая, что в рассматриваемой области изменения величины y выполняется условие $q(y)\neq 0$, перепишем уравнение (4) следующим образом:

$$\frac{dy}{q(y)}=p(x)\,dx. \tag{5}$$

Теперь левая часть содержит только y, а правая — только x, т. е. переменные, как принято говорить, разделены. Интегрируя обе части уравнения (5), получим

$$\int \frac{dy}{q(y)} = \int p(x)\,dx.$$

Если $Q(y)$ — какая-нибудь первообразная функция для $\frac{1}{q(y)}$, а $P(x)$ — первообразная для $p(x)$, то последнее равенство можно записать в виде соотношения

$$Q(y) = P(x) + C,$$

дающего, таким образом, общий интеграл уравнения (4).

Замечания. 1. Операция интегрирования обеих частей уравнения (5) нуждается в некотором обосновании, поскольку мы интегрируем, казалось бы, по разным переменным (левую часть по y, правую по x). Однако если интегрировать обе части, предполагая, что $y = y(x)$, то операция станет законной. В этом случае из (5) получаем

$$\int \frac{y'\,dx}{q(y)} = \int p(x)\,dx,$$

после чего от записи $\int \frac{y'\,dx}{q(y)}$ можно перейти к $\int \frac{dy}{q(y)}$, используя правило замены переменной в неопределенном интеграле.

2. Если ищется не общее, а частное решение уравнения (4), удовлетворяющее начальному условию $y|_{x=x_0} = y_0$, то неопределенное интегрирование в (5) можно заменить определенным; тогда получим

$$\int\limits_{y_0}^{y} \frac{dy}{q(y)} = \int\limits_{x_0}^{x} p(x)\,dx.$$

3. Если при некотором значении y_0 имеем $q(y_0) = 0$, то отыскание решения $y(x)$, удовлетворяющего условию $y(x_0) = y_0$, указанным выше методом невозможно. Однако в этом случае решением является функция $y(x)$, тождественно равная y_0. Действительно, тогда производная y' равна нулю, но и произведение $p(x)q(y_0)$ также равно нулю.

Примеры. 1. Найти все решения уравнения $y' = y^2$.

○ Это уравнение с разделяющимися переменными. Предполагая $y \neq 0$, можем записать

$$\frac{dy}{y^2} = dx, \tag{6}$$

откуда следует $-\frac{1}{y} = x + c$ или $y = -\frac{1}{x+c}$. Это общее решение уравнения. К нему следует добавить решение $y = 0$ (потерянное при переходе к уравнению (6)). ●

2. Решить уравнение $y' = \sin(x + y)$.

○ Данное уравнение не является уравнением с разделяющимися переменными, но оно приводится к нему заменой неизвестной функции $y(x)$ на $u(x) = x + y(x)$. Тогда $u' = 1 + y'$, и уравнение принимает вид $u' - 1 = \sin u$ или $u' = 1 + \sin u$. Предполагая $1 + \sin u \neq 0$, запишем это уравнение в виде

$$\frac{du}{1 + \sin u} = dx, \tag{7}$$

т. е. получим уравнение с разделяющимися переменными.

Интегрируя обе части равенства (7), имеем

$$\int \frac{du}{1 + \sin u} = x + C.$$

Чтобы найти интеграл, записанный слева, используем подстановку $\operatorname{tg}\frac{u}{2}=v$. Получаем

$$\int\frac{du}{1+\sin u}=\int\frac{2dv}{(1+v)^2}=-\frac{2}{1+v}=-\frac{2}{1+\operatorname{tg}\frac{u}{2}}.$$

Таким образом, общий интеграл уравнения (7) имеет вид

$$-\frac{2}{1+\operatorname{tg}\frac{u}{2}}=x+C.$$

Подставляя вместо функции u ее выражение $x+y$, находим

$$x+\frac{2}{1+\operatorname{tg}\frac{x+y}{2}}+C=0. \tag{8}$$

Это соотношение не охватывает тех решений y, для которых $1+\sin u=0$ или, что то же самое, $\sin(x+y)=-1$. Для таких решений имеем $x+y=-\frac{\pi}{2}+2\pi k$ $(k\in\mathbf{Z})$. Итак, данное уравнение имеет общий интеграл (8), а также решения $y=-x-\frac{\pi}{2}+2\pi k$ $(k\in\mathbf{Z})$. ●

К уравнениям с разделяющимися переменными сводится ряд других типов дифференциальных уравнений первого порядка. Некоторые из них будут рассмотрены ниже.

5. Однородные уравнения

Однородным дифференциальным уравнением первого порядка называется уравнение вида

$$\frac{dy}{dx}=f\left(\frac{y}{x}\right), \tag{9}$$

правая часть которого зависит только от отношения y/x.

Чтобы решить уравнение (9), перейдем от неизвестной функции $y(x)$ к функции $u(x)=\frac{y(x)}{x}$. Тогда $y=xu$, $\frac{dy}{dx}=u+x\frac{du}{dx}$, и уравнение принимает вид

$$u+x\frac{du}{dx}=f(u) \text{ или } \frac{du}{f(u)-u}=\frac{dx}{x}.$$

Мы получили уравнение с разделяющимися переменными, которое решается уже известным способом:

$$\int\frac{du}{f(u)-u}=\int\frac{dx}{x}=\ln|x|+C.$$

После нахождения $u(x)$ следует вернуться к функции $y(x)=xu(x)$.

Замечание. Если существуют такие значения u, для которых $f(u)=u$, то к найденным решениям добавляются еще решения вида $u(x)=u_0$, т. е. $y=u_0x$, где u_0 — любой из корней уравнения $f(u)=u$.

Пример. Решить уравнение $y'=\frac{y+x}{y-x}$.

○ Это уравнение является однородным, так как

$$\frac{y+x}{y-x}=\frac{\frac{y}{x}+1}{\frac{y}{x}-1}=\frac{u+1}{u-1},$$

где $u=\frac{y}{x}$. Для функции u имеем уравнение

$$\frac{du}{\frac{u+1}{u-1}-u}=\frac{dx}{x} \text{ или } \frac{(u-1)\,du}{-u^2+2u+1}=\frac{dx}{x}.$$

Интегрируя обе части, получаем

$$-\frac{1}{2}\ln|-u^2+2u+1|=\ln|x|-\frac{1}{2}\ln C,$$

где $c>0$ (произвольную постоянную для интеграла $\int\frac{dx}{x}$ удобно записать в виде $-\frac{1}{2}\ln C$), затем

$$|-u^2+2u+1|^{-1/2}=|x|\cdot\frac{1}{\sqrt{C}}$$

и, наконец,

$$x^2|1+2u-u^2|=C.$$

Знак модуля в последнем равенстве можно опустить вместе с ограничением на знак C. Возвращаясь к функции $y=xu$, получим

$$x^2+2xy-y^2=C. \tag{10}$$

Это общий интеграл данного уравнения. Из полученного соотношения нетрудно выразить y через x и найти общее решение.

В данном примере существуют такие значения u, для которых $f(u)-u=0$; это корни уравнения $1+2u-u^2=0$, т. е. $u=1\pm\sqrt{2}$. Им соответствуют два решения $y=(1\pm\sqrt{2})\,x$, которые можно получить из соотношения (10) при $C=0$. ●

6. Пример физической задачи, приводящей к дифференциальному уравнению первого порядка

С помощью дифференциальных уравнений решаются разнообразные задачи — физические, технические, экономические и т. д. Укажем кратко общую схему решения задач с помощью дифференциальных уравнений.

Пусть требуется найти функцию $y=y(x)$, выражающую зависимость между двумя переменными величинами x и y. Зафиксируем значение x. Тогда любому приращению Δx будет отвечать определенное значение Δy. Зависимость Δy от Δx может носить в принципе сколь угодно сложный характер. Однако если ограничиться малыми значениями Δx, то эта зависимость приближенно является линейной, поскольку отношение $\frac{\Delta y}{\Delta x}$ близко к некоторому постоянному числу k (значению производной y', вычисленному в точке x). Таким образом, имеем приближенное равенство $\Delta y\approx k\Delta x$.

Коэффициент пропорциональности k может быть найден из условий задачи, для чего следует воспользоваться известными законами физики, механики, геометрии и т. д. (в зависимости от характера задачи). Во многих

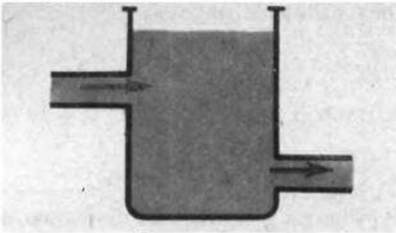


Рис. 255

случаях оказывается, что значение этого коэффициента определяется только значениями самих величин x и y, т. е. $k=f(x, y)$. Тогда для неизвестной функции $y(x)$ получим дифференциальное уравнение $\frac{dy}{dx}=f(x, y)$.

Дальнейшая часть решения задачи оказывается уже чисто математической. Она сводится к нахождению общего решения $y=\varphi(x, C)$ (или общего интеграла) полученного уравнения, а затем — к выделению определенного частного решения, поскольку задача содержит обычно те или иные дополнительные условия (например, начальные условия задачи Коши).

Пример. Котел вместимостью 1 м^3 первоначально наполнен водой, имеющей температуру 100°. По верхней трубе (рис. 255) в котел начинает поступать вода, температура которой равна 40°, одновременно по нижней трубе вода вытекает из котла. За каждую минуту втекает и вытекает 200 л. Найти закон изменения температуры воды в котле (как функции от времени). Предполагается, что вода в котле постоянно перемешивается, так что в каждый момент времени ее температура во всех точках внутри котла одна и та же. Определить, через какой промежуток времени температура воды в котле понизится до 50°.

○ Пусть $Q(t)$ — количество теплоты (в килокалориях), которое содержится в воде, наполняющей котел в момент времени t; тогда температура воды равна $Q(t)/1000$. В начальный момент (при $t=0$) имеем $Q(0)=1000\cdot 100=10^5$. Рассмотрим некоторое значение t. За последующее время Δt в котел прибывает количество теплоты, равное $200\Delta t\cdot 40$ (время Δt измеряем в минутах), и уходит количество теплоты*, равное $\frac{200}{1000}Q(t)\,\Delta t$. Следовательно, получаем равенство

$$\Delta Q=200\cdot 40\cdot\Delta t-\frac{200}{1000}Q(t)\,\Delta t,$$

откуда после деления обеих частей на Δt имеем

$$\frac{\Delta Q}{\Delta t}=8000-\frac{1}{5}Q(t).$$

В пределе при $\Delta t\to 0$ это равенство примет вид

$$\frac{dQ}{dt}=8000-\frac{1}{5}Q(t).$$

Получили дифференциальное уравнение с разделяющимися переменными

$$\frac{dQ}{\frac{1}{5}Q-8000}=-dt \text{ или } \frac{dQ}{Q-40\,000}=-\frac{1}{5}dt.$$

---

* Количество ушедшей за время Δt теплоты указано лишь приближенно. Точное значение находится в промежутке между $\frac{200}{1000}Q(t)\Delta t$ и $\frac{200}{1000}(Q(t)+\Delta Q)\Delta t$, т. е. отличается от указанного значения на величину $\alpha\Delta t$, где α — бесконечно малая при $\Delta t\to 0$. Подобного рода уточнения обычно остаются «за текстом».

Решая это уравнение, последовательно находим:

$$\ln|Q-40\,000| = -\frac{t}{5}+\ln C,\ Q-40\,000 = Ce^{-t/5},\ Q = Ce^{-t/5}+40\,000.$$

Начальное условие $Q(0)=100\,000$ дает $100\,000 = C+40\,000$, т. е. $C=60\,000$. Итак, окончательно получим

$$Q = 40\,000+60\,000e^{-t/5}.$$

Отсюда находим температуру воды в котле как функцию от t:

$$T = \frac{Q}{1000} = 40+60e^{-t/5}.$$

Обозначая через t_0 промежуток времени, за который температура понизится до 50°, можем записать $50 = 40+60e^{-t_0/5}$, т. е. $e^{-t_0/5} = 1/6$, откуда $t_0 \approx 8{,}96$ мин. ●

7. Понятие об особых точках и особых решениях дифференциального уравнения

В п. 2 была сформулирована теорема о существовании и единственности решения задачи Коши для уравнения $y' = f(x, y)$. Условием, гарантирующим как существование решения, так и его единственность, является непрерывность функции $\frac{\partial f}{\partial y}$. В отдельных точках это условие может нарушаться; через такие точки может не проходить ни одной интегральной кривой или проходить несколько интегральных кривых.

Точки, через которые не проходит ни одна интегральная кривая или проходит более одной интегральной кривой, называются *особыми точками* данного дифференциального уравнения.

Может случиться, что некоторая интегральная кривая уравнения состоит из одних особых точек. Такая кривая называется *особым решением* уравнения.

Например, для уравнения

$$y' = 3\sqrt[3]{y^2} \tag{11}$$

функция $f(x, y) = 3y^{2/3}$ определена и непрерывна на всей плоскости xy. Ее частная производная $\frac{\partial f}{\partial y} = \frac{2}{\sqrt[3]{y}}$ существует и непрерывна во всех точках, где $y \neq 0$, т. е. во всех точках, не принадлежащих оси x. Таким образом, через любую точку, не лежащую на оси x, проходит единственная интегральная кривая уравнения. Чтобы выяснить, как обстоит в этом смысле дело с точками оси x, проинтегрируем данное уравнение. Имеем $\frac{dy}{3y^{2/3}} = dx$, откуда следует $y^{1/3}+C = x$ или $y = (x-C)^3$.

Итак, общее решение представляет собой семейство кубических парабол. Однако имеется еще одно решение $y(x) = 0$. Следовательно, через любую точку $(C, 0)$ оси x проходят по крайней мере две интегральные кривые: ось x и парабола $y = (x-C)^3$. Это показывает, что точки оси x являются особыми точками уравнения (11), а функция $y(x) \equiv 0$ — особым решением.

Нетрудно установить, что через любую точку вида $(C, 0)$ проходит в действительности бесчисленное множество интегральных кривых. Любую из них можно составить из трех кусков: «нижней» половины параболы $y = (x-C_1)^3$, где C_1 — число, меньшее или равное C (рис. 256), отрезка C_1C_2 оси x, где $C_2 > C_1$, и «верхней» половины параболы $y = (x-C_2)^3$.

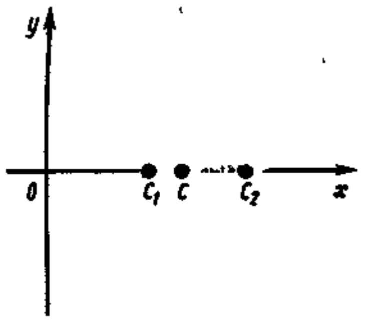


Рис. 256

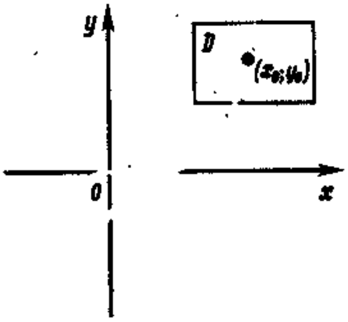


Рис. 257

Из этого примера можно понять, почему в формулировке теоремы Коши мы были вынуждены говорить о существовании и единственности решения лишь в некоторой окрестности начальной точки $(x_0;\ y_0)$, а не во всей области существования функции $f(x,\ y)$. В самом деле, пусть точка $(x_0;\ y_0)$ не является особой для уравнения (11), т. е. $y_0 \neq 0$. Если взять столь малую окрестность D точки $(x_0;\ y_0)$, чтобы она не пересекала ось x, то внутри такой окрестности через точку $(x_0;\ y_0)$ будет проходить единственная кубическая парабола вида $y=(x-C)^3$. Однако если взять достаточно большую окрестность (например, всю плоскость xy), то окажется, что внутри такой окрестности через точку $(x_0;\ y_0)$ проходит бесчисленное множество интегральных кривых. Любую из них можно составить указанным выше способом из трех кусков, один из которых проходит через точку $(x_0;\ y_0)$ (рис. 257).

8. Приближенное решение дифференциального уравнения. Метод ломаных Эйлера

В предыдущих пунктах мы рассмотрели (притом лишь для некоторых классов дифференциальных уравнений первого порядка) методы нахождения точных решений. В тех случаях, когда ни один из них не достигает цели, можно воспользоваться приближенными методами. Опишем кратко простейший из них — так называемый *метод ломаных Эйлера.*

Пусть требуется найти приближенное решение дифференциального уравнения $y'=f(x,\ y)$, удовлетворяющее начальному условию $y(a)=y_0$. При этом неизвестная функция $y(x)$ ищется на отрезке $[a,\ b]$.

Разобьем отрезок $[a,\ b]$ точками деления $x_1,\ x_2,\ \ldots,\ x_{n-1}$ на n частей $[x_0,\ x_1]$, $[x_1,\ x_2]$, ..., $[x_{n-1},\ x_n]$, где $x_0=a$, $x_n=b$. В точке x_0 график искомой функции $y(x)$ имеет угловой коэффициент $k_0=f(x_0,\ y_0)$. Заменим график на участке $[x_0,\ x_1]$ приближенно отрезком прямой, выходящим из точки $(x_0;\ y_0)$ и имеющим угловой коэффициент k_0. Тогда получим

$$y(x_1)\approx y_0+f(x_0,\ y_0)(x_1-x_0).$$

Правую часть этого равенства обозначим y_1.

Аналогично, заменив график функции $y(x)$ на участке $[x_1,\ x_2]$ приближенно отрезком прямой, выходящим из точки $(x_1;\ y_1)$ и имеющим угловой коэффициент $k_1=f(x_1,\ y_1)$, находим

$$y(x_2)\approx y_1+f(x_1,\ y_1)(x_2-x_1),$$

и т. д. В результате после n шагов получим числа

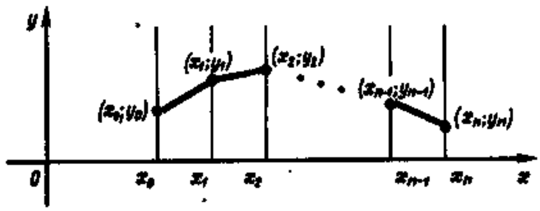


Рис. 258

$$\begin{aligned} y_1 &= y_0 + f(x_0,\ y_0)(x_1 - x_0), \\ y_2 &= y_1 + f(x_1,\ y_1)(x_2 - x_1), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= y_{n-1} + f(x_{n-1},\ y_{n-1})(x_n - x_{n-1}), \end{aligned} \tag{12}$$

приближенно равные значениям $y(x_1)$, $y(x_2)$, ..., $y(x_n)$ искомого решения $y(x)$.

Геометрический смысл рассмотренного метода состоит в замене кривой $y = y(x)$ ломаной линией, состоящей из n звеньев (рис. 258) и называемой *ломаной Эйлера*.

На практике удобно делить отрезок $[a, b]$ на равные части. Тогда $x_i - x_{i-1} = \Delta x = (b-a)/n$, и расчетные формулы (12) принимают вид

$$y_i = y_{i-1} + f(x_{i-1},\ y_{i-1})\,\Delta x \quad (i = 1,\ 2,\ \ldots,\ n). \tag{13}$$

Можно показать, что если правая часть $f(x, y)$ дифференциального уравнения удовлетворяет условиям теоремы 1, то ломаные Эйлера с ростом n неограниченно приближаются к искомому решению $y(x)$.

Пример. Найти приближенное решение уравнения $y' = 1 + y$, удовлетворяющее начальному условию $y(0) = 1$. Значения $y(x)$ вычислить для 10 равноотстоящих точек $x_1 = 0{,}1$, $x_2 = 0{,}2$, ..., $x_{10} = 1$. Сравнить приближенное решение с точным.

○ Используя формулы (13), находим

$$y_1 = 1 + (1+1)\cdot 0{,}1 = 1{,}2, \quad y_2 = 1{,}2 + (1 + 1{,}2)\cdot 0{,}1 = 1{,}42$$

и т. д. Вычисления удобно располагать в виде таблицы, которая в данном случае имеет вид

i	x_i	y_i	$f(x_i, y_i)\Delta x$	i	x_i	y_i	$f(x_i, y_i)\Delta x$
0	0	1	0,2	6	0,6	2,5431	0,3543
1	0,1	1,2	0,22	7	0,7	2,8974	0,3897
2	0,2	1,42	0,242	8	0,8	3,2871	0,4287
3	0,3	1,662	0,2662	9	0,9	3,7158	0,4715
4	0,4	1,9282	0,2928	10	1,0	4,1873	
5	0,5	2,2210	0,3221				

Точное решение указанного уравнения при заданном начальном условии есть $y = 2e^x - 1$. Для сравнения приведем таблицу значений этой функции:

x	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
y	1,2103	1,4428	1,6997	1,9836	2,2974	2,6642	3,0275	3,4511	3,9192	4,4366

§ 73. Линейные дифференциальные уравнения первого порядка. Уравнение Бернулли

1. Линейное уравнение первого порядка

Дифференциальное уравнение первого порядка $F(x, y, y')=0$ называется *линейным*, если его левая часть линейно зависит от y и y'. Таким образом, линейное уравнение имеет вид

$$\alpha(x)y'+\beta(x)y+\gamma(x)=0.$$

Предполагая, что $\alpha(x)\neq 0$, и разделив обе части на $\alpha(x)$, приведем уравнение к виду

$$y'+p(x)\,y=f(x), \tag{1}$$

где $p(x)=\beta(x)/\alpha(x)$, $f(x)=-\gamma(x)/\alpha(x)$.

Интегрирование уравнения (1) обычно проводят в два этапа. Сначала находят общее решение уравнения

$$y'+p(x)\,y=0, \tag{2}$$

получаемого из уравнения (1) заменой функции $f(x)$, стоящей в правой части, нулем. Уравнение (2) называется *линейным однородным** уравнением, соответствующим уравнению (1); в противоположность этому само уравнение (1) называется (в случае, когда $f(x)\not\equiv 0$) *неоднородным*. После того как получено общее решение однородного уравнения, находят какое-либо частное решение $y_*(x)$ неоднородного уравнения (1). Тогда общее решение уравнения (1) можно получить с помощью следующей теоремы.

Теорема. *Общее решение неоднородного уравнения* (1) *есть сумма частного решения $y_*(x)$ этого уравнения и общего решения соответствующего однородного уравнения* (2).

□ Прежде всего убедимся, что сумма $y_*(x)$ и любого решения $y_0(x)$ однородного уравнения также является решением уравнения (1). Положим $y(x)=$ $=y_0(x)+y_*(x)$. Тогда

$$y'+p(x)\,y=(y_0'+y_*')+p(x)\,(y_0+y_*)=(y_0'+p(x)\,y_0)+(y_*'+p(x)\,y_*)=$$
$$=0+f(x)=f(x),$$

что и требовалось установить.

Теперь остается показать, что всякое решение $y(x)$ неоднородного уравнения есть сумма $y_*(x)$ и некоторого решения $y_0(x)$ однородного уравнения; иначе говоря, что разность $y(x)-y_*(x)$ является решением однородного уравнения. Имеем

$$(y-y_*)'+p(x)\,(y-y_*)=(y'+p(x)\,y)-(y_*'+p(x)\,y_*)=f(x)-f(x)=0,$$

что и завершает доказательство теоремы. ■

Решим однородное уравнение (2). Оно представляет собой уравнение с разделяющимися переменными

$$\frac{dy}{y}=-p(x)\,dx$$

* Не следует смешивать с однородным уравнением в смысле п. 5 § 72.

и его общее решение имеет вид

$$y = Ce^{-P(x)}, \tag{3}$$

где $P(x)$ обозначает одну из первообразных для функции $p(x)$.

Теперь найдем частное решение уравнения (1). Воспользуемся для этого приемом, который называется *вариацией произвольной постоянной*. А именно, будем искать решение $y_*(x)$ уравнения (1) в виде произведения:

$$y_*(x) = u(x)e^{-P(x)},$$

которое получается из (3) заменой постоянной C некоторой функцией $u(x)$ (отсюда и название «вариация произвольной постоянной»). Подставляя это выражение для $y_*(x)$ в уравнение (1), для неизвестной функции $u(x)$ получим уравнение

$$u'e^{-P(x)} - uP'(x)\,e^{-P(x)} + p(x)\,ue^{-P(x)} = f(x).$$

Поскольку $P'(x) = p(x)$, второе и третье слагаемое в левой части взаимно уничтожаются, и для функции u получается уравнение

$$u'e^{-P(x)} = f(x) \text{ или } \frac{du}{dx} = f(x)\,e^{P(x)},$$

из которого следует $u(x) = \int f(x)\,e^{P(x)}dx$ (одна из первообразных).

Пример. Решить уравнение

$$y' - 2xy = 2x. \tag{4}$$

○ Это линейное уравнение. Соответствующее однородное уравнение имеет вид $y' - 2xy = 0$. Решая его, получаем

$$\frac{dy}{y} = 2x\,dx, \tag{5}$$

откуда

$$\ln|y| = x^2 + \ln C \ (C > 0),\ y = \pm Ce^{x^2} = Ae^{x^2},$$

где постоянная A может быть как положительной, так и отрицательной или нулем (случай $A = 0$ позволяет учесть решение $y = 0$, потерянное при переходе к уравнению (5)).

Теперь находим частное решение $y_*(x)$ исходного уравнения в виде произведения:

$$y_*(x) = u(x)\,e^{x^2} = u(x)\,y_0(x),$$

где $y_0(x) = e^{x^2}$. Подставляя это выражение для $y_*(x)$ в уравнение (4), получим

$$u'y_0 + uy_0' - 2xy_0u = 2x$$

откуда, учитывая, что $y_0' - 2xy_0 = 0$, находим

$$u' = 2xe^{-x^2}.$$

Следовательно, $u = -e^{-x^2}$ (берем частное решение). Итак, $y_*(x) = -e^{-x^2}\cdot e^{x^2} = -1$ (заметим, что это частное решение уравнения (4) можно было обнаружить непосредственно). Теперь на основании теоремы находим общее решение уравнения (4). Оно записывается в виде $y_*(x) + Cy_0(x)$, т. е. в виде $y(x) = -1 + Ce^{x^2}$. ●

2. Уравнение Бернулли

Метод, использованный выше для решения линейного уравнения, позволяет решать и некоторые нелинейные уравнения. В частности, с его помощью решается уравнение вида

$$y' + p(x)\,y = q(x)\,y^n, \tag{6}$$

называемое *уравнением Бернулли.* Как и выше, сначала находим какое-нибудь решение $y_0(x)$ однородного уравнения $y'+p(x)y=0$; затем полагаем $y(x)=u(x)y_0(x)$. Подставляя это выражение в уравнение (6), для функции $u(x)$ получаем уравнение

$$u'(x)\,y_0(x)=q(x)\,u^n(x)\,y_0^n(x),$$

которое решаем как уравнение с разделяющимися переменными:

$$\frac{du}{u^n}=q(x)\,y_0^{n-1}(x)\,dx.$$

Пример. Решить уравнение

$$y'+2y=y^2e^{2x}. \qquad (7)$$

○ Сначала находим решение уравнения $y'+2y=0$. В качестве такого решения можно взять функцию $y_0(x)=e^{-2x}$. Затем полагаем $y(x)=e^{-2x}u(x)$. Подставляя это выражение в уравнение (7), получаем

$$e^{-2x}u'=e^{-4x}e^{2x}u^2 \text{ или } \frac{du}{u^2}=1,$$

откуда $\frac{1}{u}=-x+C$, или $u=\frac{1}{-x+C}$. Окончательно имеем

$$y(x)=u(x)\,y_0(x)=\frac{e^{-2x}}{C-x}. \text{ ●}$$

§ 74. Линейные дифференциальные уравнения второго порядка

1. Начальные условия

Дифференциальное уравнение второго порядка имеет вид

$$F(x,\ y,\ y',\ y'')=0, \qquad (1)$$

где $y(x)$ — неизвестная функция. К уравнениям второго порядка приводят, в частности, различные задачи механики. Пусть, например, материальная точка с массой m движется вдоль оси x, причем сила, действующая на точку, задана как функция времени: $F=F(t)$. Закон движения выражается с помощью функции $x=x(t)$, задающей положение точки на оси в произвольный момент времени t. Согласно второму закону Ньютона, имеем $ma=F$, где a — ускорение точки, т. е. $a=\frac{d^2x}{dt^2}$. Таким образом, функция $x(t)$ должна удовлетворять дифференциальному уравнению

$$\frac{d^2x}{dt^2}=\frac{1}{m}F(t). \qquad (2)$$

В более общем случае сила F может зависеть не только от момента времени t, но и от положения точки в момент t (таковы, например, сила тяготения или сила упругости), а также от ее скорости. Тогда $F=F\left(t,\ x,\ \frac{dx}{dt}\right)$, и вместо уравнения (2) имеем уравнение более общего вида

$$\frac{d^2x}{dt^2}=F\left(t,\ x,\ \frac{dx}{dt}\right). \qquad (3)$$

Для дифференциального уравнения важное значение имеет вопрос о дополнительных условиях, позволяющих получить какое-то одно определенное

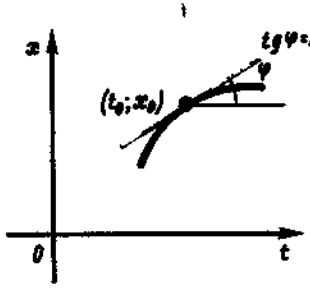


Рис. 259

решение уравнения. Какой характер могут носить эти условия для уравнения (3)? Возможный ответ на этот вопрос подсказывает рассмотренная выше механическая модель. Интуитивно ясно, что если задано положение точки в некоторый момент времени t_0, а также ее скорость в момент t_0, то дальнейшее движение точки однозначно определяется этими условиями. Поэтому имеются основания выбрать дополнительные условия в виде

$$x\Big|_{t=t_0}=x_0, \quad \frac{dx}{dt}\Big|_{t=t_0}=x'_0,$$

где t_0, x_0, x'_0 — заданные числа. Геометрически это означает, что ищется интегральная кривая уравнения (3), проходящая через заданную точку $(t_0; x_0)$ и имеющая в этой точке заданное направление, которое характеризуется угловым коэффициентом x'_0 (рис. 259). Поставленная таким образом задача называется *задачей Коши* для уравнения (1).

В дальнейшем независимую переменную будем, как и ранее, обозначать через x, а неизвестную функцию — через y; таким образом, мы возвращаемся к записи уравнения второго порядка в форме (1) (а не в форме (3)).

Обычно уравнение (1) удается разрешить относительно y'', т. е. привести его к виду

$$y''=f(x, y, y'). \qquad (4)$$

Сформулируем теорему о существовании и единственности решения задачи Коши для уравнения (4).

Теорема 1. *Если в некоторой окрестности значений x_0, y_0, y'_0 функция $f(x, y, y')$ определена, непрерывна и имеет непрерывные частные производные $\frac{\partial f}{\partial y}$ и $\frac{\partial f}{\partial y'}$, то существует такая окрестность точки $(x_0; y_0; y'_0)$ (в пространстве $\mathbf{R}^3$), в которой задача Коши для уравнения (4) с начальными условиями $y|_{x=x_0}=y_0$, $y'|_{x=x_0}=y'_0$ имеет решение, и притом единственное.*

Эту теорему мы оставляем без доказательства.

Как уже отмечалось в § 71, общее решение уравнения второго порядка зависит от двух произвольных постоянных, т. е. имеет вид $y=\varphi(x, C_1, C_2)$. Это вполне согласуется с существованием и единственностью решения задачи Коши: из равенств $\varphi(x_0, C_1, C_2)=y_0$, $\varphi'(x_0, C_1, C_2)=y'_0$, вообще говоря, однозначно определяются значения C_1 и C_2, а значит, и частное решение уравнения (1)

Например, общее решение уравнения $y''=0$ имеет вид $y=C_1x+C_2$. Таким образом, интегральные кривые представляют собой прямые на плоскости.

Через данную точку $(x_0; y_0)$ плоскости в данном направлении, характеризуемом угловым коэффициентом y_0', проходит интегральная кривая, и притом единственная.

В качестве другого примера рассмотрим уравнение $y''+y=0$. Нетрудно проверить, что оно имеет частные решения $y_1=\cos x$ и $y_2=\sin x$. Функция

$$y=C_1y_1+C_2y_2=C_1\cos x+C_2\sin x$$

при любых значениях постоянных C_1 и C_2 также является решением, поскольку

$$(C_1y_1+C_2y_2)''+(C_1y_1+C_2y_2)=C_1(y_1''+y_1)+C_2(y_2''+y_2)=C_1\cdot 0+C_2\cdot 0=0.$$

Это решение зависит от двух произвольных постоянных C_1 и C_2. Каковы бы ни были числа x_0, y_0, y_0' (начальные условия задачи Коши), существует единственная функция вида $C_1\cos x+C_2\sin x$, удовлетворяющая условиям $y|_{x=x_0}=$ $=y_0$ и $y'|_{x=x_0}=y_0'$. Например, если $x_0=0$, $y_0=1$, $y_0'=1$, то для нахождения C_1 и C_2 имеем условия

$$y|_{x=x_0}=C_1\cos 0+C_2\sin 0=1,\ y'|_{x=x_0}=-C_1\sin 0+C_2\cos 0=1,$$

из которых следует $C_1=1$, $C_2=1$, т. е. $y=\cos x+\sin x$.

2. Линейные дифференциальные уравнения второго порядка

Дифференциальное уравнение второго порядка называется *линейным*, если оно имеет вид

$$\alpha(x)y''+\beta(x)y'+\gamma(x)y=\delta(x),$$

где левая часть линейна по отношению к y, y', y''. Предполагая, что $\alpha(x)\neq 0$, и разделив обе части уравнения на $\alpha(x)$, придем к уравнению

$$y''+p(x)y'+q(x)y=f(x). \tag{5}$$

Будем считать функции $p(x)$, $q(x)$, $f(x)$ непрерывными; согласно теореме 1 это обеспечивает существование и единственность решения задачи Коши для уравнения (5).

При рассмотрении уравнения (5) важную роль играет уравнение

$$y''+p(x)y'+q(x)y=0, \tag{6}$$

которое называется *линейным однородным уравнением,* соответствующим уравнению (5). Множество решений уравнения (6) обладает рядом особенностей, которые делают оправданным специальное изучение этого уравнения. Такое изучение будет проведено в пп. 3 и 4.

3. Линейно независимые и линейно зависимые функции. Определитель Вронского

Пусть $y_1(x)$, $y_2(x)$, ..., $y_k(x)$ — какие-то функции. Рассмотрим выражение

$$y(x)=C_1y_1(x)+C_2y_2(x)+...+C_ky_k(x),$$

где C_1, C_2, ..., C_k — постоянные числа. Любая функция $y(x)$ такого вида называется *линейной комбинацией* функций $y_1(x)$, $y_2(x)$, ..., $y_k(x)$.

Определение 1. Функции $y_1(x)$, $y_2(x)$, ..., $y_k(x)$ называется *линейно независимыми,* если ни одна из них не является линейной комбинацией остальных. В противном случае, т. е. если какая-то из данных функций может быть представлена как линейная комбинация остальных, эти функции называются *линейно зависимыми.*

Какими способами можно установить линейную независимость нескольких функций? Один из способов связан с так называемым *определителем Вронского**. Для системы, состоящей из двух функций $y_1(x)$, $y_2(x)$, определитель Вронского имеет вид

$$W(y_1, y_2)=\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix},$$

для системы из трех функций y_1, y_2, y_3 — вид

$$W(y_1, y_2, y_3)=\begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix}$$

и т. д. Отметим, что как сами функции $y_1(x)$, $y_2(x)$, ... так и их определитель Вронского являются функциями от x.

Теорема 2. *Если функции y_1, y_2, ..., y_k линейно зависимы, то их определитель Вронского тождественно (т. е. при всех x) равен нулю.*

□ Для сокращения записи положим $k=3$. Пусть функции y_1, y_2, y_3 линейно зависимы; например, пусть y_3 есть линейная комбинация y_1 и y_2: $y_3=C_1y_1+C_2y_2$. Тогда имеем

$$W(y_1, y_2, y_3)=\begin{vmatrix} y_1 & y_2 & C_1y_1+C_2y_2 \\ y_1' & y_2' & C_1y_1'+C_2y_2' \\ y_1'' & y_2'' & C_1y_1''+C_2y_2'' \end{vmatrix},$$

но такой определитель равен нулю, поскольку его третий столбец является линейной комбинацией первого и второго столбцов. ■

Следствие. *Если $W(y_1, y_2, \dots, y_k)\neq 0$ (хотя бы при одном значении x), то функции y_1, y_2, ..., y_k линейно независимы.*

Заметим, что обратное утверждение, а именно: если y_1, y_2, ..., y_k линейно независимы, то $W(y_1, y_2, \dots, y_k)\neq 0$, вообще говоря, неверно.

4. Фундаментальная система решений и общее решение линейного однородного уравнения второго порядка

Займемся теперь подробно изучением однородного линейного уравнения

$$y''+p(x)y'+q(x)y=0. \tag{7}$$

Теорема 3. *Линейная комбинация нескольких решений уравнения (7) также является решением этого уравнения.*

□ Пусть каждая из функций $y_1(x)$, ..., $y_k(x)$ является решением уравнения (7). Рассмотрим какую-нибудь линейную комбинацию этих функций:

$$y(x)=C_1y_1(x)+\dots+C_ky_k(x).$$

Дифференцируя это равенство, находим

$$y'=C_1y_1'+\dots+C_ky_k',\quad y''=C_1y_1''+\dots+C_ky_k''.$$

* Ю. Вронский (1776—1853) — польский математик.

Отсюда получим
$$y''+p(x)y'+q(x)=(C_1y_1''+\ldots+C_ky_k'')+p(x)(C_1y_1'+\ldots+C_ky_k')+$$
$$+q(x)(C_1y_1+\ldots+C_ky_k)=C_1(y_1''+p(x)y_1'+q(x)y_1)+\ldots+C_k(y_k''+p(x)y_k'+$$
$$+q(x)y_k)=C_1\cdot 0+\ldots+C_k\cdot 0=0,$$
т. е. функция $y(x)$ также является решением уравнения (7). ■

Теорема 4. *Если $y_1(x)$ и $y_2(x)$ — два линейно независимых решения уравнения (7), то их определитель Вронского $W(y_1, y_2)$ ни при одном значении x не обращается в нуль.*

□ Рассуждая от противного, допустим, что в некоторой точке x_0 справедливо равенство $W=0$. Найдем такие два числа C_1 и C_2, не равные одновременно нулю, чтобы выполнялись равенства
$$\begin{cases} C_1y_1(x_0)+C_2y_2(x_0)=0, \\ C_1y_1'(x_0)+C_2y_2'(x_0)=0. \end{cases} \tag{8}$$
Искомые числа C_1 и C_2 обязательно существуют, так как определитель системы (8), имеющий вид
$$\begin{vmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{vmatrix},$$
есть определитель Вронского $W(y_1, y_2)$ в точке x_0 и, следовательно, равен нулю. Далее, рассмотрим функцию
$$\varphi(x)=C_1y_1(x)+C_2y_2(x),$$
которая в силу теоремы 3 является решением уравнения (7). Имеем $\varphi(x)\not\equiv 0$: в противном случае тождественно по x выполнялось бы равенство $C_1y_1(x)+$ $+C_2y_2(x)=0$, что означало бы линейную зависимость функций $y_1(x)$ и $y_2(x)$ (здесь существенно используется тот факт, что C_1 и C_2 не равны одновременно нулю). Равенства (8) означают, что $\varphi(x_0)=0$, $\varphi'(x_0)=0$. Однако тем же начальным условиям удовлетворяет и другое решение уравнения (7), а именно, функция $y(x)$, тождественно равная нулю. Это противоречит единственности решения задачи Коши для уравнения (7). Таким образом, $W(y_1, y_2)\neq 0$ для всех x. ■

Из теорем 2 и 4 вытекает, что для любой пары решений $y_1(x)$ и $y_2(x)$ уравнения (7) имеются только две возможности:

$y_1(x)$, $y_2(x)$ линейно зависимы; тогда (по теореме 2) $W(y_1, y_2)=0$ при любом значении x;

$y_1(x)$, $y_2(x)$ линейно независимы, тогда (по теореме 4) $W(y_1, y_2)\neq 0$ при любом значении x.

Мы располагаем теперь всем необходимым для доказательства следующей теоремы, занимающей центральное место в теории линейных однородных дифференциальных уравнений второго порядка.

Теорема 5 (о структуре множества всех решений однородного уравнения). *Если два решения $y_1(x)$ и $y_2(x)$ уравнения (7) линейно независимы, то любое решение уравнения можно представить в виде их линейной комбинации.*

□ Пусть $\varphi(x)$ — произвольное решение уравнения (7). Выберем некоторое значение x_0 и обозначим числа $\varphi(x_0)$ и $\varphi'(x_0)$ соответственно через y_0 и y_0'. Если мы докажем, что существует решение вида $y=C_1y_1+C_2y_2$, удовлетворяющее начальным условиям $y|_{x=x_0}=y_0$, $y'|_{x=x_0}=y_0'$, то в силу единственности решения задачи Коши для уравнения (7) отсюда будет следовать $\varphi(x)=y(x)$, т. е. что заданное решение $\varphi(x)$ есть линейная комбинация решений y_1 и y_2.

Для нахождения искомых чисел C_1 и C_2 имеем следующие соотношения:

$$\begin{aligned} C_1y_1(x_0)+C_2y_2(x_0)&=y_0,\\ C_1y_1'(x_0)+C_2y_2'(x_0)&=y_0'. \end{aligned} \qquad (9)$$

Эта система двух линейных алгебраических уравнений с двумя неизвестными C_1, C_2. Определитель, составленный из коэффициентов при неизвестных, равен

$$\begin{vmatrix} y_1(x_0) & y_2(x_0) \\ y_1'(x_0) & y_2'(x_0) \end{vmatrix},$$

т. е. совпадает с определителем Вронского для функций $y_1(x)$ и $y_2(x)$ в точке x_0. Ввиду линейной независимости функций $y_1(x)$ и $y_2(x)$ этот определитель отличен от нуля. Следовательно решение C_1, C_2 системы (9) существует. ■

Определение 2. Набор из двух линейно независимых решений $y_1(x)$ и $y_2(x)$ уравнения (7) называется *фундаментальной системой решений* этого уравнения.

Согласно предыдущему, для фундаментальности системы $y_1(x)$, $y_2(x)$ необходимо и достаточно выполнение условия $W(y_1, y_2)\neq 0$. Используя данное определение, теорему 5 можно сформулировать по-другому. Ее новая формулировка выглядит так: *если $y_1(x)$, $y_2(x)$ — какая-либо фундаментальная система решений однородного уравнения (7), то общее решение уравнения имеет вид $y=C_1y_1+C_2y_2$, где C_1 и C_2 — произвольные постоянные.*

Примеры. 1. Для уравнения $y''-y=0$ функции $y_1(x)=e^x$ и $y_2(x)=e^{-x}$ являются частными решениями. Эти решения линейно независимы (образуют фундаментальную систему), так как их определитель Вронского

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}=\begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix}=-2$$

отличен от нуля. Следовательно, общее решение имеет вид

$$y=C_1e^x+C_2e^{-x}.$$

2. Для уравнения $y''+y=0$ очевидными частными решениями являются функции $y_1=\cos x$ и $y_2=\sin x$. Их определитель Вронского

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}=\begin{vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{vmatrix}=1;$$

следовательно, y_1 и y_2 линейно независимы. Общее решение есть

$$y=C_1\cos x+C_2\sin x.$$

3. Найти общее решение уравнения $y''-y'+\frac{1}{x}y=0$.

○ В данном случае одним из решений является функция $y_1=x$. Будем искать второе частное решение с помощью подстановки $y=y_1u$, где u — новая неизвестная функция (можно показать, что такой способ нахождения второго решения применим к любому линейному уравнению). Имеем $y'=u+xu'$, $y''=2u'+xu''$; уравнение принимает вид

$$2u'+xu''-u-xu'+u=0, \text{ или } xu''+(2-x)u'=0.$$

Так как в полученное уравнение входит не сама неизвестная функция u, а лишь ее производные, то можно понизить порядок уравнения с помощью подстановки $v=u'$. Получим

$$xv'+(2-x)v=0, \text{ или } \frac{dv}{v}=\frac{x-2}{x}dx,$$

откуда $v=\frac{e^x}{x^2}$. Следовательно, $u=\int v\,dx=\int\frac{e^x}{x^2}dx$, и искомое решение $y_2(x)=x\int\frac{e^x}{x^2}dx$. Линейная независимость $y_1(x)$ и $y_2(x)$ очевидна. Общее решение данного уравнения имеет вид

$$y(x)=C_1x+C_2x\int\frac{e^x}{x^2}dx$$

(написанный справа интеграл не выражается в элементарных функциях).●

5. Решение неоднородного уравнения методом вариации произвольных постоянных

Рассмотрим линейное неоднородное уравнение

$$y''+p(x)y'+q(x)y=f(x). \tag{10}$$

Соответствующее ему однородное уравнение (7) было подробно изучено в предыдущем пункте.

Пусть нам известна какая-то фундаментальная система частных решений $y_1(x)$ и $y_2(x)$ однородного уравнения. Тогда общее решение однородного уравнения, как мы знаем, имеет вид $C_1y_1+C_2y_2$. Для отыскания же общего решения уравнения (10) теперь достаточно найти какое-нибудь частное решение этого уравнения. Это вытекает из следующей теоремы.

Теорема 6. *Общее решение неоднородного уравнения* (10) *есть сумма частного решения* $y_*(x)$ *этого уравнения и общего решения соответствующего однородного уравнения* (7).

Доказательство этой теоремы ничем не отличается от доказательства аналогичной теоремы из § 73.

Итак, требуется найти частное решение уравнения (10). Будем искать такое решение в виде

$$y_*=u_1(x)y_1(x)+u_2(x)y_2(x),$$

т. е. в виде линейной комбинации функций $y_1(x)$ и $y_2(x)$, но не с постоянными коэффициентами C_1 и C_2, а с переменными коэффициентами $u_1(x)$ и $u_2(x)$ (отсюда и название «метод вариации постоянных»). Так как одно уравнение связывает две неизвестные функции (u_1 и u_2), то имеется возможность по ходу решения наложить на функции u_1 и u_2 еще одно ограничение, чем мы вскоре и воспользуемся. Дифференцируя равенство $y_*=u_1y_1+u_2y_2$, находим

$$y'_*=u_1y'_1+u_2y'_2+u'_1y_1+u'_2y_2.$$

Используем теперь возможность, о которой говорилось выше: введем ограничение на выбор неизвестных функций u_1 и u_2. В качестве такого ограничения примем условие

$$u'_1y_1+u'_2y_2=0.$$

Тогда получим

$$y'_*=u_1y'_1+u_2y'_2.$$

Дифференцируя еще раз, имеем

$$y''_*=u_1y''_1+u_2y''_2+u'_1y'_1+u'_2y'_2.$$

Подставляя теперь выражения для y, y', y'' в уравнение (10), после очевидных преобразований получим

$$u_1(y''_1+p(x)y'_1+q(x)y_1)+u_2(y''_2+p(x)y'_2+q(x)y_2)+u'_1y'_1+u'_2y'_2=f(x).$$

Оба выражения, заключенные в скобки, равны нулю, так как y_1 и y_2 являются решениями однородного уравнения. Следовательно, приходим к соотношению

$$u_1'y_1'+u_2'y_2'=f(x).$$

Итак, неизвестные функции $u_1(x)$ и $u_2(x)$ должны удовлетворять системе уравнений

$$\begin{cases} u_1'y_1+u_2'y_2=0, \\ u_1'y_1'+u_2'y_2'=f(x). \end{cases} \tag{11}$$

Пользуясь тем, что определитель из коэффициентов при u_1' и u_2', равный $\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$, есть определитель Вронского для функций y_1, y_2 и потому отличен от нуля (по условию y_1 и y_2 образуют фундаментальную систему решений однородного уравнения), находим из соотношений (11) функции u_1' и u_2'. Затем простым интегрированием находим сами функции u_1 и u_2, а вслед за ними и искомое решение $y_*=u_1y_1+u_2y_2$.

Пример. Найти частное решение уравнения $y''-\frac{y'}{x}=1$ $(x>0)$, используя тот факт, что однородное уравнение $y''-\frac{y'}{x}=0$ имеет линейно независимые частные решения $y_1(x)=1$, $y_2(x)=x^2$.

○ Представим искомое решение y в виде $y=u_1\cdot 1+u_2x^2$. Для нахождения u_1 и u_2 имеем систему уравнений (11), которая в данном случае принимает вид

$$\begin{cases} u_1'+u_2'x^2=0, \\ u_2'\cdot 2x=1. \end{cases}$$

Отсюда $u_2'=\frac{1}{2x}$, $u_1'=-\frac{1}{2}x$; следовательно,

$$u_2=\frac{1}{2}\ln x,\ u_1=-\frac{1}{4}x^2.$$

Искомое решение имеет вид $y_*(x)=-\frac{1}{4}x^2+\frac{1}{2}x^2\ln x$. ●

§ 75. Линейные дифференциальные уравнения с постоянными коэффициентами

1. Линейные однородные дифференциальные уравнения второго порядка с постоянными коэффициентами. Характеристическое уравнение

Фактическое решение однородного линейного дифференциального уравнения второго порядка требует знания какой-нибудь фундаментальной системы частных решений. Если коэффициенты уравнения не постоянны, т. е. действительно зависят от x, то нахождение такой системы представляет, вообще говоря, трудную задачу. Значительно проще обстоит дело в случае *уравнения с постоянными коэффициентами,* т. е. уравнения вида

$$y''+py'+qy=0, \tag{1}$$

где p и q — постоянные. Для этого случая можно указать простой способ построения фундаментальной системы решений.

Будем искать частное решение уравнения (1) в виде показательной функции $y=e^{\lambda x}$. Дифференцируя дважды функцию y, получим $y'=\lambda e^{\lambda x}$, $y''=\lambda^2e^{\lambda x}$.

Подставляя выражения для функции y и ее производных в уравнение (1), приходим к соотношению

$$(\lambda^2+p\lambda+q)e^{\lambda x}=0.$$

Так как $e^{\lambda x}\neq 0$, то полученное соотношение равносильно уравнению

$$\lambda^2+p\lambda+q=0. \qquad (2)$$

Определение. Алгебраическое уравнение (2) называется *характеристическим уравнением* для дифференциального уравнения (1).

Дальнейшая схема построения фундаментальной системы решений для уравнения (1) такова. Алгебраическое уравнение (2) имеет два корня — действительных или комплексных. Обозначим их λ_1 и λ_2. Таким образом, каждая из функций $y_1=e^{\lambda_1 x}$, $y_2=e^{\lambda_2 x}$ является решением уравнения (1). Если эти функции линейно независимы, то общее решение уравнения, согласно теореме 5 § 74, имеет вид $C_1e^{\lambda_1 x}+C_2e^{\lambda_2 x}$. В случае линейной зависимости указанных функций необходимы дополнительные рассуждения. Более подробно этот вопрос будет обсужден в следующем пункте.

2. Построение общего решения

Рассмотрим все случаи, которые могут представиться при решении характеристического уравнения (2).

I случай: корни λ_1 и λ_2 — действительные и различные. Соответствующие им решения $y_1=e^{\lambda_1 x}$ и $y_2=e^{\lambda_2 x}$ линейно независимы. Действительно, их определитель Вронского

$$W(y_1,y_2)=\begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x}\end{vmatrix}=e^{\lambda_1 x}e^{\lambda_2 x}\begin{vmatrix} 1 & 1 \\ \lambda_1 & \lambda_2\end{vmatrix}=(\lambda_2-\lambda_1)e^{(\lambda_1+\lambda_2)x},$$

ввиду $\lambda_1\neq\lambda_2$ отличен от нуля. Следовательно, y_1 и y_2 образуют фундаментальную систему решений.

Пример 1. Найти общее решение уравнения $y''-5y'+6y=0$.

○ Характеристическое уравнение имеет вид $\lambda^2-5\lambda+6=0$. Его корнями являются числа $\lambda_1=2$ и $\lambda_2=3$. Следовательно, имеем линейно независимые частные решения $y=e^{2x}$ и $y_2=e^{3x}$. Общее решение запишется в виде $y=C_1e^{2x}+Ce^{3x}$. ●

II случай: корни λ_1 и λ_2 — комплексно сопряженные: $\lambda_1=\alpha+\beta i$ и $\lambda_2=\alpha-\beta i$, где $\beta\neq 0$. Соответствующие им комплексные решения обозначим z_1 и z_2:

$$z_1=e^{(\alpha+\beta i)x},\quad z_2=e^{(\alpha-\beta i)x}.$$

Они линейно независимы, поскольку $\lambda_1\neq\lambda_2$. Используя формулу Эйлера (см. п. 8 § 67), можем записать

$$z_1=e^{\alpha x}(\cos\beta x+i\sin\beta x),\quad z_2=e^{\alpha x}(\cos\beta x-i\sin\beta x),$$

откуда видно, что при любом x числа z_1 и z_2 сопряжены. Составим линейные комбинации

$$y_1=\frac{1}{2}(z_1+z_2)=e^{\alpha x}\cos\beta x,\quad y_2=\frac{1}{2i}(y_1-y_2)=e^{\alpha x}\sin\beta x.$$

Функции y_1 и y_2 являются действительными решениями уравнения (1). Эти решения также линейно независимы: в противном случае мы имели бы тождественное равенство $y_1=Cy_2$ (или $y_2=Cy_1$), откуда следовала бы линейная зависимость между z_1 и z_2. Итак, комплексно сопряженным корням $\alpha+\beta i$, $\alpha-\beta i$ характеристического уравнения можно сопоставить два линейно независимых частных решения $y_1=e^{\alpha x}\cos\beta x$, $y_2=e^{\alpha x}\sin\beta x$. Следовательно, общее решение имеет вид

$$y=C_1e^{\alpha x}\cos\beta x+C_2e^{\alpha x}\sin\beta x=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x).$$

Пример 2. Найти общее решение уравнения $y''-4y'+5y=0$.

○ Характеристическое уравнение есть $\lambda_2-4\lambda+5=0$. Его корни $\lambda_1=2+i$ и $\lambda_2=2-i$. Таким образом, имеем частные решения $y_1=e^{2x}\cos x$, $y_2=e^{2x}\sin x$. Общее решение записывается в виде $y=e^{2x}(C_1\cos x+C_2\sin x)$. ●

III случай: корни λ_1 и λ_2 — равные, а значит, действительные. Будем рассуждать следующим образом (хотя это рассуждение и не имеет силы доказательства). Изменим незначительно коэффициенты p и q уравнения так, чтобы вместо двух равных корней λ_1 и λ_2 получились два неравных (но близких) корня λ_1^* и λ_2^*. Соответствующие решения $e^{\lambda_1^* x}$ и $e^{\lambda_2^* x}$ являются различными. Составим из них линейную комбинацию

$$\frac{e^{\lambda_2^* x}-e^{\lambda_1^* x}}{\lambda_2^*-\lambda_1^*}. \tag{3}$$

Если теперь представить, что коэффицненты уравнения (1) возвращаются к своим прежним значениям, т. е. $\lambda_1^*\to\lambda_1$, $\lambda_2^*\to\lambda_2$, то из (3) в пределе получим решение

$$y(x)=\lim_{\substack{\lambda_1^*\to\lambda_1\\ \lambda_2^*\to\lambda_2}}\frac{e^{\lambda_2^* x}-e^{\lambda_1^* x}}{\lambda_2^*-\lambda_1^*} \tag{4}$$

исходного дифференциального уравнения (1). Для нахождения правой части соотношения (4) воспользуемся тем, что по теореме Лагранжа для любых a и b имеет место равенство $e^a-e^b=(a-b)e^{\xi}$, где точка ξ находится между a и b. В частности, $e^{\lambda_2^* x}-e^{\lambda_1^* x}=(\lambda_2^* x-\lambda_1^* x)e^{\lambda x}$, где λ находится между λ_1^* и λ_2^*. Поэтому правая часть соотношения (4) равна $xe^{\lambda_1 x}$. Таким образом, кроме решения $e^{\lambda_1 x}$, имеем решение $xe^{\lambda_1 x}$.

Итак, равным корням $\lambda_1=\lambda_2$ характеристического уравнения можно сопоставить два частных решения $e^{\lambda_1 x}$, $xe^{\lambda_1 x}$. Они линейно независимы (проверьте это самостоятельно); следовательно, общее решение

$$y=C_1e^{\lambda_1 x}+C_2xe^{\lambda_1 x}=e^{\lambda_1 x}(C_1+C_2x).$$

Пример 3. Найти общее решение уравнения $y''-2y'+y=0$.

○ Характеристическое уравнение есть $\lambda^2-2\lambda+1=0$, или $(\lambda-1)^2=0$. Его корни $\lambda_1=\lambda_2=1$. Общее решение уравнения имеет вид $y=e^x(C_1+C_2x)$. ●

3. Неоднородное уравнение

Рассмотрим теперь уравнение

$$y''+py'+qy=f(x), \tag{5}$$

где $f(x)$ — некоторая заданная функция; коэффициенты p и q по-прежнему считаем постоянными. Согласно теореме 6 § 73, для построения общего решения уравнения (5) достаточно найти какое-нибудь частное решение этого уравнения, а также общее решение соответствующего однородного уравнения. Поскольку второе мы уже умеем делать, задача сводится к нахождению частного решения уравнения (5).

Для решения этой задачи можно воспользоваться методом вариации произвольных постоянных. Однако во многих важных для практики случаях имеется и более простой способ. Он применим в случае, когда $f(x)$ есть функция вида $P(x)e^{\alpha x}$, где $P(x)$ — многочлен. Укажем суть этого способа, не вникая в его обоснование.

I случай: число α не является корнем характеристического уравнения. Тогда решение нужно искать в виде

$$y=Q(x)e^{\alpha x},$$

где $Q(x)$ — многочлен той же степени, что и $P(x)$. Записав $Q(x)$ в виде многочлена с неопределенными коэффициентами и подставив выражение для y в уравнение (5), после сокращения обеих частей на $e^{\alpha x}$ получаем равенство двух многочленов (из которых один есть $P(x)$). Приравняв коэффициенты при одинаковых степенях x слева и справа, получим систему уравнений, из которой найдем коэффициенты многочлена $Q(x)$.

Особо отметим два частных случая:

1) $f(x)=P(x)$, т. е. правая часть уравнения (5) представляет собой многочлен от x. В этом случае имеем $\alpha=0$; следовательно, если число 0 не является корнем характеристического уравнения, то решение ищем в виде $y=Q(x)$;

2) $f(x)=e^{\alpha x}$. Тогда $P(x)=1$ есть многочлен нулевой степени, а значит, $Q(x)=c=\text{const}$. Если при этом число α не является корнем характеристического уравнения, то решение ищем в виде $y=ce^{\alpha x}$.

Пример 1. Найти частное решение уравнения $y''-5y'+6y=(12x-55)e^{-x}$.

○ В данном случае $\alpha=-1$. Корнями характеристического уравнения являются 2 и 3; число α не совпадает ни с одним из них. Поэтому решение ищем в виде $y=(ax+b)e^{-x}$. Дифференцируя выражение для y, находим

$$y'=ae^{-x}-(ax+b)e^{-x}=(-ax+a-b)e^{-x},$$
$$y''=-ae^{-x}-(-ax+a-b)e^{-x}=(ax-2a+b)e^{-x}.$$

Подставляя y, y', y'' в уравнение и сокращая обе части на e^{-x}, приходим к тождественному равенству

$$(ax-2a+b)-5(-ax+a-b)+6(ax+b)=12x-55, \text{ или } 12ax-7a+12b=12x-55,$$

откуда $12a=12$, $-7a+12b=-55$. Решая эту систему, находим $a=1$, $b=-4$. Следовательно, искомое частное решение данного уравнения есть $y=(x-4)e^{-x}$. ●

II случай: один из корней характеристического уравнения равен α, а второй корень отличен от α. Тогда частное решение следует искать в виде

$$y=xQ(x)e^{\alpha x}.$$

Пример 2. Найти частное решение уравнения $y''-2y'-3y=(8x+10)e^{3x}$.

○ Здесь $\alpha=3$. Корнями характеристического уравнения являются -1 и 3, один из них совпадает с α. Поэтому решение ищем в виде $y=x(ax+b)e^{3x}$. Находим

$$y'=(2ax+b)e^{3x}+3(ax^2+bx)e^{3x}=(3ax^2+(2a+3b)x+b)e^{3x},$$
$$y''=(6ax+(2a+3b))e^{3x}+3(3ax^2+(2a+3b)x+b)e^{3x}=(9ax^2+(12a+9b)x+(2a+6b))e^{3x}$$

и подставляем выражения для y, y', y'' в уравнение. Приходим к тождеству

$$9ax^2+(12a+9b)x+2a+6b-2(3ax^2+(2a+3b)x+b)-3(ax^2+bx)=8x+10,$$

или $8ax+2a+4b=8x+10$, откуда $8a=8$, $2a+4b=10$. Решение этой системы есть $a=1$, $b=2$. Итак, частное решение данного уравнения имеет вид $y=x(x+2)e^{3x}$. ●

III случай: оба корня характеристического уравнения равны α. Тогда частное решение следует искать в виде

$$y=x^2Q(x)e^{\alpha x}.$$

Пример 3. Найти частное решение уравнения $y''+2y'+y=3e^{-x}$.

○ Здесь $\alpha=-1$ и оба корня характеристического уравнения также равны -1. Поэтому решение ищем в виде $y=ax^2e^{-x}$. Проведя такие же вычисления, что и в примерах 1 и 2, получим $a=1,5$. Искомое решение имеет вид $y=1,5x^2e^{-x}$. ●

§ 76. Упругие колебания материальной точки

Линейные дифференциальные уравнения порядка выше, чем 1, находят многочленные применения в физике. В качестве примера рассмотрим задачу о движении материальной точки, подвешенной на пружине (рис. 260).

Массу материальной точки обозначим m. На вертикальной оси, вдоль которой движется точка m, выберем начало координат O, которое соответствует положению массы m при недеформированной (т. е. несжатой и нерастянутой) пружине.

Закон движения точки m задается равенством $x = x(t)$. Согласно второму закону Ньютона, произведение $m\frac{d^2x}{dt^2}$ должно быть равно сумме всех сил, действующих на точку m (в вертикальном направлении). Одной из таких сил является сила упругости пружины, которая по закону Гука равна $-cx$, где c — коэффициент упругости пружины; знак минус поставлен потому, что сила упругости направлена в сторону, противоположную смещению точки. Кроме силы упругости, на точку m действует сила тяжести mg. Возможны и какие-то внешние силы, которые будем считать зависящими от времени; сумму этих сил обозначим $F(t)$. Тогда для функции $x(t)$ имеем уравнение

$$m\frac{d^2x}{dt^2} = -cx + mg + F(t).$$

Полагая $c/m = \omega^2$, перепишем это уравнение в виде

$$\frac{d^2x}{dt^2} + \omega^2 x = g + \frac{1}{m}F(t). \tag{1}$$

Это линейное дифференциальное уравнение второго порядка с постоянными коэффициентами, вообще говоря, неоднородное (если $g + \frac{1}{m}F(t) \neq 0$). Соответствующее ему однородное уравнение

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \tag{2}$$

имеет фундаментальную систему решений $x_1 = \cos \omega t$, $x_2 = \sin \omega t$.

Изучим решения уравнения (1) в некоторых частных случаях.

I случай. Сила $F(t)$ постоянна: $F(t) = F_0$. Тогда имеем уравнение

$$\frac{d^2x}{dt^2} + \omega^2 x = g + \frac{F_0}{m} \tag{3}$$

с постоянной правой частью. Решим его при нулевых начальных условиях

$$x(0) = 0, \ x'(0) = 0, \tag{4}$$

т. е. когда смещение точки m, а также ее скорость в начальный момент равны нулю.

Общее решение однородного уравнения (2) имеет вид $C_1 \cos \omega t + C_2 \sin \omega t$. Нетрудно найти частное решение уравнения (3), правая часть которого есть постоянная $g + \frac{F_0}{m}$. Таким решением является постоянная функция $\frac{1}{\omega^2}\left(g + \frac{F_0}{m}\right)$. Отсюда находим общее решение уравнения (3):

$$x(t) = C_1 \cos \omega t + C_2 \sin \omega t + \frac{1}{\omega^2}\left(g + \frac{F_0}{m}\right).$$

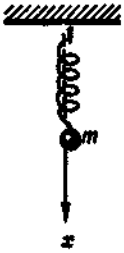


Рис. 260

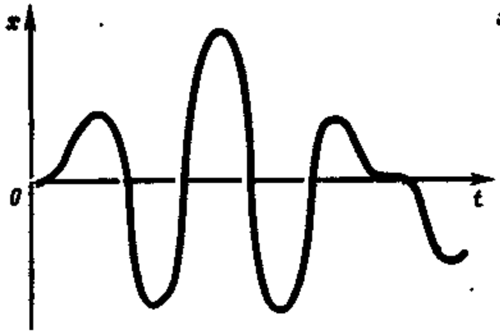


Рис. 261

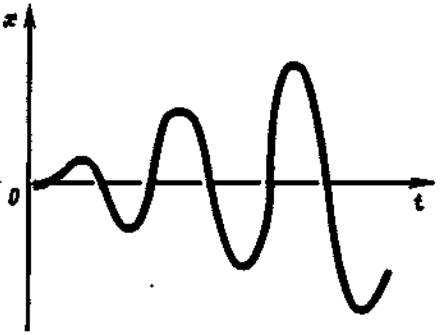


Рис. 262

Теперь остается подобрать значения постоянных C_1 и C_2 так, чтобы функция $x(t)$ удовлетворяла начальным условиям (4). Имеем

$$x(0)=C_1+\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right)=0,\quad x'(0)=\omega C_2=0,$$

откуда $C_1=-\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right)$, $C_2=0$. Следовательно, получаем решение

$$x(t)=-\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right)\cos\omega t+\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right).$$

Таким образом, материальная точка совершает гармоническое колебание с частотой ω $\left(\text{равной } \sqrt{\frac{c}{m}}\right)$ и амплитудой $\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right)$; центром колебания является точка

$$\frac{1}{\omega^2}\left(g+\frac{F_0}{m}\right)=\frac{m}{c}\left(g+\frac{F_0}{m}\right)=\frac{1}{c}(mg+F_0),$$

отвечающая статическому растяжению пружины силой $mg+F_0$.

II случай. Сила $F(t)$ изменяется по периодическому закону: $F(t)=F_0\sin\lambda t$, причем $\lambda\neq\omega$. Тогда уравнение (1) примет вид

$$\frac{d^2x}{dt^2}+\omega^2x=g+\frac{F_0}{m}\sin\lambda t. \tag{5}$$

Частное решение уравнения (5) можно найти, используя метод вариации произвольных постоянных. Однако проще с самого начала искать его в виде гармонического колебания с той же частотой λ, с какой изменяется внешняя сила $F(t)$, точнее, в виде

$$x_*(t)=a\cos\lambda t+b\sin\lambda t+c,$$

где a, b, c — постоянные (такой способ отыскания частного решения можно обосновать, но сейчас это не столь важно). Подставляя выражение для $x_*(t)$ в уравнение (5), получим

$$(\omega^2-\lambda^2)\,x_*(t)=g+\frac{F_0}{m}\sin\lambda t,$$

т. е. $a=0$, $b=\frac{F_0}{m(\omega^2-\lambda^2)}$, $c=\frac{g}{\omega^2-\lambda^2}$. Итак, общее решение уравнения (5) записывается в виде

$$x(t)=C_1\cos\omega t+C_2\sin\omega t+\frac{F_0}{m(\omega^2-\lambda^2)}\sin\lambda t+\frac{g}{\omega^2-\lambda^2}. \qquad (6)$$

Используя начальные условия $x(0)=0$ и $x'(0)=0$, находим значения постоянных C_1 и C_2:

$$x_*(0)=C_1+\frac{g}{\omega^2-\lambda^2}=0,\ x'(0)=\omega C_2+\frac{\lambda F_0}{m(\omega^2-\lambda^2)}=0,$$

т. е. $C_1=-\frac{g}{\omega^2-\lambda^2}$, $C_2=-\frac{\lambda F_0}{\omega m(\omega^2-\lambda^2)}$.

Подставляя эти значения в равенство (6), находим искомое решение:

$$x_*(t)=\frac{1}{\omega^2-\lambda^2}\left(-g\cos\omega t-\frac{\lambda F_0}{\omega m}\sin\omega t+\frac{F_0}{m}\sin\lambda t+g\right).$$

Примерный график функции $x_*(t)$ изображен на рис. 261. Функция $x_*(t)$, вообще говоря, не является периодической. Лишь в том случае, когда числа ω по λ соизмеримы (т. е. когда отношение ω/λ рационально), $x_*(t)$ имеет некоторый период τ.

Выражение для $x_*(t)$ содержит множитель $\frac{1}{\omega^2-\lambda^2}$. Отсюда ясно, что в случае близости частот ω и λ функция $x_*(t)$ может принимать весьма большие значения.

III случай. Пусть $F(t)=F_0\sin\omega t$, т. е. частота внешней силы $F(t)$ совпадает с «собственной» частотой ω («собственной» называется частота колебаний, обусловленных только силой упругости пружины). Такой случай носит в физике название *резонанса*. Тогда уравнение (1) принимает вид

$$\frac{d^2x}{dt^2}+\omega^2x=g+\frac{F_0}{m}\sin\omega t. \qquad (7)$$

Можно показать (сделайте это самостоятельно), что уравнение (7) имеет частное решение вида $at\sin\omega t+b$, где a и b — некоторые постоянные числа. После этого снова остается так подобрать значения C_1 и C_2, чтобы функция

$$x_*(t)=C_1\cos\omega t+C_2\sin\omega t+at\sin\omega t+b$$

удовлетворяла начальным условиям (4).

Примерный график функции $x_*(t)$ изображен на рис. 262. Характерным для поведения $x_*(t)$ является то, что с каждым новым «всплеском» графика амплитуда становится все больше (причиной этого является множитель t перед $\sin\omega t$). Размах «колебания» (если так можно назвать изменение $x_*(t)$) может в этом случае достичь размеров, представляющих опасность для конструкции.

§ 77. Системы дифференциальных уравнений

1. Общие сведения о системах дифференциальных уравнений

Решение той или иной задачи может потребовать нахождения не одной, а сразу нескольких неизвестных функций. Для этого необходимо располагать, вообще говоря, таким же числом уравнений. Если каждое из этих уравнений является дифференциальным, т. е. имеет вид соотношения, связывающего неизвестные функции и их производные, то говорят о *системе дифференциальных уравнений*. Ограничимся рассмотрением систем дифференциальных уравнений первого порядка: это означает, что в уравнения не входят производные порядка выше, чем 1.

При рассмотрении систем дифференциальных уравнений аргумент обозначают, как правило, через t, а сами неизвестные функции — через $x_1(t)$, $x_2(t)$ и т. д. Так, система дифференциальных уравнений первого порядка с двумя неизвестными функциями записывается обычно в виде

$$\begin{cases} \varphi\left(t,\ x_1,\ x_2,\ \dfrac{dx_1}{dt},\ \dfrac{dx_2}{dt}\right)=0, \\ \psi\left(t,\ x_1,\ x_2,\ \dfrac{dx_1}{dt},\ \dfrac{dx_2}{dt}\right)=0. \end{cases} \tag{1}$$

На системы дифференциальных уравнений естественным образом обобщается постановка задачи Коши для одного уравнения. Например, в случае системы (1) задача Коши состоит в нахождении решения $x_1(t)$, $x_2(t)$, удовлетворяющего начальным условиям $x_1(t_0)=x_1^0$, $x_2(t_0)=x_2^0$, где t_0, x_1^0, x_2^0 — заданные числа. Для случая системы может быть доказана теорема существования и единственности решения задачи Коши, аналогичная теореме из § 71.

К системам дифференциальных уравнений первого порядка в известном смысле сводятся уравнения (и системы уравнений) любого порядка. Проиллюстрируем это на примере уравнения третьего порядка. Пусть дано уравнение

$$y'''=f(x,\ y,\ y',\ y'').$$

Если обозначить функции y' и y'' соответственно через u и v, то уравнение можно заменить системой

$$\begin{cases} y'=u, \\ u'=v, \\ v'=f(x,\ y,\ u,\ v), \end{cases}$$

состоящей из трех уравнений первого порядка с тремя неизвестными функциями $y(x)$, $u(x)$, $v(x)$. Аналогичное истолкование допускает любое другое дифференциальное уравнение (или система уравнений).

2. Системы линейных дифференциальных уравнений первого порядка. Сведение к одному уравнению более высокого порядка

Ограничимся рассмотрением систем дифференциальных уравнений специального вида, называемых *линейными системами*. В случае двух неизвестных функций $x_1(t)$, $x_2(t)$ линейная система имеет вид

$$\begin{cases} \dfrac{dx_1}{dt}=a_{11}x_1+a_{12}x_2, \\ \dfrac{dx_2}{dt}=a_{21}x_1+a_{22}x_2, \end{cases} \tag{2}$$

где коэффициенты a_{ij} являются, вообще говоря, функциями независимой переменной t. Будем считать эти функции непрерывными; тогда для заданной системы заведомо выполняются условия теоремы о существовании и единственности решения задачи Коши.

Один из методов интегрирования системы (2) заключается в сведении системы к одному уравнению второго порядка с одной неизвестной функцией (о сведении одного уравнения произвольного порядка к системе уравнений первого порядка было сказано в п. 1). Дифференцируя (по t) обе части первого уравнения системы (2), находим

$$\frac{d^2x_1}{dt^2}=\alpha_{11}\frac{dx_1}{dt}+\alpha_{12}\frac{dx_2}{dt}+\frac{d\alpha_{11}}{dt}x_1+\frac{d\alpha_{12}}{dt}x_2,$$

откуда, заменяя производные $\frac{dx_1}{dt}$, $\frac{dx_2}{dt}$ их выражениями из самой системы, имеем

$$\frac{d^2x_1}{dt^2}=\alpha_{11}(\alpha_{11}x_1+\alpha_{12}x_2)+\alpha_{12}(\alpha_{21}x_1+\alpha_{22}x_2)+\frac{d\alpha_{11}}{dt}x_1+\frac{d\alpha_{12}}{dt}x_2.$$

Группируя в правой части все члены с x_1, а также с x_2, получим уравнение вида

$$\frac{d^2x_1}{dt^2}=\beta_1x_1+\beta_2x_2, \tag{3}$$

где коэффициенты β_1 и β_2 определенным образом выражаются через коэффициенты α_{ij} и их производные (записывать эти выражения не будем). Комбинируя уравнение (3) с первым уравнением системы (2), получаем

$$\begin{cases} \frac{dx_1}{dt}=\alpha_{11}x_1+\alpha_{12}x_2, \\ \frac{d^2x_1}{dt^2}=\beta_1x_1+\beta_2x_2. \end{cases} \tag{4}$$

Предположим, что в рассматриваемой области изменения t определитель $\begin{vmatrix}\alpha_{11} & \alpha_{12}\\ \beta_1 & \beta_2\end{vmatrix}$ отличен от нуля. Тогда систему (4) можно решить относительно x_1 и x_2, т. е. выразить x_1 и x_2 через $\frac{dx_1}{dt}$ и $\frac{d^2x_1}{dt^2}$. В результате приходим к уравнениям вида

$$x_1=a\frac{dx_1}{dt}+b\frac{d^2x_1}{dt^2}, \tag{5}$$

$$x_2=c\frac{dx_1}{dt}+d\frac{d^2x_1}{dt^2} \tag{6}$$

(выражения для a, b, c, d приводить не будем). Первое из них представляет собой линейное дифференциальное уравнение второго порядка с одной неизвестной функцией $x_1(t)$. К нему приложима вся теория, изложенная в § 74. Заметим, что если в исходной системе (2) все коэффициенты α_{ij} постоянны, то уравнение (5) также является уравнением с постоянными коэффициентами; для решения таких уравнений имеется эффективный метод (см. § 75).

Пример. Решить систему
$$\begin{cases} \frac{dx_1}{dt}=3x_1-2x_2, \\ \frac{dx_2}{dt}=2x_1-x_2. \end{cases}$$

○ Дифференцируя обе части первого уравнения, имеем

$$\frac{d^2x_1}{dt^2}=3\frac{dx_1}{dt}-2\frac{dx_2}{dt}=3(3x_1-2x_2)-2(2x_1-x_2)=5x_1-4x_2.$$

В комбинации с первым уравнением данной систем это приводит к системе

$$\begin{cases} \frac{dx_1}{dt}=3x_1-2x_2, \\ \frac{d^2x_1}{dt}=5x_1-4x_2. \end{cases}$$

Отсюда находим выражения для x_1 и x_2 через $\frac{dx_1}{dt}$ и $\frac{d^2x_1}{dt^2}$:

$$x_1 = 2\frac{dx_1}{dt} - \frac{d^2x_1}{dt^2}, \tag{7}$$

$$x_2 = \frac{5}{2}\frac{dx_1}{dt} - \frac{3}{2}\frac{d^2x_1}{dt^2}. \tag{8}$$

В результате приходим к уравнению второго порядка для неизвестной функции $x_1(t)$: $\frac{d^2x_1}{dt^2} - 2\frac{dx_1}{dt} + x_1 = 0$. Решая это уравнение известным способом, получим $x_1 = (C_1 + C_2t)e^t$, после чего из выражения (8) находим $x_2 = \frac{1}{2}(2C_1 - C_2 + 2C_2t)e^t$. ●

3. Матричная запись системы

При изучении линейных систем удобно использовать матричные обозначения. Условимся для краткости записывать $\dot{x}$ вместо $\frac{dx}{dt}$. Далее, введем матрицы

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \dot{X} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix}.$$

Тогда систему (2) можно записать в виде одного матричного уравнения

$$\dot{X} = AX. \tag{9}$$

Например, система

$$\begin{cases} \dot{x}_1 = 3x_1 - 5x_2, \\ \dot{x}_2 = 4x_1 - 2x_2. \end{cases}$$

в матричной записи выглядит так:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

4. Линейная система дифференциальных уравнений с постоянными коэффициентами. Нахождение решений вида $x_i = e^{\lambda t}$ $(i = 1, 2, \ldots)$

Предположим, что коэффициенты α_{ij} в системе (2) постоянны. В этом случае для решения системы можно использовать методы линейной алгебры. Дадим некоторое представление о таких методах, ограничиваясь системами двух уравнений с двумя неизвестными, т. е. системами вида (2).

Прежде всего заметим, что система (2) имеет очевидное частное решение $x_1(t) \equiv 0$, $x_2(t) \equiv 0$. Это решение называется *нулевым*. Интерес представляют, конечно, ненулевые решения. Будем искать такие решения в виде $x_1 = p_1e^{\lambda t}$, $x_2 = p_2e^{\lambda t}$ или, используя матричную запись, в виде $X = Pe^{\lambda t}$, где $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ — матрица (вектор) с постоянными элементами p_1, p_2, не равными одновременно нулю.

Имеем, очевидно, $\dot{X} = \lambda Pe^{\lambda t}$. Подставляя выражения для X и $\dot{X}$ в уравнение (9), получим $\lambda Pe^{\lambda t} = APe^{\lambda t}$, откуда после сокращения на $e^{\lambda t}$ находим

$$AP = \lambda P. \tag{10}$$

Итак, для отыскания пары λ, P необходимо решить уравнение (10).

5. Собственные векторы и собственные значения матрицы

Определение. Пусть A — квадратная матрица размера $n\times n$. Если ненулевой вектор $P\in\mathbb{R}^n$ и число λ удовлетворяют соотношению (10), то λ называется *собственным значением* матрицы A, а вектор P — *собственным вектором*, отвечающим (соответствующим, принадлежащим) собственному значению λ.

Укажем способ отыскания собственных значений и собственных векторов. Соотношение (10) в подробной записи означает

$$\begin{cases}\alpha_{11}p_1+\alpha_{12}p_2=\lambda p_1,\\ \alpha_{21}p_1+\alpha_{22}p_2=\lambda p_2\end{cases}\quad\text{или}\quad\begin{cases}(\alpha_{11}-\lambda)p_1+\alpha_{12}p_2=0,\\ \alpha_{21}p_1+(\alpha_{22}-\lambda)p_2=0.\end{cases}\tag{11}$$

Таким образом, для нахождения пары λ, P (собственного значения и собственного вектора) имеет систему (11), содержащую два уравнения с тремя неизвестными λ, p_1, p_2; при этом нас интересуют лишь такие решения, где хотя бы одно из чисел p_1, p_2 отлично от нуля.

Для неизвестного λ можно получить отдельное уравнение. Для этого примем во внимание следующее обстоятельство. По отношению к неизвестным p_1, p_2 соотношения (11) представляют собой систему двух (алгебраических) однородных линейных уравнений с двумя неизвестными. Нас интересуют ненулевые решения p_1, p_2 этой системы. Как показывается в курсе линейной алгебры, однородная система 2×2 (вообще, система $n\times n$) имеет ненулевое решение тогда и только тогда, когда ее определитель равен нулю. Следовательно, для того чтобы число λ было собственным значением матрицы A, необходимо и достаточно выполнение равенства

$$\begin{vmatrix}\alpha_{11}-\lambda & \alpha_{12}\\ \alpha_{21} & \alpha_{22}-\lambda\end{vmatrix}=0.\tag{12}$$

Равенство (12) называется *характеристическим уравнением для матрицы A*. Это квадратное уравнение относительно λ (а в случае матрицы A размера $n\times n$ — уравнение n-й степени). В подробной записи оно имеет вид

$$(\alpha_{11}-\lambda)(\alpha_{22}-\lambda)-\alpha_{12}\alpha_{21}=0\ \text{ или }\ \lambda^2-(\alpha_{11}+\alpha_{22})\lambda+(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})=0.\tag{13}$$

Его корни могут быть как действительными, так и комплексно сопряженными.

Если найдено некоторое собственное значение λ_0 (корень характеристического уравнения), то для отыскания соответствующих ему собственных векторов P используются соотношения (11) — система двух линейных однородных уравнений с двумя неизвестными. *Любое ненулевое решение этой системы дает собственный вектор P, отвечающий собственному значению λ_0, и все собственные векторы, отвечающие λ_0, находятся таким способом.*

Пример. Найти собственные значения и собственные векторы матрицы

$$A=\begin{pmatrix}7 & 6\\ -1 & 2\end{pmatrix}.$$

○ Составляем характеристическое уравнение:

$$\begin{vmatrix}7-\lambda & 6\\ -1 & 2-\lambda\end{vmatrix}=0,\ \text{ или }\ \lambda^2-9\lambda+20=0.$$

Его корни $\lambda_1=5$, $\lambda_2=4$ и являются собственными значениями матрицы A.

Собственные векторы, отвечающие λ_1, находим из системы (11), которая в данном случае имеет вид

$$\begin{cases}(7-5)p_1+6p_2=0,\\ -p_1+(2-5)p_2=0\end{cases}\quad\text{или}\quad\begin{cases}2p_1+6p_2=0,\\ -p_1-3p_2=0.\end{cases}$$

Здесь второе уравнение лишь повторяет первое, поэтому его можно отбросить. Остается уравнение $2p_1+6p_2=0$, или $p_1=-3p_2$. Его ненулевые решения имеют вид $p_2=s$, $p_1=-3s$, где s — любое не равное нулю число. Итак, собственные векторы, отвечающие λ_1, имеют вид $s\begin{pmatrix}-3\\1\end{pmatrix}$, где $s\neq0$. Аналогично найдем, что собственными векторами, отвечающими λ_2, являются $s\begin{pmatrix}-2\\1\end{pmatrix}$, где $s\neq0$. ●

6. Нахождение общего решения линейной системы с постоянными коэффициентами

Пусть λ_1 и λ_2 — корни характеристического уравнения (12). Предположим сначала, что $\lambda_1\neq\lambda_2$. Если P_1 — какой-нибудь собственный вектор, отвечающий λ_1, а P_2 — собственный вектор, отвечающий λ_2, то формулы $X_1=P_1e^{\lambda_1t}$, $X_2=P_2e^{\lambda_2t}$ определяют два частных решения уравнения (9). Общее же решение, как можно показать, имеет вид

$$X=C_1X_1+C_2X_2,$$

где C_1, C_2 — произвольные постоянные.

Разумеется, числа λ_1 и λ_2 не обязательно являются действительными. Может оказаться, что $\lambda_1=\alpha+\beta i$, $\lambda_2=\alpha-\beta i$, где $\beta\neq0$. Пусть P — собственный вектор, отвечающий λ_1 (разумеется, комплексный). Тогда имеем $AP=\lambda_1P$. Из этого равенства следует $\bar A\bar P=\bar\lambda_1\bar P$ (черта обозначает комплексную сопряженность) или, учитывая, что матрица A состоит из действительных чисел, $A\bar P=\bar\lambda_1\bar P$. Таким образом, $A\bar P=\lambda_2\bar P$, т. е. вектор $\bar P$ является собственным, отвечающим собственному значению λ_2. Частные решения $X_1=Pe^{\lambda_1t}$, $X_2=\bar Pe^{\bar\lambda_1t}$ комплексно сопряжены. Чтобы получить действительные решения, заменим X_1 и X_2 их линейными комбинациями

$$X_1^*=X_1+X_2,\quad X_2^*=\frac{1}{2i}(X_1-X_2).$$

Наконец, возможен случай, когда корни λ_1, λ_2 характеристического уравнения совпадают. Если $\lambda_1=\lambda_2$, то вместо двух частных решений X_1 и X_2 имеем только одно решение X_1. Однако рассуждая по аналогии с § 75, можно показать, что в этом случае вектор tX_1 также является решением; таким образом, и здесь получаются два различных частных решения X_1 и tX_1.

Пример. Решить систему
$$\frac{dx_1}{dt}=7x_1+6x_2,\qquad \frac{dx_2}{dt}=-x_1+2x_2.$$

○ Здесь $A=\begin{pmatrix}7&6\\-1&2\end{pmatrix}$. Такую матрицу A мы рассмотрели в примере из п. 5, где было показано что $P_1=\begin{pmatrix}-3\\1\end{pmatrix}$, $P_2=\begin{pmatrix}-2\\1\end{pmatrix}$.

Частными решениями являются $X_1=\begin{pmatrix}-3\\1\end{pmatrix}e^{5t}$, $X_2=\begin{pmatrix}-2\\1\end{pmatrix}e^{4t}$. Общее решение в матричной записи имеет вид $X=C_1\begin{pmatrix}-3\\1\end{pmatrix}e^{5t}+C_2\begin{pmatrix}-2\\1\end{pmatrix}e^{4t}$, а в развернутой форме оно запишется следующим образом:

$$x_1(t)=-3C_1e^{5t}-2C_2e^{4t},\quad x_2(t)=C_1e^{5t}+C_2e^{4t}.\ ●$$

Основные формулы

Функции одной переменной

$a^{\log_a x}=x\ (a>0,\ a\neq 1,\ x>0)$ — определение логарифма

$\log_a(x_1x_2)=\log_a x_1+\log_a x_2\quad (x_1>0,\ x_2>0)$, $\log_a\frac{x_1}{x_2}=\log_a x_1-\log_a x_2\quad (x_1>0,\ x_2>0)$,
$\log_a x^k=k\log_a x\quad (x>0,\ k\in\mathbb{R})$, $\log_a x=\frac{\log_b x}{\log_b a}\quad (x>0)$, $\log_a x=\log_{a^k}x^k\quad (x>0,\ k\in\mathbb{R})$
— свойства логарифмов

$\sin x=a;\ x=(-1)^n\arcsin a+\pi n\quad (n\in\mathbb{Z})$, $\cos x=a;\ x=\pm\arccos a+2\pi h\quad (n\in\mathbb{Z})$,
$\operatorname{tg} x=a;\quad x=\operatorname{arctg} a+\pi n\quad (n\in\mathbb{Z})$, $\operatorname{ctg} x=a;\ x=\operatorname{arcctg} a+\pi n\quad (n\in\mathbb{Z})$
— формулы для решения простейших тригонометрических уравнений

$\lim\limits_{x\to 0}\frac{\sin x}{x}=1$ — первый замечательный предел

$\lim\limits_{x\to 0}(1+x)^{1/x}=e$ — второй замечательный предел

Производная

$y'=f'(x)=\lim\limits_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$ — определение производной функции $f(x)$ в точке x

$dy=f'(x)dx$ — дифференциал функции $f(x)$ в точке x

$(u+v)'=u'+v'$, $(uv)'=u'v+uv'$,
$(Cu)'=Cu'(C=\text{const})$, $\left(\frac{u}{v}\right)'=\frac{u'v-uv'}{v^2}(v\neq 0)$
— правила дифференцирования

$y'=f'(g(x))g'(x)$ — производная сложной функции

$\varphi'(y)=\frac{1}{f'(x)}$ — производная функции $\varphi(y)$, обратной по отношению к $f(x)$

$(x^\alpha)'=\alpha x^{\alpha-1}$, $(a^x)'=a^x\ln a$, $(e^x)'=e^x$, $(\log_a x)'=\frac{1}{x\ln a}$, $(\ln x)'=\frac{1}{x}$,
$(\sin x)'=\cos x$, $(\cos x)'=-\sin x$, $(\operatorname{tg} x)'=\frac{1}{\cos^2 x}$, $(\operatorname{ctg} x)'=-\frac{1}{\sin^2 x}$,
$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$, $(\arccos x)'=-\frac{1}{\sqrt{1-x^2}}$, $(\operatorname{arctg} x)'=\frac{1}{1+x^2}$, $(\operatorname{arcctg} x)'=-\frac{1}{1+x^2}$,
$(\operatorname{sh} x)'=\operatorname{ch} x$, $(\operatorname{ch} x)'=-\operatorname{sh} x$, $(\operatorname{th} x)'=\frac{1}{\operatorname{ch}^2 x}$, $(\operatorname{cth} x)'=-\frac{1}{\operatorname{sh}^2 x}$
— формулы дифференцирования основных элементарных функций

$y'_x=\frac{y'(t)}{x'(t)}$ — производная функции, заданной параметрически: $x=x(t)$, $y=y(t)$

$y=f(a)+f'(a)(x-a)$ — уравнение касательной к графику функции $f(x)$ в точке $(a;\ f(a))$

Определенный интеграл

$\int dx=x+C$, $\int x^{\alpha}dx=\frac{x^{\alpha+1}}{\alpha+1}+C$ $(\alpha\neq-1)$, $\int\frac{dx}{x}=\ln|x|+C$, $\int\frac{dx}{a^2+x^2}=\frac{1}{a}\operatorname{arctg}\frac{x}{a}+C$,

$\int\frac{dx}{\sqrt{a^2-x^2}}=\arcsin\frac{x}{a}+C$, $\int a^x dx=\frac{a^x}{\ln a}+C$, $\int e^x dx=e^x+C$,

$\int\sin x\,dx=-\cos x+C$, $\int\cos x\,dx=\sin x+C$, $\int\frac{dx}{\sin^2 x}=-\operatorname{ctg}x+C$, $\int\frac{dx}{\cos^2 x}=\operatorname{tg}x+C$,

$\int\frac{dx}{x^2-a^2}=\frac{1}{2}\ln\left|\frac{x-a}{x+a}\right|+C$, $\int\frac{dx}{\sqrt{x^2\pm a^2}}=\ln|x+\sqrt{x^2\pm a^2}|+C$,

$\int\operatorname{sh}x\,dx=\operatorname{ch}x+C$, $\int\operatorname{ch}x\,dx=\operatorname{sh}x+C$, $\int\frac{dx}{\operatorname{sh}^2 x}=-\operatorname{cth}x+C$, $\int\frac{dx}{\operatorname{ch}^2 x}=\operatorname{th}x+C$

— таблица основных интегралов

$\int u\,dv=uv-\int v\,du$ — формула интегрирования по частям в неопределенном интеграле

$\int_a^b u\,dv=uv\Big|_a^b-\int_a^b v\,du$ — формула интегрирования по частям в определенном интеграле

$\int f(\varphi(t))\varphi'(t)dt=F(\varphi(t))+C$ — формула замены переменной в неопределенном интеграле; здесь $F'(x)=f(x)$

$\int_a^b f(x)dx=\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt$ — формула замены переменной в определенном интеграле; здесь $\varphi(\alpha)=a$, $\varphi(\beta)=b$

$\int_a^b f(x)\,dx=F(b)-F(a)$ — формула Ньютона—Лейбница; здесь $F(x)$ — первообразная для $f(x)$

$S=\int_a^b f(x)\,dx$ — площадь криволинейной трапеции $a\leqslant x\leqslant b$, $f(x)\geqslant 0$

$l=\int_a^b\sqrt{1+(f'(x))^2}dx$ — длина дуги кривой $y=f(x)$, где $a\leqslant x\leqslant b$

$V=\int_a^b S(x)dx$ — объем тела; здесь $S(x)$ — площадь поперечного сечения

Функции нескольких переменных

$z-z_0=f'_x(x_0,y_0)(x-x_0)+f'_y(x_0,y_0)(y-y_0)$ — уравнение касательной плоскости к поверхности $z=f(x,y)$ в точке $(x_0;y_0;z_0)$

$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x}$, $\frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y}$ — производные сложной функции $z=f(u(x,y),v(x,y))$ в точке $(x;y)$

$\frac{\partial z}{\partial l}=\frac{\partial z}{\partial x}\cos\alpha+\frac{\partial z}{\partial y}\cos\beta$ — производная функции $z=f(x, y)$ по направлению $\vec{l}=(\cos\alpha;\ \cos\beta)$

$\frac{\partial z}{\partial l}=\operatorname{grad} z\cdot\vec{l}$ — связь между производной по направлению и градиентом

$\mu=\iint\limits_D \rho(x, y)\,dx\,dy$ — масса плоской пластинки D с линейной плотностью $\rho(x, y)$

$V=\iint\limits_D f(x, y)\,dx\,dy$ — объем цилиндрического бруса, ограниченного снизу плоской областью D, а сверху — графиком функции $f(x, y)\geqslant 0$

$\iint\limits_D f(x, y)\,dx\,dy=\int\limits_a^b dx\int\limits_{g(x)}^{h(x)} f(x, y)\,dy$ — формула для вычисления двойного интеграла по области D: $a\leqslant x\leqslant b$, $g(x)\leqslant y\leqslant h(x)$

$A=\int\limits_L P(x, y)\,dx+Q(x, y)\,dy$ — работа силового поля $\vec{F}=(P(x, y),\ Q(x, y))$ вдоль дуги L

$\int\limits_L P(x, y)\,dx+Q(x, y)\,dy=\iint\limits_D\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\,dy$ — формула Грина; здесь L — контур, ограничивающий область D

Функциональные ряды

$f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+\ldots$ — ряд Тейлора для функции $f(x)$

$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\ldots\quad(-\infty<x<\infty),$

$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots+(-1)^n\frac{x^{2n}}{2n)!}+\ldots\quad(-\infty<x<\infty),$

$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}+\ldots\ (-\infty<x<\infty),$

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

$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{1\cdot 2}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot 2\cdot 3}x^3+\ldots\ (-1<x<1)$

— разложения функций $\sin x$, $\cos x$, e^x, $\ln(1+x)$, $(1+x)^\alpha$ в ряд Тейлора

$f(x)=\frac{a_0}{2}+\sum\limits_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$ — ряд Фурье для функции $f(x)$ с периодом 2π, где

$a_n=\frac{1}{\pi}\int\limits_0^{2\pi} f(x)\cos nx\,dx\ (n=0, 1, 2, \ldots),\quad b_n=\frac{1}{\pi}\int\limits_0^{2\pi} f(x)\sin nx\,dx\quad(n=1,\ 2,\ \ldots)$

Функции комплексной переменной

$r=\sqrt{a^2+b^2}$ — модуль комплексного числа $a+bi$

$\cos\varphi=a/r$, $\sin\varphi=b/r$ — определение аргумента комплексного числа

$\alpha=r(\cos\varphi+i\sin\varphi)$ — тригонометрическая форма комплексного числа

$[r(\cos\varphi+i\sin\varphi)]^n=r^n(\cos n\varphi+i\sin n\varphi)$ — формула Муавра

$\sqrt[n]{\alpha}=\sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n}+i\sin\frac{\varphi+2\pi k}{n}\right)$ $(k\in\mathbb{Z},\ n=0, 1, \ldots, n-1)$ — корень n-й степени из комплексного числа α

$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\ \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}$ — необходимые и достаточные условия дифференцируемости функции $f(z)=u(x, y)+iv(x, y)$ в точке $z=x+yi$

$e^{iz}=\cos z+i\sin z$ — формула Эйлера

$\cos z=\frac{e^{iz}+e^{-iz}}{2},\ \sin z=\frac{e^{iz}-e^{-iz}}{2i}$ — выражения для $\sin z$ и $\cos z$ в комплексной области

$\ln z=\ln|z|+i\arg z$ — выражение для $\ln z$ в комплексной области

$f(a)=\frac{1}{2\pi i}\int\limits_{\Gamma}\frac{f(z)}{z-a}dz$ — интегральная формула Коши; контур Γ ориентирован положительно относительно области D, содержащей точку a

$\mathrm{Res}(f;\, a)=\frac{1}{(m-1)!}\frac{d^{m-1}[(z-a)^m f(z)]}{dz^{m-1}}\Big|_{z=a}$ — формула для нахождения вычета относительно полюса a порядка m

$\int\limits_{\gamma}f(z)dz=2\pi i\sum\limits_{k=1}^{n}\mathrm{Res}(f;\, a_k)$ — теорема о вычетах

Дифференциальные уравнения

$y'=p(x)q(y)$ — уравнение с разделяющимися переменными

$y'+p(x)y=f(x)$ — линейное уравнение первого порядка

$y'+p(x)y=q(x)y^n$ — уравнение Бернулли

$y''+p(x)y'+q(x)y=f(x)$ — линейное уравнение второго порядка

$y''+py'+qy=0$ — линейное однородное уравнение второго порядка с постоянными коэффициентами

$x^2+px+q=0$ — характеристическое уравнение; в зависимости от его корней λ_1 и λ_2 общее решение уравнения $y''+py'+qy=0$ имеет вид:

$y=C_1e^{\lambda_1 x}+C_2e^{\lambda_2 x}$, если λ_1 и λ_2 — различные действительные корни;

$y=e^{\lambda_1 x}(C_1+C_2x)$, если λ_1 и λ_2 — равные действительные корни;

$y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x)$, если $\lambda_1=\alpha+\beta i$, $\lambda_2=\alpha-\beta i$ — комплексные корни

$y''+py'+qy=f(x)$ — линейное неоднородное уравнение второго порядка с постоянными коэффициентами; в случае $f(x)=P(x)e^{ax}$, где $P(x)$ — многочлен, его общее решение имеет вид:

$y=Q(x)e^{ax}$, где $Q(x)$ — многочлен той же степени, что и $P(x)$, если $a\neq\lambda_1$, $a\neq\lambda_2$ (λ_1 и λ_2 — корни характеристического уравнения);

$y=xQ(x)e^{ax}$, если $a=\lambda_1$, $a\neq\lambda_2$;

$y=x^2Q(x)e^{ax}$, если $a=\lambda_1=\lambda_2$

Предметный указатель

У

Учебное издание

Мордкович Александр Григорьевич
Солодовников Александр Самуилович

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Зав. редакцией Е. С. Гридасова. Редактор А. М. Суходский. Оформление художника Ю. Д. Федичкина. Художественный редактор В. И. Пономаренко. Технический редактор З. А. Муслимова. Корректор Г. И. Кострикова

ИБ № 7432

Изд. № ФМ-943 Сдано в набор 13.10.89. Подп. в печать 13.08.90. Формат $70 \times 100^1/_{16}$. Бум. тип. № 2. Гарнитура литературная. Печать офсетная. Объем 33,80 усл. печ. л. + форзац 0,16 усл. печ. л., 67,76 усл. кр.-отт., 30,26 уч.-изд. л. + форзац 0,06 уч.-изд. л. Тираж 40 000 экз. Зак. № 583. Цена 1 р. 30 коп.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Неглинная ул., д. 29/14.

Отпечатано с диапозитивов Ярославского полиграфкомбината Госкомпечати СССР. 150014, Ярославль, ул. Свободы, 97, в Московской типографии № 4 Госкомпечати СССР. 129041, Москва, Б. Переяславская ул., 46. Зак. 583.